Quantifying the Burden of Malnutrition in Children with Orofacial Clefts: SUPPLEMENTARY APPENDIX

FINAL Report

September 27, 2022

Table of Contents

Appendix 1: Methods Appendix

Section 1.1: Modeling Approaches for GBD

Overview of GBD: Dimensions and Metrics

GBD Locations

GBD 2020 produced estimates for 204 countries and territories, which includes all WHO member states that were grouped into 21 regions and seven super-regions. The super-regions[26] are Central Europe, Eastern Europe, and Central Asia; High Income; Latin America and the Caribbean; North Africa and the Middle East; South Asia; Southeast Asia, East Asia, and Oceania; and Sub-Saharan Africa. GBD completed subnational analyses for a number of countries including Brazil, China, Ethiopia, India, Indonesia, Iran, Italy, Japan, Kenya, Mexico, New Zealand, Nigeria, Norway, Pakistan, the Philippines, Poland, Russia, South Africa, Sweden, the UK, and the USA. All analyses are at the first level of administrative organization within each country except for New Zealand (by Māori ethnicity), Sweden (by Stockholm and non-Stockholm), the UK (by local government authorities), and the Philippines (by provinces). *We completed all analyses at the national level for this project*.

GBD Cause List

The GBD cause and sequelae list is organized in hierarchical nested categories designated as "levels" (Appendix Table 1). The highest level (Level 1) splits causes into broad categories: communicable, maternal, neonatal, and nutritional diseases; non-communicable diseases; and injuries. Each of these categories is then decomposed into specific causes of health loss, with finer resolution at each subsequent level. As can be seen in Appendix Table 1, orofacial clefts is a Level 4 cause in the GBD cause list, falling under congenital birth defects (Level 3), other non-communicable diseases (Level 2), and non-communicable diseases (Level 1).The cause list is mutually exclusive and collectively exhaustive.

GBD Cause Level	GBD Cause Name
1	Non-communicable diseases
2	Other-non-communicable diseases
3	Congenital birth defects
4	Neural tube defects
4	Congenital heart anomalies
4	Orofacial clefts
4	Down syndrome
4	Klinefelter syndrome
4	Turner syndrome
4	Other chromosomal anomalies
4	Congenital urogenital anomalies
4	Congenital digestive anomalies
4	Congenital musculoskeletal anomalies
4	Other congenital birth defects

Appendix Table 1: Location of orofacial clefts within the GBD cause hierarchy

GBD Risk Factors List

The GBD risk factor hierarchy contains three broad categories within its highest level: behavioral; metabolic; and environmental/occupational. This list is not collectively exhaustive, but rather contains the subset of risk factors that have been found to be significant causal factors in incident and/or fatal diseases or injuries. Stunting, wasting, and underweight are each Level 4 risks, falling under child growth failure (CGF, Level 3), child and maternal malnutrition (Level 2), and behavioral risk factors (Level 1) as shown in Appendix Table 2.

GBD Risk Level	GBD Risk Name
1	Behavioral risks
2	Child and maternal malnutrition
3	Child growth failure
4	Child underweight
4	Child wasting
4	Child stunting

Appendix Table 2: Location of child growth failure within the GBD risk hierarchy

GBD Age and Sex Groups

The GBD produces estimates for 25 age groups and two sexes (males and females). Ages are grouped into ages ranges that include 0–7 days, 7–28 days, 1–5 months, 6–11 months, 12–23 months, 2–4 years, and then 5–year age bins from ages 5–9 years through age 95+ years. The under-5 age groups are more granular to capture the rapid physical and physiological changes during this phase of life.

For certain causes and risk factors, there are only a subset of demographic groups that are most relevant for comprehensively describing their epidemiology and health burden. These are operationalized through the assignment of "age and sex restriction." For example, prostate cancer is sex-restricted to only occur in males. Our analysis applies several such restrictions to the estimations. For orofacial clefts, cause-specific mortality burden was estimated for all age groups <5 years, with nonfatal disease burden (incidence, prevalence, and YLDs) estimated from birth to 69 years. (Incidence was assigned to be zero after birth for all congenital birth defects.) For CGF (encompassing stunting, wasting, and underweight), anthropometry distributions were estimated only for children younger than 5 years, and the same restrictions were applied when estimating the attributable burden of CGF.

Overview of Cause-specific Burden Estimation

Orofacial clefts, as described above, are quantified in GBD as a source of both cause-specific mortality and disability. What follows is a brief overview of the associated modeling processes, with additional details in the "Details of GBD Cause Estimation: Orofacial Clefts" section below.

Causes of Death (deaths, years of life lost [YLLs])

The GBD estimates cause-specific deaths and years of life lost (YLLs) for each cause by location, age group, sex, and year. Within the GBD framework, each death is assigned to a single underlying cause according to certification guidelines from the WHO and International Classification of Diseases.[27] The GBD estimates mortality due to all fatal diseases and injuries in the GBD cause list using a unified approach. There are seven primary data sources used in the modeling of cause-specific mortality. These include vital registration (VR), verbal autopsy (VA), cancer registries, police records, sibling history, surveillance, survey/census, and minimally invasive tissue sample (MITS) diagnoses. Not all of these data types are relevant for every cause. For orofacial clefts, the sources with usable data include VR, a small subset of detailed VA studies, selected surveillance systems, and MITS.

Causes of Death (CoD) input data are standardized in a process that involves 1) mapping International Classification of Disease (ICD) codes to translate causes found in the input data to the GBD cause list, 2) age-sex splitting, 3) correcting for miscoding and misclassification of select causes, 4) redistributing garbage codes (deaths assigned to causes that should not be considered the underlying cause of death), 5) dropping non-representative VR and other select adjusted data, cause aggregation, noise reduction and outlier identification.

Cause-specific death rates are modeled using the cause of death ensemble modeling (CODEm) framework, which is comprised of four components. An initial set of individual models are run and ranked based on out-of-

sample predictive validity. That ranking is used to construct different weighted combinations, and the resulting ensembles and individual models are then ranked on a different set of out-of-sample data. Finally, CODEm chooses the single model with the highest rank, which is usually an ensemble model. After all cause-specific cause of death modeling is complete and there are models chosen for each cause, we run a process called CoDCorrect to ensure that the sum of all models is equal to corresponding all-cause mortality estimates. This process guarantees internal consistency across causes, in addition to confirming that the resulting models sum to the parent models (i.e., all congenital birth defects models, including cleft, sum to the inclusive congenital birth defects model). Following CoDCorrect, YLLs are computed by multiplying the number of estimated deaths by the standard life expectancy at age of death. This metric captures premature deaths by applying a larger weight to deaths that occur in younger age groups.

Nonfatal Disease Burden (incidence, prevalence, years of life lived with disability [YLDs])

The GBD estimates incidence, prevalence, and years lived with disability (YLDs) for each cause by location, age group, sex, and year. There are several primary data sources used for nonfatal burden estimation, which generally include published literature from systematic reviews, survey data, disease registries, and hospital and claims data. These data go through adjustment and standardization processes, including extracting all relevant information and metadata, splitting data points into GBD age groups and sexes (described above), and adjusting for bias and alternative case definitions or study methods (internally referred to as crosswalking). Crosswalks allow us to impute the implied value of a data point if it were to meet our case definition, thereby rendering it comparable to the other data points in the model and allowing us to incorporate as much input data as possible. The crosswalk betas used to adjust data points are generated using a subset of data to fit a model using a tool called Meta-regression—Bayesian, Regularized, Trimmed (MR-BRT)). Clinical data also go through specific adjustment processes described in further in the "Details of GBD Cause Estimation: Orofacial Clefts" section below.

After standardization, data are combined to produce estimates of prevalence and incidence for each condition. Across GBD, a variety of modeling approaches are used, with selection dependent on the unique epidemiological aspects and availability of data for each condition. For orofacial clefts, we used DisMod-MR 2.1, a meta-regression tool that uses a compartmental model to generate estimates of incidence, prevalence, remission, and mortality that are all internally consistent with one another (e.g., there cannot be more deaths than people with the condition). Whenever possible, additional information is used to inform this model. Examples for orofacial clefts include a constraint that incidence is zero after birth and utilization of results from the cause-specific mortality analysis above to inform the model. After modeling within DisMod-MR 2.1, we then estimate the distribution of severity for the sequela of causes. The number of people living with a health outcome is multiplied by a disability weight, which represents the magnitude of health loss associated with that outcome to calculate unadjusted YLDs. We then run a comorbidity correction (COMO), which is a micro-simulation that adjusts for comorbidities and computes final YLDs.

Overview of Risk Factor Burden Estimation

Child growth failure, as described above, is quantified in GBD as a risk factor. What follows is a brief overview of the risk factor modeling process, with additional information in the "Details of GBD Risk Factor Estimation: Child Growth Failure (Stunting, Wasting, Underweight)" section below.

The three components needed for risk factor estimation in GBD are exposure estimates (often expressed as prevalence but can be a distribution or a continuous value), relative risks (RR; increased likelihood of developing, or dying from, adverse health outcomes), and theoretical minimum relative exposure level (TMREL; the

exposure level at which RR changes from baseline). Risk factors in GBD can be continuous, polytomous, or dichotomous, depending upon the nature of the risk and the available data on exposure and RR. Data sources for exposure estimation include cross-sectional surveys, gray literature and reports, and published studies identified from systematic literature reviews. Data sources for relative risk estimation consist of intervention trials, cohort studies, and in some instances case-control studies. Data must come from individual-level studies, and ecological or cross-sectional observations are not included. Risk-outcome pairs are only included in GBD if they meet criteria including sufficient evidence to indicate a direct causal relationship, lack of reverse causality, and biological plausibility. TMREL values are selected for each risk factor, either via empirical assessment of the exposure for when elevated risk occurs (preferred, but not always possible) or via expert assessment of where risk would be minimal (e.g., the TMREL for air pollution is zero).

Using exposure, RR, and TMREL estimates, GBD estimates population attributable fractions (PAFs) using the following formula.

$$PAF = \frac{p(RR-1)}{p(RR-1)+1}$$

Where p is prevalence of exposure and RR is the relative risk corresponding to that level of exposure. PAFs are calculated separately for each location, age group, sex, and year for each risk-outcome pair. The PAF represents the proportion of risk that would be reduced if the exposure to a risk factor were reduced to the TMREL. After calculating a PAF for a specific risk-outcome pair, we then calculate the attributable burden of that risk factor by multiplying the PAF and the observed rates of deaths, YLLs, YLDs, or DALYs for that specific outcome. The attributable burden reflects the reduction in disease burden (YLLs, YLDs, DALYs, or deaths) for a specific outcome that would have been possible if the risk factor exposure had been at the TMREL level. Attributable burden can be summed across demographic groups and causes for an assessment of the aggregate health impact of that risk factor, but PAFs cannot be summed, and attributable burden cannot be directly summed across risk factors. To assess the aggregate impact of groups of risks, the GBD pipeline includes a simulation to account for comorbidity risk factor exposure and also reports PAFs and attributable burdens for groups of risks as outputs of this simulation.

Each of the three components of CGF (stunting, wasting, and underweight) have exposure estimates that are modeled as continuous distributions of age-sex specific Z scores. These distributions are summarized into polytomous bins of <-3 (severe), -3 to <-2 (moderate), and -2 to <-1 (mild) to reflect the available data to inform RR for specific outcomes associated with CGF. Z scores for TMREL are assigned as -1 or greater to reflect the evidence available to inform the GDB risk assessment.

Uncertainty Estimation in GBD

Throughout the GBD, uncertainty is captured in the 95% uncertainty intervals (UIs), propagated by sampling 1000 draws of the mean and standard error of the distribution and summarizing this for each quantity estimated. The UIs correspond to the 2.5th and 97.5th percentiles of the draws. The uncertainty within CoD is propagated for all demographics and includes uncertainty in both modeled cause-specific and all-cause mortality estimates. Within nonfatal modeling, the uncertainty around YLDs incorporates uncertainty associated with prevalence and disability weights.

Details of GBD Cause Estimation: Orofacial Clefts

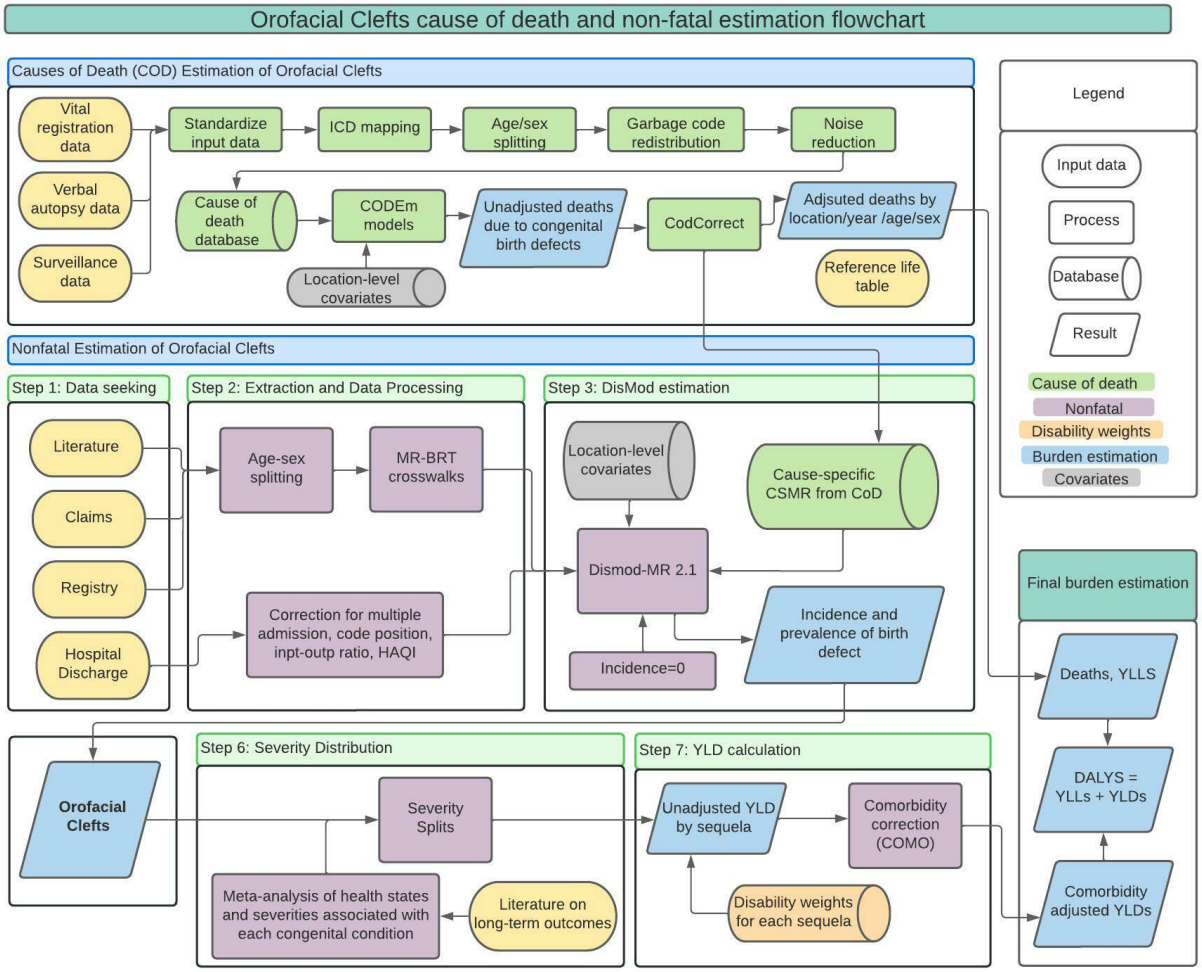
Case Definition

Orofacial clefts within the GBD include isolated cleft lip, isolated cleft palate, and combined cleft lip and cleft palate. Cleft lip is an opening in the upper lip that may extend into the nose, and cleft palate is an opening in the roof of the mouth into the nose. Both conditions are the result of the tissues of the face not joining properly during fetal development. This GBD case definition of orofacial clefts includes ICD-10 codes for isolated cleft palate Q35.2, Q35.3, Q35.5, Q35.6, Q35.7, Q35.8, and Q35.9, and ICD-10 codes Q36.0, Q36.1, Q36.9, Q37.1, Q37.5, Q37.8, and Q37.9, which correspond to cleft palate with or without cleft lip. Craniofacial clefts that do not include the oropharynx are excluded.

Orofacial clefts can be successfully treated by surgery, which is typically performed during the first few months or years of life but may occasionally be completed later in life. The GBD considers a child to remit from orofacial clefts if they undergo corrective surgery. The sequelae associated with orofacial clefts are disfigurement level 1, disfigurement level 2, and disfigurement level 2 with speech problems (see Assigning health states and sequela for YLD calculation section below for additional details).

Summary and Flowchart

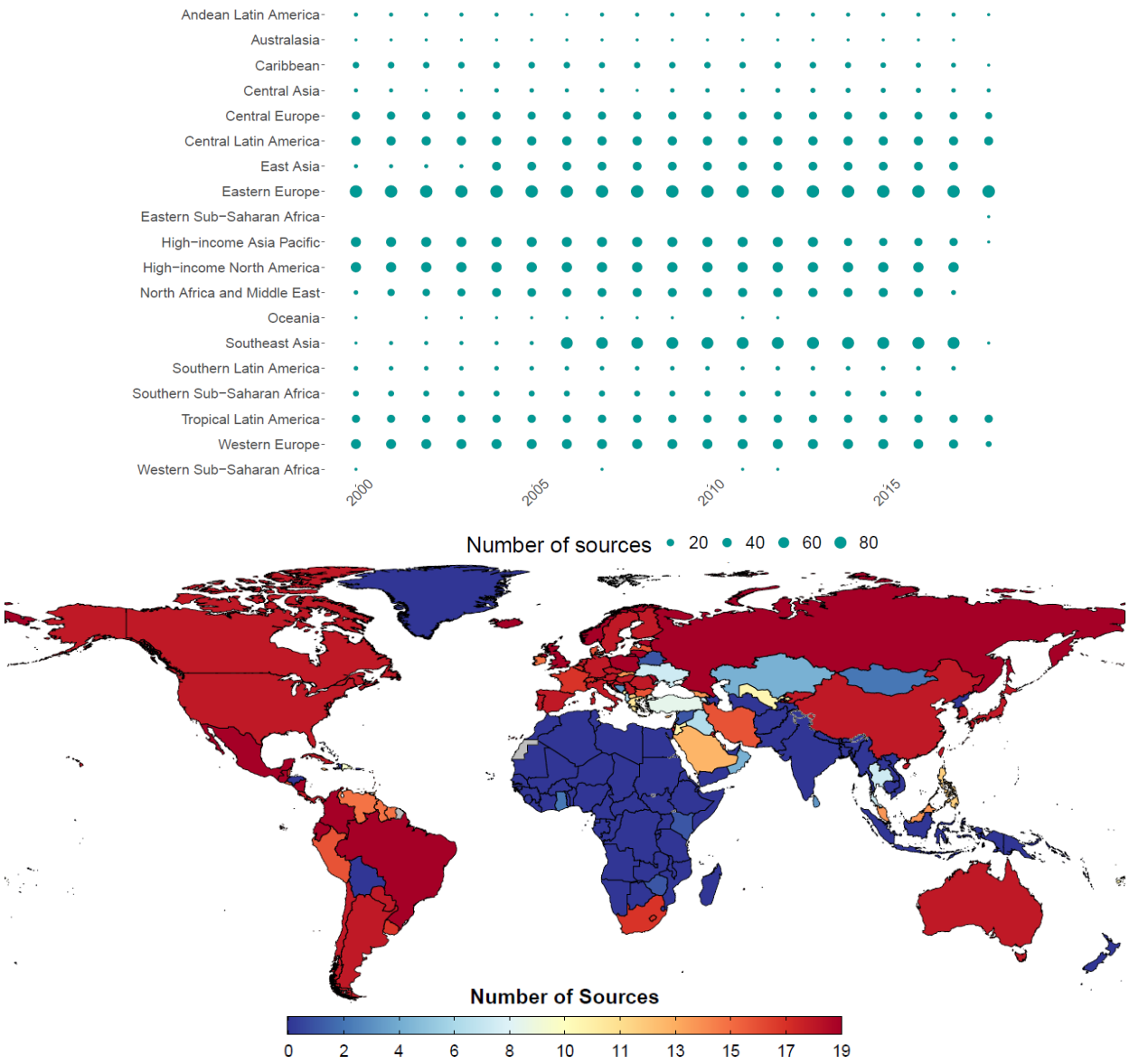
The GBD estimates both fatal and nonfatal burden due to orofacial clefts. Models of cause-specific mortality are used to inform models of prevalence and incidence, and together these measures allow us to estimate YLLs, YLDs, and DALYs due to orofacial clefts. Remission is estimated based on prevalence, incidence, and cause-specific mortality results, incorporating expert priors to guide the model fit in the absence of data. We estimate these measures for every age, location, sex, and year in the GBD, as described above. Appendix Figure 1 summarizes the orofacial cleft modeling process for both fatal and nonfatal burden, detailing the data sources used, data processing steps taken, and modeling strategies employed.



Appendix Figure 1. Analytical flowchart for the estimation of fatal and non-fatal orofacial cleft burden. Ovals represent data inputs, boxes represent analytical steps, cylinders represent databases, and parallelograms represent results.

Cause-specific Mortality Estimation

The GBD estimates cause-specific mortality rate due to orofacial clefts using CODEm for children up to 5 years of age. This model captures all deaths for which orofacial clefts are coded as the primary cause of death, and therefore does not represent all deaths in those with orofacial clefts. Data used in the modeling of orofacial clefts went through the CoD standardization process, which involves 1) mapping ICD codes to translate causes found in the input data to the GBD cause list, 2) age-sex splitting, 3) correcting for miscoding and misclassification of select causes, 4) redistributing garbage codes (deaths assigned to causes that should not be considered the underlying cause of death), 5) dropping non-representative VR and other select adjusted data, and performing cause aggregation, noise reduction, and outlier identification. As is shown in Appendix Figure 2, input data coverage for this model was high in many regions of the world, with the exception of South Asia and sub-Saharan Africa. As mentioned previously, most data incorporated in the model are VR data, with a small subset of VA and MITS data.



Appendix Figure 2. Cause-specific mortality input data by (a) region and year and (b) total number of source-years of data from 2000 to 2020 by country for orofacial clefts.

The GBD 2020 CODEm model tested several covariates, including abortion legality, antenatal care coverage proportions, indoor air pollution, skilled birth attendant coverage, average maternal education (in years), socio-demographic index (SDI), maternal alcohol consumption, healthcare access quality index (HAQI), two measures of folic acid, smoking rate, fruit and vegetable consumption, liters of alcohol consumed per capita, and fasting-plasma glucose. Many of these are population-level covariates that vary by location and year, and sometimes by age and sex. All covariates are tested for predictive validity and are selected using a linear mixed effects model. A set of component models are built with the covariates that were selected, and these models are assessed for prediction quality using root mean square error of prediction. Our final ensemble CODEm model ranks the component models based on their prediction quality and out-of-sample performance. In the model of orofacial cleft mortality used in this analysis, the covariates that explained the greatest variation and were therefore

included in the selected component models were SDI, legality of abortion, one folic acid covariate, smoking, and HAQI.

The results of this model are incorporated in nonfatal burden estimation of orofacial clefts as cause-specific mortality rate. Our compartmental nonfatal model incorporates this cause-specific mortality rate, prevalence data, and priors on incidence and remission. This allows us to model prevalence and birth prevalence with all of the information and data that we have available. This also ensures internal consistency between causes of death and nonfatal estimates for the same condition.

Nonfatal Estimation

Input Data

The GBD uses several types of data sources in the estimation of congenital anomalies: (1) prevalence and with-condition mortality rate (all cause deaths per prevalent person-year) data from literature sources, (2) birth prevalence from several international birth defects registries, (3) birth prevalence/prevalence data surveillance systems, inpatient hospital and Marketscan insurance claims data, and (4) cause-specific mortality rate estimates produced within the GBD.

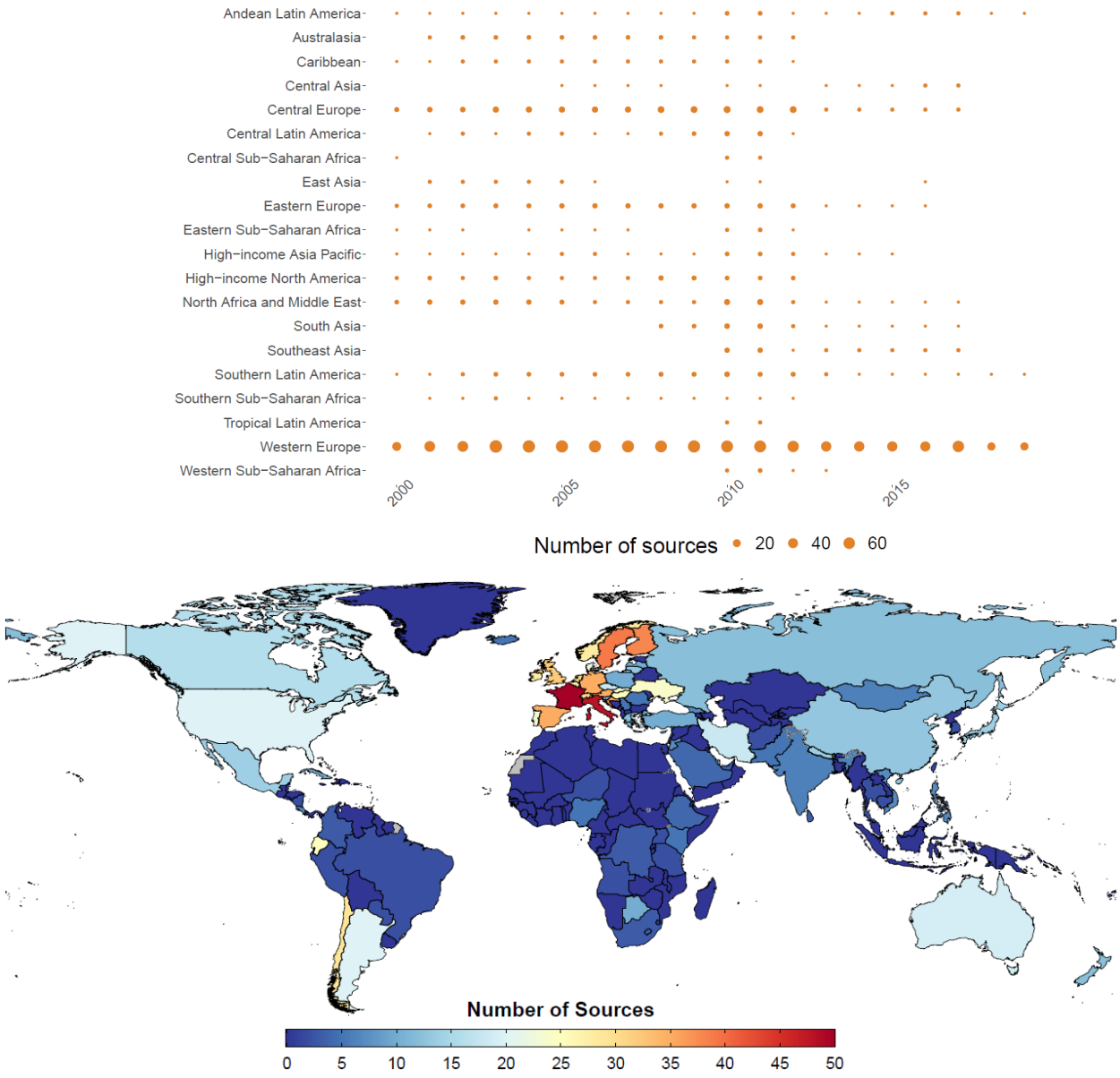
Most birth prevalence data come from international birth defects registries. The International Clearinghouse for Birth Defects Surveillance and Research (ICBDSR)[28] reports birth prevalence from international member registries. The World Atlas Report[29] also publishes birth prevalence estimates from these international registries prior to the publication of ICBDSR reports. The European Surveillance of Congenital Anomalies and Twins (EUROCAT)[30] reports the birth prevalence of anomalies for a variety of locations in Western Europe as provided by participating member registries. China's Maternal and Child Health Surveillance survey (MCHS)[31] reports birth prevalence and early neonatal mortality data for all subnational locations of China. The National Birth Defects Prevention Network (NBDPN)[32] reports birth prevalence estimates as compiled by numerous subnational registries within the United States. The Birth Defects Registry of India (BDRI)[33] reports congenital anomalies from select hospitals within India.

Inpatient hospital and claims data (from USA, Taiwan, and Singapore) used in modeling cleft are collected and corrected through a centralized process across GBD.[34] Briefly, aggregated claims data (Marketscan and Singapore, Taiwan, and Poland data) were derived from the Truven database of USA private health insurance and subset of public insurance schemes of Medicaid and Medicare for the years 2000 and 2010-2016. Inpatient hospital admissions data were extracted from 4401 location-years in 45 countries. Four rounds of data bias correction were employed in the processing of clinical data. These included 1) adjustment for readmission to account for multiple admissions for a single case of disease, 2) correction of primary diagnoses to all diagnoses to account for cases of any cause that were non-primary reasons for admission, 3) adjustment for inpatient-to-outpatient ratio to account for additional cases that did not warrant an inpatient admission, and 4) adjustment based on HAQI to account for differences in access and quality of health care across time and space.

Additionally, we performed a systematic review of the available literature for all types of congenital birth defects, which was completed in GBD 2016. The studies for inclusion were determined by constructing search strings designed to capture information on the prevalence, associated mortality, and long-term health outcomes associated with each sub-category of congenital anomalies. All results were screened—first abstracts, then full-text screenings—to ensure availability of required information, representativeness of the reported population, and exclusion of duplicate data also reported as part of the birth registry data inputs. The systematic review for orofacial clefts was performed most recently on September 13, 2016, using the following search string:

("cleft lip"[Title/Abstract] OR “cleft palate”[Title/Abstract] OR "Cleft Lip"[MeSH] OR "Cleft Palate"[MeSH]) **AND** (prevalen*[Title/Abstract] OR inciden*[Title/Abstract] OR stillbirth[Title/Abstract] OR mortality[Title/Abstract] OR survival[Title/Abstract]) **AND** ("2013/11/01"[PDAT]: "2016"[PDAT]) **NOT** (“Case Reports”[Publication Type] OR “Review”[Publication Type] OR “case report”[Title/Abstract] OR “rat”[Title/Abstract] OR “mice”[Title/Abstract] OR “mouse”[Title/Abstract]).

Appendix Figure 3 shows the dataset coverage including all data from registries, clinical administrative sources, and published studies.



Appendix Figure 3: Prevalence input data by (a) region and year and (b) total number of source-years of data by country for orofacial clefts.

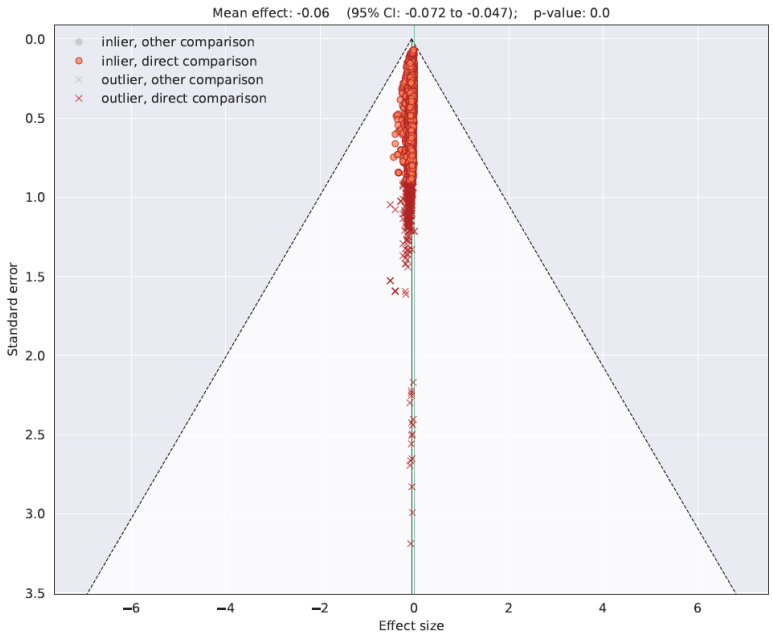
Data Processing

Any data that were not sex-specific or did not fit within GBD age-groups were age- and sex-split to fit these 10 groups prior to modeling using empirical age- and sex-patterns derived from previous DisMod-MR 2.1 models of the same condition.

Several of the input data sources used for the estimation of congenital birth defects are known to have biases leading to underreporting or overreporting relative to the true prevalence of congenital anomalies, as well as different case definitions than we use on the GBD. We applied Meta Regression – Bayesian, Regularised Trimmed (MR-BRT) to develop statistical models that were used to adjust data that had a different case definition—crosswalking. We generated crosswalks to adjust for differences in inclusion of stillbirths and chromosomal conditions, as detailed below.

Exclusion of chromosomal conditions: Some sources report birth defects in isolation (i.e., excluding any persons who have a coexisting genetic or chromosomal disorder). Our reference definition includes chromosomal diagnoses. As such, we developed a crosswalk in order to adjust data points that systematically excluded those who had a chromosomal diagnosis in addition to their orofacial cleft (or other birth defect). No splines were used in these crosswalks. The result of this crosswalk model for orofacial clefts is shown in Appendix Figure 4 and Appendix Table 3 below.

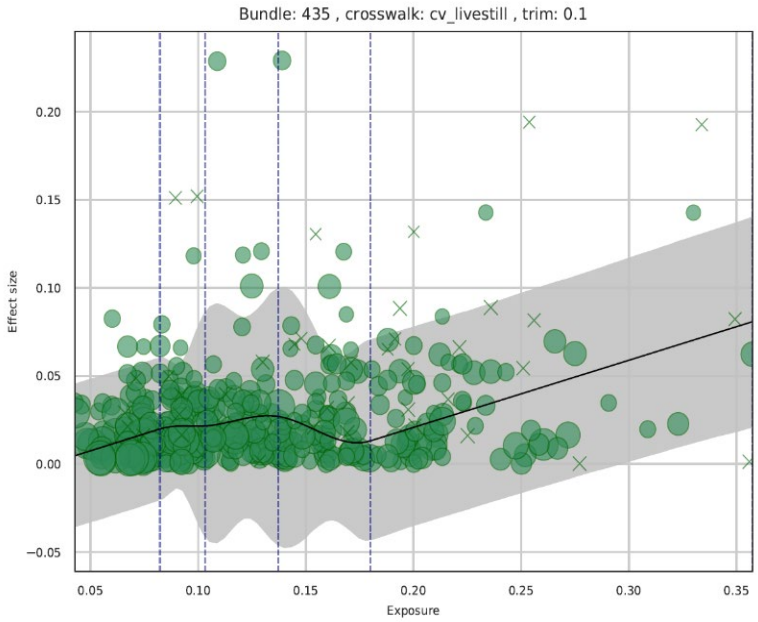
Live/Stillbirths: The GBD case definition for congenital birth defects includes only live births. As such, when a data source included stillbirths in their case definition, we used a crosswalk to adjust for the inclusion of stillbirths in the reported birth prevalence estimates in literature and registry data sources. This crosswalk essentially allowed us to impute the value of that data point in the case that it was to meet our case definition of live births only. This crosswalk used a linear covariate on log-transformed neonatal mortality rate. The result of this crosswalk model for orofacial clefts is shown in Appendix Figure 5 below.



Appendix Figure 4: Funnel plot illustrating MR-BRT meta-analyzed crosswalk result of alternate definition (chromosomal diagnoses excluded) to reference definition (chromosomal diagnoses included).

Category of Data	Beta (Standard Error)	Exponentiated Value
Including chromosomal diagnoses (reference)	Ref	—
Excluding chromosomal diagnoses adjustment (alternate)	-0.055 (0.012)	0.946 (0.012)

Appendix Table 3: MR-BRT crosswalk betas for alternate definitions (reference = livebirths including those with chromosomal anomalies) of Orofacial Clefts



Appendix Figure 5: MR-BRT crosswalk of alternate definition (livebirths and stillbirths included) with spline on log-transformed all-cause neonatal mortality rate (Orofacial Clefts)

Identifying Outliers and Data Thresholds

Underreporting of congenital birth defects is common and can vary by source, location, year, sex, and age. In order to have an empirical, systematic approach to outliering of data, we adapted the non-zero floor approach used by the GBD cause-specific mortality analysis. Briefly, after all age-sex splitting and crosswalking were completed, the first step was to calculate median absolute deviation (MAD) for the age group of birth, where registry and literature data were combined with all clinical data for the early neonatal age group (0 to 6 days) using birth prevalence data from the EUROCAT data,[30] which is generally considered our most reliable data source. The thresholds chosen were -0.5*median and +3 MAD, with any data outside of these bounds being identified as outliers. These cutoff criteria were determined based on the right skewed distribution observed in most of the congenital data and the expert prior that underreporting is far more prevalent than overreporting, making the bias asymmetric. When the lower MAD bound was negative, we used a lower bound of 0.

To evaluate data for older age groups, we employed two approaches. First, we outliered data from any location-year-source that was outliered for the first stage MAD algorithm. Second, using all clinical and literature data, we developed a model with fixed effects by age to estimate implied MAD bounds for each non-zero age group, and again applied the same thresholds of -0.5*median and +3 MAD.

Nonfatal Modeling

All available prevalence input data, along with cause-specific mortality rate (CSMR) results from the GBD cause of death (CoD) analysis of orofacial clefts, were utilized in a DisMod-MR 2.1 model in order to estimate the prevalence of orofacial clefts for each location/age/sex combination. The DisMod-MR 2.1 model of orofacial clefts had random effects on prevalence limited to ±0.8, as we expected limited variation in birth prevalence of orofacial clefts. The model settings allow increased smoothness on both excess mortality rate (EMR) and remission (maximum Xi = 5.0) in order to fit steep changes in the rates of mortality and remission during the first few years of life. Incidence was set to 0 from birth onward since orofacial clefts occur only at the time of birth, and by GBD case definition, congenital cases do not occur after birth.

When CSMR data are used as an input, DisMod-MR 2.1 pairs each CSMR datum with a matching prevalence data point by age, sex, location, and year. After matching, CSMR is divided by prevalence to calculate an implied EMR datum. All EMR data are then used in driving the model. Priors on EMR were set at a maximum of 2.5 for the early neonatal period, 0.01 for ages 5–10, and 0.000001 for ages 10+. These limits on excess mortality reflect our priors that up to 5% of individuals with orofacial clefts die in the first week of life; up to 5% die in the following three weeks; up to 20% die in the next 11 months; another maximum of 20% die before 5 years of age; and a maximum of 5% of the remaining individuals die between the ages of 5 and 10 years.

Remission was allowed in the orofacial cleft models and is an approximation cause for which surgical intervention or spontaneous remission can completely eliminate the disability due to that congenital condition. Remission was set to 0 for the first three months of life, as cleft lip and/or palate are rarely corrected during that time frame. A maximum remission rate of 0.8 per person-year was set for ages 3 months to 2 years—the age range in which cleft repair is most commonly performed—allowing up to 75% of cleft cases to be repaired. Remission was bounded from 0 to 0.07 for ages 2–5 years, 0 to 0.004 for ages 5–20 years, then bounded from 0 to 0.002 for ages 20–50 years and set at 0 for ages 50+ years. These limits on remission reflect our priors that up to 20% of remaining cleft cases are repaired between 2 and 5 years of age, while another 5% may be repaired between 5 and 20 years of age, and a maximum 5% of remaining cases are surgically repaired between ages 20 and 50 years.

Location-level covariates were used in the orofacial cleft model based on published information about the risk factors for these birth defects. A folic acid fortification covariate was used in the cleft model, which was modelled based on data from the Global Fortification Data Exchange.[35] Additionally, we incorporated covariates of the natural log of lag-distributed income per capita (LDI), a covariate measuring the folic acid intake per person per day in µg, and a covariate estimating access to care, HAQI. The HAQI and LDI covariates were used to guide the global pattern of with-condition mortality and excess mortality. Covariate effects on the model are illustrated in Appendix Table 4.

Covariate Name	Measure	Beta value	Exponentiated value
HAQI	Prevalence	-0.000097 (-0.00019 to -0.000015)	1.00 (1.00 - 1.00)
Folic acid unadjusted (ug)	Prevalence	-0.00016 (-0.00026 to -0.000071)	1.00 (1.00 - 1.00)
Composite + folic acid fortification standard	Prevalence	-0.0077 (-0.014 to -0.0015)	0.99 (0.99 — 1.00)
LN-LDI (I$ per capita)	EMR	-0.75 (-0.75 to -0.75)	0.47 (0.47 - 0.47)

Appendix Table 4. Location-level covariate effects for orofacial clefts.

Assigning Health States and Sequelae for YLD Calculation

The sequelae associated with orofacial clefts include disfigurement level 1, disfigurement level 2, and disfigurement level 2 with speech problems, which has a combined disability weight. Descriptions of these

health states can be found in Appendix Table 5 below. To determine the distribution of health outcomes associated with orofacial clefts, we performed a review of available literature on the long-term health outcomes of survivors in cohorts born with cleft abnormalities. The health states were included in the disability weight calculations that correspond to the post-surgery outcomes reported in cohorts of individuals born with these life-threatening congenital conditions. Where data were available from multiple cohorts, we pooled these cohorts to calculate the proportion of individuals with each health state. Where data on the joint distribution of the long-term health outcomes were not available, we assumed independence of each long-term health outcome. For orofacial clefts, we were able to find qualitative descriptions of the domains of disability typically associated with cleft conditions, namely disfigurement and speech problems), but no population-level assessments of the relative distributions of disability. We therefore implemented an assumption of equal split between three different sequelae. Combined disability weights were calculated for all necessary combinations of existing disability weights.

Health State Name	Health State Description	Disability Weight Mean (upper, lower bound)	Proportion
Disfigurement level 1	Has a slight, visible physical deformity that others notice, which causes some worry and discomfort	0.011 (0.071–0.006)	0.33
Disfigurement level 2	Has a visible physical deformity that causes others to stare and comment; subsequently the person is worried and has trouble sleeping and concentrating	0.067 (0.096–0.044)	0.33
Speech problems with Disfigurement level 2	Combined DW	0.115 (0.164–0.076)	0.33

Appendix Table 5. Severity splits, health states and disability weights for orofacial clefts

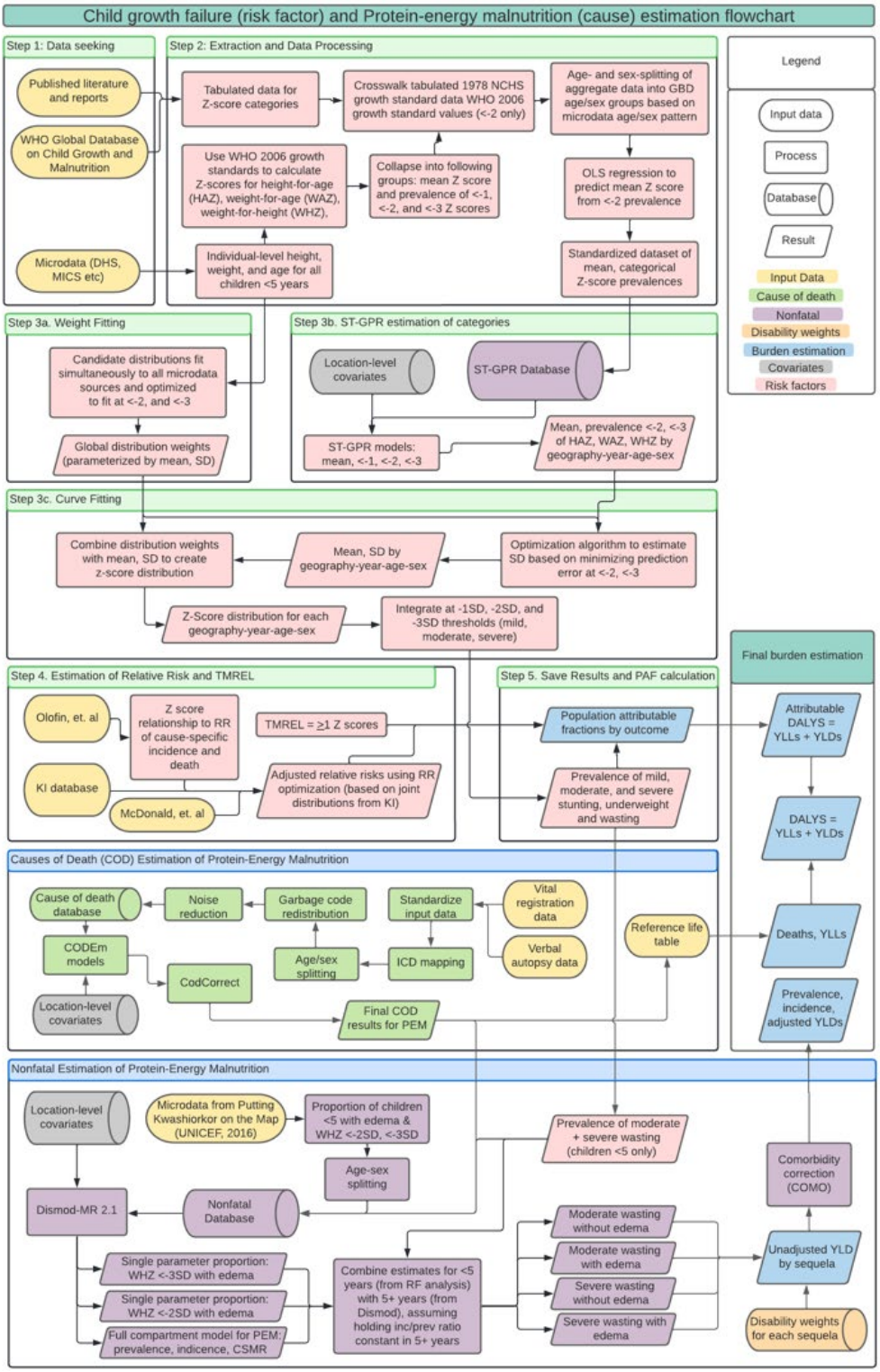
Details of GBD Risk Factor Estimation: Child Growth Failure (Stunting, Wasting, Underweight) and Protein Energy Malnutrition

Case Definition

Child growth failure (CGF) is a GBD risk factor in that it increases the risk of subsequent death and disability from other conditions. There is also a GBD cause named Protein Energy Malnutrition which represents clinical conditions such as moderate and severe acute malnutrition (MAM, SAM), marasmus, and kwashiorkor from which there is direct death and disability. The case definition of PEM is equivalent to wasting and the data used to estimate each are the same for children <5 years. CGF is estimated using three indicators—stunting, wasting, and underweight—all of which are based on categorical definitions using the WHO 2006 growth standards for children 0–59 months.[21] Mild (<-1 to -2 Z score), moderate (<-2 to -3 Z score), and severe (<3 Z score) categorical prevalence were estimated for each indicator.

Summary and Estimation Flowchart

We incorporated representative stunting, wasting, and underweight data in ST-GPR (spatiotemporal Gaussian process regression) models to estimate the prevalence of CGF below -1, -2, and -3 Z scores from the 2006 WHO Growth Standards[17] median. We then leveraged surveys that included individual-level stunting, wasting, and underweight observations to parameterize characteristic HAZ, WHZ, and WAZ curve shapes using an ensemble distribution modeling strategy. By optimizing characteristically shaped curves to align with age-, sex-, location-, and year-specific regression outputs, we estimated age- and sex-specific continuous HAZ, WHZ, and WAZ distributions. Final estimates of mild (<-1 Z score), moderate (<-2 Z score), severe (<-3 Z score) CGF reflect the prevalence of each severity level as integrated from the estimated continuous distributions.



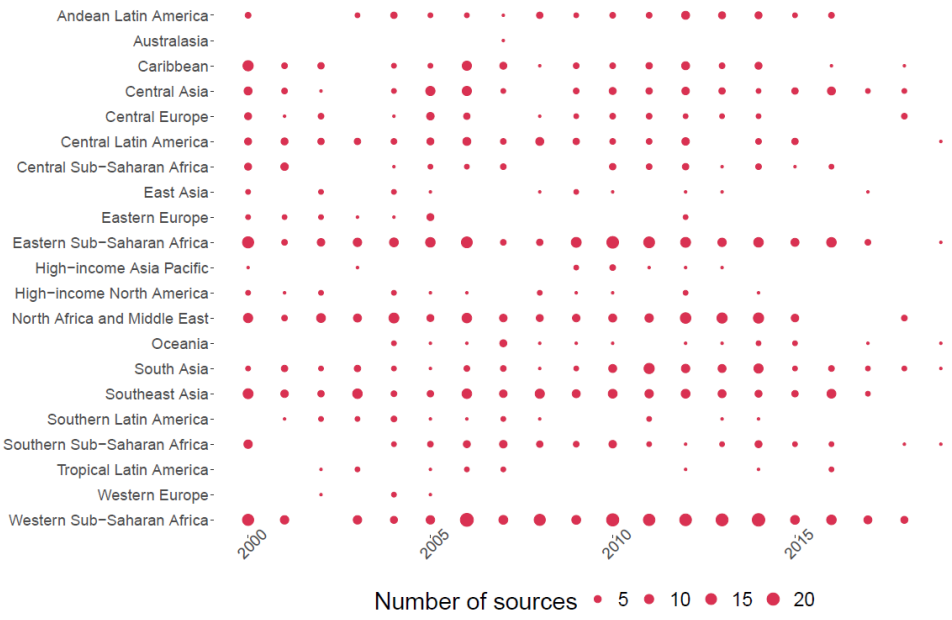
Appendix Figure 6. Analytical flowchart for the estimation of child growth failure (CGF; stunting, was5ting, and underweight) and protein-energy malnutrition (PEM) Ovals represent data inputs, boxes represent analytical steps, cylinders represent databases, and parallelograms represent results.

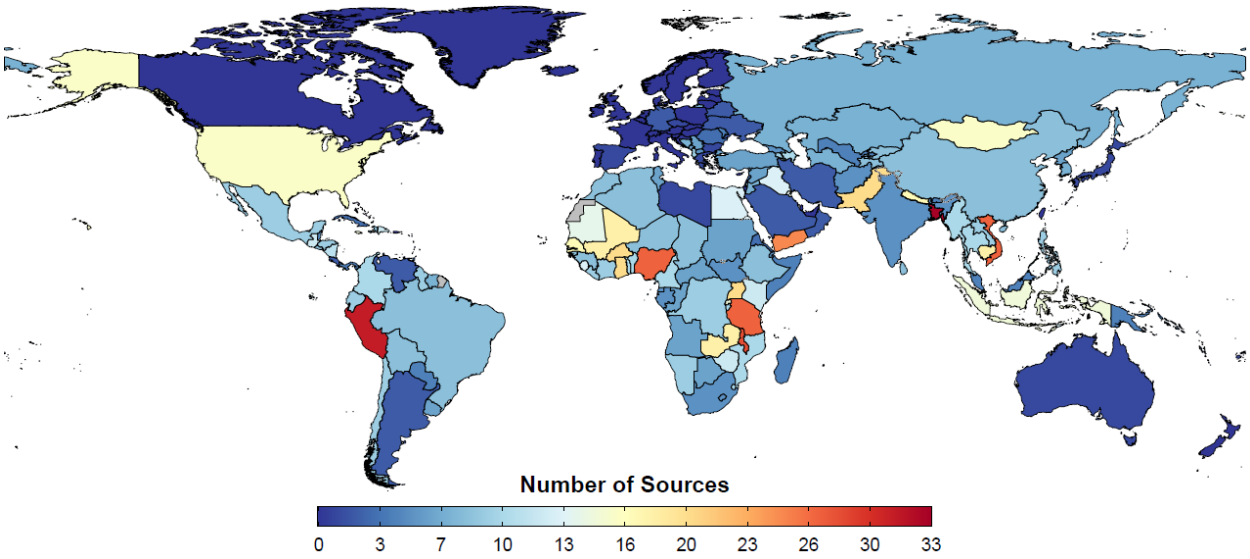
Exposure Estimation

Exposure Input Data

There are three main inputs for the GBD CGF models: microdata from population surveys, tabulated data from reports and published literature, and the WHO Global Database on Child Growth and Malnutrition.[36] The primary data additions in GBD 2020 for CGF were from population surveys that include anthropometry. Population surveys include a variety of multi-country and country-specific survey series such as Multiple Indicator Cluster Surveys (MICS), Demographic and Health Surveys (DHS), Living Standards Measurement Surveys (LSMS), and the China Health and Nutrition Survey (CHNS), as well as other one-time, country-specific surveys such as the Indonesia Family Life Survey and the Brazil National Demographic and Health Survey of Children and Women. These microdata contain information about each individual child's age, as well as their associated height and/or weight. From this information, height-for-age Z score (HAZ), weight-for-age Z score (WAZ), and weight-for-height Z score (WHZ) are calculated using the WHO 2006 Child Growth Standards[17] and the LMS method.[37] Data were dropped if found to contain invalid Z scores (HAZ, WAZ, WHZ) and/or impossible values (negative height, weight, or age), or Z scores outside the range of +6 to -6.

All available data from the WHO Global Database on Child Growth and Malnutrition[36] were extracted, much of which are from published studies. Data exclusions contained examination date prior to 1985, non-population-representative studies, and self-reported information. We looked for four metrics from all sources with tabulated data: mean Z score, prevalence <-1 Z score, prevalence <-2 Z score, and prevalence <-3 Z score. All data for each metric were extracted for each of stunting (HAZ), wasting (WHZ), and underweight (WAZ). Appendix Figure 7 below show the final input data sources for moderate underweight prevalence.





Appendix Figure 7. Exposure input data by (a) region and year and (b) total number of source-years of data by country for underweight(<-2 weight-for-age Z scores)

Exposure: Data Processing

For any data that were presented as both sexes combined or for 0–59 months combined, we used the age and sex pattern from all data sources that included that detail to split into corresponding and age- and sex-specific data. No crosswalks were performed for CGF.

Exposure: Modeling

The following four-step modelling process was applied in parallel to each of stunting, wasting, and underweight estimates. First, all microdata distributions were fitted using ensemble modelling. A series of 10 individual distributions (normal, log-normal, log-logistic, exponential, gamma, mirror gamma, inverse gamma, gumbel, mirror gumbel, and Weibull) were fitted simultaneously to each microdata source in the dataset. All component distributions that were used to derive weights were parameterized using "method of moments," meaning that each corresponding probability density function (PDF) could be described as a function of the mean and variance of the quantity of interest. From these distribution families, an ensemble distribution was parameterized. These ensemble distributions were specifically fit on the portions of the distributions that constitute mild, moderate, and severe CGF. This optimization process considers the fit across all microdata sources simultaneously. Therefore, the algorithm targets the set of ensemble weights that minimizes the predictive error across all microdata sources, collectively.

Second, we modeled prevalence of mild, moderate, and severe CGF and mean Z scores using spatiotemporal Gaussian process regression (ST-GPR), a common modelling framework used across GBD. This modeling approach generated estimates for each age group, sex, year, and location, incorporating all data sources and covariates including Socio-demographic Index (SDI), maternal care and immunization (a composite indicator of pregnancy care and vaccination coverage), Healthcare Access and Quality Index,[23] age-standardized prevalence of severe anemia, age- and sex-specific summary exposure values for unsafe sanitation, and all-age energy unadjusted (kcal/person/day available from food supply).

Third, we combined estimates of mean and prevalence (for both moderate and severe CGF) with ensemble distributions in an optimization framework to derive the variance that would best correspond to the predicted mean and prevalence values. This variance was paired with the mean. Then, using the method of moments

equation for each of the component distributions of the ensemble, PDFs of the distribution of Z scores were calculated for each location, year, age group, and sex.

Fourth, PDFs were integrated to determine the prevalence between -1 and -2 Z score (mild), between -2 and -3 Z score (moderate), and below -3 Z score (severe). These were categorical exposures used for the subsequent attributable risk analysis.

Theoretical Minimum-risk Exposure Level

Theoretical minimum risk exposure level (TMREL) is defined as the level of risk exposure that minimizes risk at the population level or at the level of risk that captures the maximum attributable burden. For underweight, stunting, and wasting, the TMREL was assigned to be greater than or equal to -1 SD of the WHO 2006[17] standard weight-for-age, height-for-age, and weight-for-height curves, respectively. Relative risks for mild, moderate, and severe CGF are therefore assessed in comparison to risk levels experienced by children above -1 SD.

Relative Risk: Input Data and Modeling

There is a high degree of correlation between stunting, wasting, and underweight. Failing to account for their covariance and assuming independence would overestimate the total burden significantly and misrepresent the attributable burden of individual CGF indicators. Data to inform the univariate relative risks (RRs) associated with each type of CGF were derived from a pooled analysis of 10 prospective cohort studies by Olofin and colleagues.[18] Next, to account for the differing effects of indicators on the same outcomes, we used a constrained optimization method to adjust the observed univariate RRs data from a pooled analysis of all-cause mortality studies by McDonald and colleagues that evaluated overlapping categories of stunting, wasting, and underweight.[19]

The process of adjusting univariate RRs had four steps. First, we created a joint distribution of stunting, underweight, and wasting from populations of children derived from 15 longitudinal studies shown in Appendix Table 8 which were accessed from the Bill and Melinda Gates Foundation Knowledge integration (Ki) database.[20]

Study Name	Country Code	Sample Size	Years Conducted
Zimbabwe Vitamin A for Mothers and Babies Trial	ZWE	14,110	1997-2001
CMC Vellore Birth Cohort Study	IND	373	2002-2006
International Lipid-Based Nutrient Supplements Project	MWI	1,206	2011-2014
Malnutrition and Enteric Disease Study	BGD	265	2009-2017
Malnutrition and Enteric Disease Study	IND	251	2009-2017
Malnutrition and Enteric Disease Study	NEP	240	2009-2017
Malnutrition and Enteric Disease Study	PER	303	2009-2017
Malnutrition and Enteric Disease Study	BRA	233	2009-2017
Malnutrition and Enteric Disease Study	ZAF	314	2009-2017
Malnutrition and Enteric Disease Study	TZA	262	2009-2017
Medical Research Council Keneba	GMB	2,867	—
Performance of Rotavirus and Oral Polio Vaccines In Developing Countries	BGD	700	2011-2014
Community-based Intervention Trial to Compare the Impact of Preventive and Therapeutic Zinc Supplementation Programs Among Young Children in Burkina Faso	BFA	7,634	2010-2012
WASH Benefits Bangladesh	BGD	4,423	2011-2014
WASH Benefits Kenya	KEN	5,649	2012-2016
Promotion of Breastfeeding Intervention Trial	BLR	16,897	1996-1998
Childhood Malnutrition and Infection Network	BGD	477	1993-1996
Childhood Malnutrition and Infection Network	BRA	119	1989-1998
Childhood Malnutrition and Infection Network	GNB	350	1987-1990
Childhood Malnutrition and Infection Network	GNB	885	1996-1997

Childhood Malnutrition and Infection Network	PER	210	1989-1991
Childhood Malnutrition and Infection Network	PER	224	1995-1998
Delhi Infant Vitamin D Study	IND	2,100	2007-2010
Characterization of Respiratory pathogens endemic to Pakistan in pregnant women and newborns in urban settings	PAK	380	2012-2013
Impact of Zinc Supplementation in Low Birth Weight Infants on Severe Morbidity, Mortality and Zinc Status: A Randomized Controlled Trial	IND	2,052	2005-2007
A Trial of Zinc and Micronutrients in Tanzanian Children	TZA	2,400	2007-2012

Appendix Table 8: Bill and Melinda Gates Foundation Knowledge Integration (KI) database study details

Second, we generated one thousand Relative Risk (RR) draws for each univariate indicator and severity based upon the cause-specific RRs and their deviations from Olofin et al.[18] Third, we altered these univariate RRs for the three causes (diarrhea, LRI, and measles) based upon interactions among the CGF indicators. An interaction occurs when the effect of one CGF indicator variable (e.g., stunting) has a different effect on the outcome depending on the value of another CGF indicator variable (e.g., underweight). Interaction terms alter the risk of the outcome among children with more than one indicator of CGF. These interaction terms were extracted from a pooled cohort analysis of all-cause mortality published by McDonald et al.[19] Fourth, we optimized the adjusted relative risks by minimizing the error between the observed RRs (generated from Olofin et al.[18]) and the altered RRs derived from the joint distribution and accounting for the interaction terms. The final list of outcomes paired with CGF risks included lower respiratory infections (LRI), diarrhea, measles, and protein-energy malnutrition, as shown in Appendix Table 9. There were no specific relative risks derived for PEM because this condition is definitionally equivalent to wasting and underweight. Instead, 100% of deaths and years lived with disability (YLDs) for PEM were assigned as attributable to each of wasting and underweight.

Cause	CGF Type	28d-5m			6m-11m		
		<-3	-3,-2	-2,-1	<-3	-3,-2	-2,-1
Diarrhea	HAZ	3.6 (2.6, 4.8)	1.7 (1.4, 2.1)	1.3 (1.1, 1.6)	2.8 (2.1, 3.9)	1.5 (1.2, 1.8)	1.2 (1.1, 1.4)
	WAZ	9.5 (6.9, 13.7)	2.5 (1.7, 3.6)	1.6 (1.2, 2.2)	9.9 (6.9, 14.3)	2.6 (1.7, 3.7)	1.6 (1.2, 2.3)
	WHZ	12.7 (8.4, 19.5)	3.5 (2.3, 5.2)	1.6 (1.2, 2.3)	12.5 (8.3, 18.4)	3.5 (2.3, 5)	1.6 (1.2, 2.2)
LRI	HAZ	4.5 (2.9, 7.0)	1.8 (1.3, 2.6)	1.4 (1.0, 1.9)	3.3 (2.1, 5.5)	1.5 (1.2, 2.2)	1.2 (1.0, 1.7)
	WAZ	8.8 (5.5, 14.6)	2.8 (1.8, 4.5)	1.8 (1.1, 2.7)	9.0 (5.3, 15.7)	2.9 (1.7, 4.8)	1.8 (1.1, 2.8)
	WHZ	10.1 (5.5, 18.4)	4.8 (2.8, 8.2)	2.0 (1.2, 3.2)	10.0 (5.5, 17.1)	4.8 (2.8, 7.8)	2.0 (1.2, 3.0)
Measles	HAZ	5.0 (2.7, 10.3)	2.4 (1.4, 4.9)	1.2 (0.6, 2.3)	4.3 (2.1, 9.6)	2.2 (1.3, 4.7)	1.2 (0.6, 2.2)
	WAZ	7.4 (4.0, 13.9)	3.1 (1.7, 5.5)	1.1 (0.5, 1.9)	7.4 (3.6, 15.0)	3.1 (1.6, 5.8)	1.1 (0.5, 2.0)
	WHZ	10.9 (4.9, 23.8)	2.9 (1.3, 6.0)	1.1 (0.6, 2.1)	10.5 (4.7, 21.5)	2.8 (1.3, 5.7)	1.1 (0.6, 2.0)
Cause	CGF Type	12m-23m			2y-4y		
		<-3	-3,-2	-2,-1	<-3	-3,-2	-2,-1
Diarrhea	HAZ	2.5 (1.9, 3.4)	1.4 (1.2, 1.6)	1.2 (1.1, 1.3)	2.5 (1.9, 3.4)	1.4 (1.2, 1.6)	1.2 (1.1, 1.3)
	WAZ	9.8 (6.9, 14.1)	2.6 (1.7, 3.7)	1.6 (1.2, 2.2)	9.9 (7, 14.1)	2.6 (1.7, 3.7)	1.6 (1.2, 2.2)
	WHZ	12.5 (8.5, 18.1)	3.5 (2.4, 4.9)	1.6 (1.2, 2.2)	12.6 (8.5, 18.1)	3.5 (2.4, 4.9)	1.6 (1.2, 2.2)
LRI	HAZ	2.7 (1.7, 4.4)	1.4 (1.1, 1.9)	1.2 (1.0, 1.5)	2.7 (1.8, 4.5)	1.4 (1.1, 1.9)	1.2 (1.0, 1.5)

	WAZ	9.2 (5.2, 16)	2.9 (1.7, 5.0)	1.8 (1.1, 2.9)	9.0 (5.2, 15.8)	2.9 (1.7, 4.8)	1.8 (1.1, 2.8)
	WHZ	10 (5.4, 17.2)	4.8 (2.8, 7.7)	2.0 (1.2, 3.0)	10.0 (5.3, 17.4)	4.8 (2.7, 7.8)	2.0 (1.2, 3.0)
Measles	HAZ	3.7 (2.0, 8.0)	2.0 (1.3, 3.5)	1.1 (0.6, 1.8)	3.7 (2.0, 8.0)	2.0 (1.2, 3.8)	1.1 (0.6, 1.9)
	WAZ	7.2 (3.7, 14.5)	3.0 (1.6, 5.5)	1.1 (0.5, 2.0)	7.2 (3.5, 14.5)	3.0 (1.6, 5.6)	1.1 (0.5, 2.0)
	WHZ	10.5 (4.8, 21.3)	2.8 (1.3, 5.6)	1.1 (0.6, 2.0)	10.5 (4.7, 21.8)	2.8 (1.3, 5.7)	1.1 (0.6, 2.0)

Appendix Table 9: Age-Specific Adjusted RRs for each risk-outcome pair for child growth failure

Protein-Energy Malnutrition Mortality and Nonfatal Estimation

CoD models were developed to estimate the mortality caused by PEM using VR and VA data. These data go through an adjustment and standardization process that is the same as what is described for orofacial clefts above and then were modeled in CODEm. PEM nonfatal burden estimation incorporates the moderate wasting estimates from the risk factor exposure analysis, combining them with cause-specific mortality rates to quantify incidence and prevalence of PEM in four mutually exclusive and collectively exhaustive categories. These categories reflect distinct gradations of disability that occur: moderate wasting *without edema* (WHZ < -2 SD to < -3 SD), moderate wasting *with edema* (WHZ < -2 SD to < -3 SD), severe wasting *without edema* (WHZ < -3 SD), and severe wasting *with edema* (WHZ < -3 SD). The aggregate of categories that include edema can be considered equivalent to the disease state commonly referred to as kwashiorkor, while severe wasting can be considered equivalent to marasmus. The split between moderate and severe wasting was informed by the relative rates in the risk factor exposure analysis. Data sources used to inform the proportion of children younger than 5 years with edema came from the compiled survey dataset associated with the 2016 "Putting Kwashiorkor on the Map"[38] technical brief. Data were adjusted and standardized, including splitting data points into GBD age groups and sexes, and modeled for all locations using DisMod-MR 2.1.

Population attributable fraction and Attributable burden

Using exposure, RR, and TMREL estimates, GBD estimates population attributable fractions (PAFs) using the following formula.

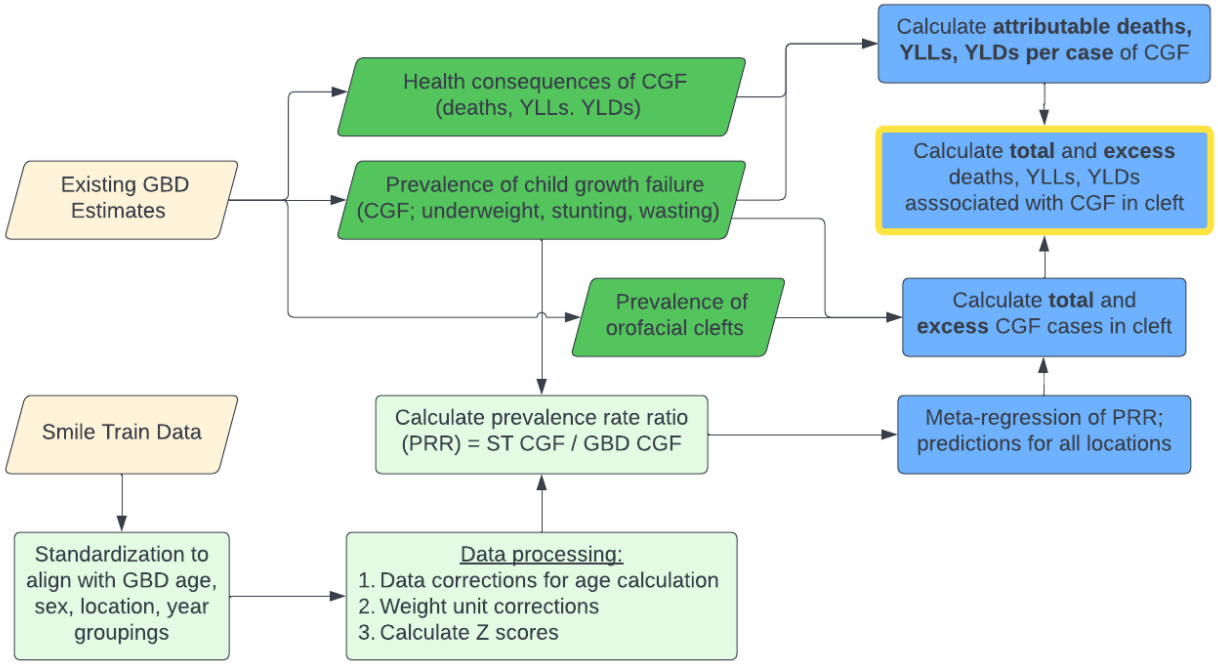
$$PAF = \frac{p(RR-1)}{p(RR-1)+1}$$

Each of the PAF values are specific to a given risk-outcome pair (e.g., underweight and diarrhea). PAF values were then multiplied by the estimated rates of specific outcomes (e.g., diarrhea deaths by age, sex, location, and year) to calculate attributable burden (e.g., diarrhea deaths *attributable to* underweight). PAF values were combined with estimates for diarrhea, LRI, and measles for each of deaths, YLLs, and YLDs to calculate attributable burden. As mentioned above, 100% of PEM deaths, YLLs, and YLDs were assigned as attributable to each of underweight and wasting. Aggregate burden attributable to each of the risk factors was quantified by summing across causes and demographic groups.

Section 1.2: Modeling Approach to Estimate the Burden of Malnutrition in Orofacial Clefts using Smile Train Database

Summary and Flowchart

We extracted, processed, and standardized cleft patient data from Smile Train Express. These data included weight and age observations for the past 20 years, and height since 2021. Because of the long time series weight data, we concentrated only on underweight condition for this analysis. Appendix Figure 8 illustrates our approach to estimation.



Appendix Figure 8. Estimation flowchart for quantifying the burden of malnutrition in *children with orofacial clefts.*

We modified Smile Train data as necessary to align with IHME's GBD modeling demographic groups, which included assigning ages to GBD age ranges, designating sex as male or female, matching location information to GBD regions, and binning according to year. We retained only individuals coming for primary cleft repair based on an assumption that these children would represent the baseline nutritional status of unremitted clefts in each location. We then completed a series of data correction steps to adjust dates (used to calculate ages) and implausible weight values (likely data entry errors). We calculated Z scores using the 2006 WHO Growth Standards[17] and applied the same +6 and -6 Z score restriction for inclusion as was used in GBD estimation. We paired calculated underweight rates with corresponding GBD underweight rates from matching year, location, age group, and sex to calculate prevalence rate ratio (PRR). This was followed by a Bayesian meta-regression model of PRR to predict PRR for each country. These were then paired with GBD 2020 estimates of orofacial clefts, underweight prevalence, and attributable burden of underweight to calculate the following quantities:

1. **Comparative risk of malnutrition in those with clefts:** We calculated the prevalence rate ratio (PRR) by dividing the rate of underweight condition in the under-5 population with orofacial clefts by underweight rate in the entire under-5 population.
2. **Total and excess underweight in those with clefts:** We compared the total number of children with malnutrition and the number with **both** cleft and malnutrition to calculate total and excess malnutrition in children with orofacial clefts. We used this proportion, as well as the PRR above, to calculate the number of "excess" cases of malnutrition in those with clefts.

3. **Total and excess deaths, years of life lost (YLLs), and years lived with disability (YLDs) related to malnutrition in those with clefts:** We evaluated the health consequences of malnutrition in those with clefts by leveraging GBD-analyzed relationships between malnutrition and subsequent malnutrition-related illness and death. This was completed by calculating the number of deaths/YLLs/YLDs per case of underweight by location, year, age, and sex then multiplying those by the estimates of total and excess malnutrition in cleft from above. This allowed us to estimate the health consequences of malnutrition specifically in the cleft population.

The outputs of this work are estimates of the burden of underweight condition in children younger than 5 years with orofacial clefts. These estimates encompass relative rates of underweight in those with clefts, total children with underweight status and cleft occurrence, excess malnutrition cases in those with clefts, and the associated malnutrition-related consequences including deaths and disease burden. These results are detailed by country, age, sex, and year. We also provide the extracted and harmonized datasets, written reports, analytic code, and summary tables and figures for use by Smile Train in their 2022 State of the World's Cleft Care.

Smile Train Surgical Database

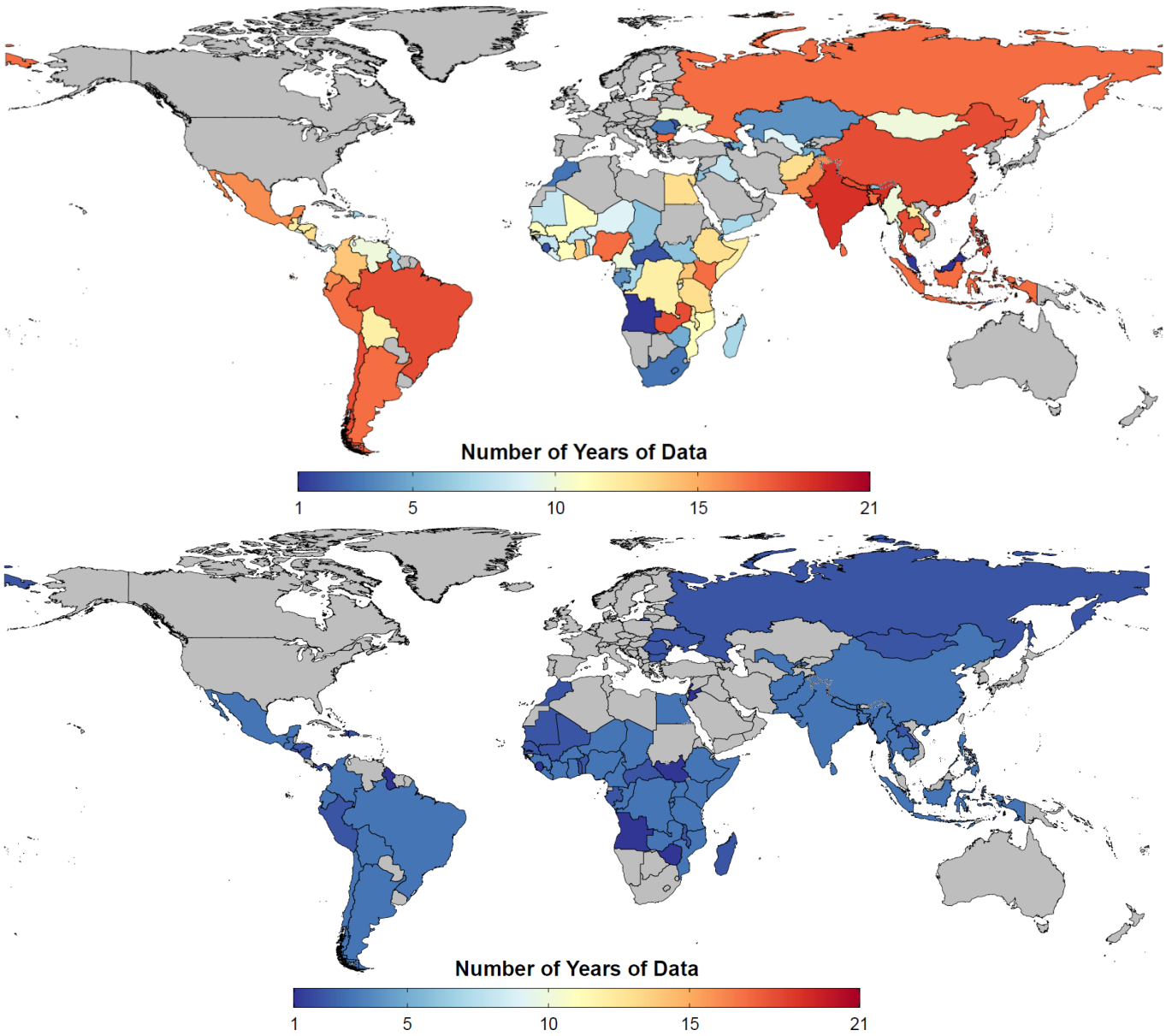
Smile Train Express is a digital database created in 2001 to allow partners all over the world to upload patient records. The database was designed as a forward-thinking approach to storing electronic data—in part to sustain programs in lower- and middle-income countries with limited resources who might face challenges with the storing, shipping, and handling of paper records. It also provides a mechanism of accountability to Smile Train sponsors, enabling traceable payments to partners who provide cleft treatment. Another benefit of this database is the ability to utilize medical-quality validations to review and monitor the surgical work of Smile Train partners. Before and after photos are examined to ensure a standard of quality is met across all treatment sites. Furthermore, analyses of available records allow Smile Train to assess program needs across the globe.

What follows is a description of our data processing and modeling approach using Smile Train data on surgical encounters and anthropometrics. Representative visuals of the effects of different data processing steps are illustrated using an example from Brazil. Data from all countries are shown in the figures of Appendix 4.

Data extraction and processing

Step 1: Data Extraction and Cleaning

We imported Smile Train's patient dataset and surgical dataset using R statistical software (version 4.1.3), which was used for all subsequent analyses. The patient dataset contains patient date of birth and anthropometric data from patients' evaluation encounters, while the surgical dataset contains anthropometric data from patients' surgical encounters and treatment details. We merged these two datasets on patient ID and extracted the following variables: unique patient ID, patient ID, treatment ID, date of birth, upload date, sex, country, evaluation weight, evaluation height, evaluation date, admission weight, admission height, admission date, and operation date. We parsed the sex and country variables to match GBD sex and location standards. Finally, we dropped Vietnam data because of a systematic error, where patients' heights and weights were from birth instead of evaluation or admission (personal communication with Smile Train). Appendix Figure 9 maps the number of country-years of weight and height data, respectively, within the Smile Train database. Histograms of the number of observations from primary surgical encounters per country and year are shown in Appendix 4. Figure 0.



Appendix Figure 9: Map of total number of country-years of weight (top) and height (bottom) data in Smile Train patient database.

Step 2: Age Analysis

Age assignment was based on three dates, evaluation, admission, and surgery, with age initially calculated as days between date of birth and each of the three extracted dates. For each patient, we assigned GBD age group IDs and years for all three of the initial age calculations. We conducted an analysis of the dates by calculating the following descriptive variables: days between evaluation and admission, admission and operation, and operation and upload. The results of the date analysis were later used to determine which version of the age calculation to extract for child growth Z score values. Appendix Figure 10 illustrates, for each location, the comparison between each of the three ages and the evaluation weight and height entries, as well as a comparison of evaluation and admission weight and height.

Step 3: Identifying Primary and Subsequent Surgical Encounters

We identified three sources of duplicated patients in the dataset: patients who received multiple operations across a time span (different treatment ID and different surgery date), patients who received multiple treatments on the same day (different treatment ID and same surgery date), and exact duplicate records (all fields identical). The third category were suspected to be the result of data input errors, so these duplicates were dropped. For the others, we completed a series of steps to identify primary and subsequent surgical encounters. All non-dropped data were processed together, but only primary encounters were retained for merging with GBD results and calculating attributable malnutrition burden in cleft.

Starting in 2015, Smile Train partners began routinely measuring height and weight upon admission for surgery. Prior to that, height and weight measurement were only reliably completed at the time of initial evaluation. For initial/ primary operation, this was not an issue, but because non-primary operations could have happened any number of days, months, or years after the initial evaluation, measurements taken at time of evaluation would not be a reliable assessment of a child's anthropometric status at the time of surgery. Therefore, these observations were dropped. With regards to subsequent surgeries involving patients with multiple treatments on the same day, it is possible for patients to have undergone multiple surgeries on the same day or for the surgeon to have decided that parts of the operation were distinct enough to be coded as separate treatments (e.g., rhinoplasty and cleft palate repair). We determined that if a subsequent surgery described malnutrition for a different age group or year other than the primary surgery, it would be kept. However, all same-day surgeries were subsetted down to one to reflect a single occurrence. We identified a patient's primary surgery as the surgical encounter with the earliest upload date and secondary surgeries as the next earliest upload date. The same logic was applied to all subsequent surgeries. For patient surgeries with the same upload date, we sequenced them according to operation date. We flagged the remaining surgeries occurring on the same day as same-day treatments. These same-day, repeat surgeries were collapsed to one encounter and a single observation of weight, height, and age. Appendix Figures 11 and 12 illustrate, for each location, the number of primary and subsequent surgical encounters.

Step 4: Date Corrections

Dates in the data were ordered in this sequence: evaluation date, admission date, operation date, and upload date. Appendix Table 6 displays summary statistics for the gap (measured in days) between these different dates both before and after date corrections. Prior to correction, the first quartile of days between evaluation and admission yielded negative values, yet there are no scenarios where evaluation occurs after admission. Thus, we determined that there are errors in these dates. Per our conversations with Smile Train colleagues, we came to understand that evaluation, admission, and operation date were entered by hand, allowing for errors. Because upload date is a digital signature and inherently error-free, it served as a guide for us to correct errors in the other three dates. We outliered operation dates occurring after the upload. In the other direction, we outliered operation dates that were too many days before upload. Then we applied a correction approach that involved calculating the median and median absolute deviation (MAD) of the days between each date in a stepwise fashion. First, when any individual record had an operation date that was further away from upload date than median plus two-times-MAD, we shifted operation date, admission date, and evaluation date together so that the gap between operation and upload date was equal to the median. We then repeated this process with the gap between admission and operation date. Finally, we repeated the correction for days between evaluation and admission date; however, because original evaluation data were regularly copied forward for subsequent surgical encounters, the corrections were only applied to primary surgeries. We therefore limited our analysis to

only data from primary surgery in subsequent steps. Appendix Figure 13 illustrates the cumulative effect of date corrections by comparing raw and corrected dates.

	1th	5th	25th	50th	75th	95th	99th	MAD	Mean	Min	Max
All Surgeries: Days Between Evaluation and Admission	−78	−22	−1	0	0	37	833	0	29.94	−2189000	36810
All Surgeries: Days Between Admission and Operation	0	0	1	1	2	6	12	1.483	2.114	−74	3971
All Surgeries: Days Between Operation and Upload	0	1	7	16	31	99	248	16.31	29.85	−3286	4448
Primary Surgeries: Days Between Evaluation and Admission	−81	−23	−1	0	0	19	365	0	15.26	−2189000	36810
Primary Surgeries: Days Between Admission and Operation	0	0	1	1	2	6	12	1.483	2.128	−74	3971
Primary Surgeries: Days Between Operation and Upload	0	1	7	16	32	100	248	16.31	30.11	−3286	4448
Subsequent Surgeries: Days Between Evaluation and Admission	−45	−10	0	0	5	973.1	3496	0	175.1	−10970	14640
Subsequent Surgeries: Days Between Admission and Operation	0	0	1	1	2	5	10	0	1.981	−27	2227
Subsequent Surgeries: Days Between Operation and Upload	0	1	6	14	28	90	250	14.83	27.37	−358	3818

	1th	5th	25th	50th	75th	95th	99th	MAD	Mean	Min	Max
All Surgeries: Days Between Evaluation and Admission	−1	0	0	0	0	0	362	0	16.08	−10970	14640
All Surgeries: Days Between Admission and Operation	0	0	1	1	1	3	3	0	1.113	0	3
All Surgeries: Days Between Operation and Upload	0	1	7	16	21	38	46	11.86	15.45	0	48
Primary Surgeries: Days Between Evaluation and Admission	0	0	0	0	0	0	0	0	0	0	0
Primary Surgeries: Days Between Admission and Operation	0	0	1	1	1	3	3	0	1.121	0	3
Primary Surgeries: Days Between Operation and Upload	0	1	7	16	21	38	46	11.86	15.55	0	48
Subsequent Surgeries: Days Between Evaluation and Admission	−45	−10	0	0	5	973.1	3496	0	175.1	−10970	14640
Subsequent Surgeries: Days Between Admission and Operation	0	0	1	1	1	3	3	0	1.036	0	3
Subsequent Surgeries: Days Between Operation and Upload	0	1	6	14	20	36	45	10.38	14.41	0	48

Appendix Table 6: Percentiles of Days Between Dates pre-correction (top) and post-correction (bottom)

Step 5: Weight Imputation and Weight Corrections

For primary surgeries, heights and weights at admission are measured on the same-day as or copied from the evaluation encounter. We imputed primary surgeries' missing evaluation heights and weights with admission heights and weights when available, and vice versa. Appendix Table 7 displays missingness of heights and weights in primary surgeries and Appendix Figure 14 illustrates each of corrected evaluation age, admission age, and operation age versus weight and height.

	Percent	Total
Only Evaluation Weight Missing	31.17	485845
Only Admission Weight Missing	0	2
Both Evaluation and Admission Weight Missing	0.02	298
Only Evaluation Height Missing	0.04	573
Only Admission Height Missing	16.21	252627
Both Evaluation and Admission Height Missing	77.8	1212618

Appendix Table 7: Missingness of heights and weights in primary surgeries

After addressing missingness, we applied a set of criteria illustrated in Appendix Table 8 to address apparent decimal errors in patient weights. The top panel lists the criteria by age and the bottom panel shows the number of observations that were affected by each adjustment. The cutoffs for action were defined by the 0th percentile and 100th percentile of age-specific WHO growth charts.[17] For example, if a weight appeared in the dataset as 1500 units and that patient was under 1 year of age, the value was assumed to be miscoded in grams and therefore was converted to kilograms by dividing by 1000. If a patient was 4 to 5 years of age and their weight appeared in the dataset as 0.4 units, that weight was assumed to have a misplaced decimal and was corrected by multiplying by 10. Appendix Figure 15 illustrates age compared to weight and height after weight corrections.

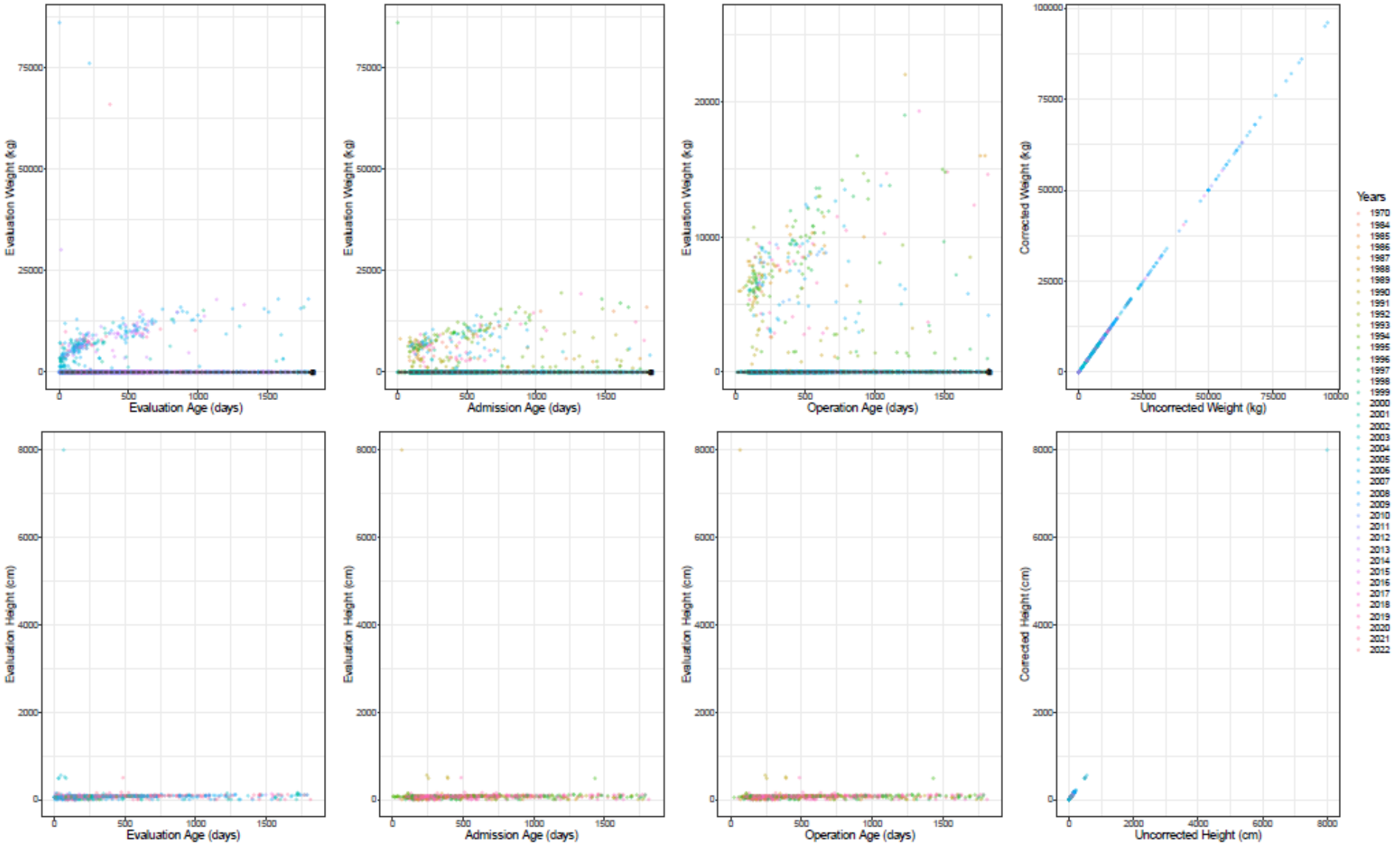
Action Taken	Under 1 year	12 to 18 months	18 to 24 months	2 to 3 years	3 to 4 years	4 to 5 years
Multiply by 1000	< 0.014	< 0.016	< 0.018	< 0.022	< 0.026	< 0.03
Multiply by 100	0.014 to 0.14	0.016 to 0.16	0.018 to 0.18	0.022 to 0.22	0.026 to 0.26	0.03 to 0.3
Multiply by 10	0.14 to 1.4	0.16 to 1.6	0.18 to 1.8	0.22 to 2.2	0.26 to 2.6	0.3 to 3.0
No Action Needed	1.4 to 14	1.6 to 16	1.8 to 18	2.2 to 22	2.6 to 26	3.0 to 30
Divide by 10	14 to 140	16 to 160	18 to 180	22 to 220	26 to 260	30 to 300
Divide by 100	140 to 1400	160 to 1600	180 to 1800	220 to 2200	260 to 2600	300 to 3000
Divide by 1000	> 1400	> 1600	> 1800	> 2200	> 2600	> 3000

Action Taken	Under 1 year	12 to 18 months	18 to 24 months	2 to 3 years	3 to 4 years	4 to 5 years
Multiply by 1000	2108 (0.44%)	621 (0.37%)	426 (0.43%)	548 (0.43%)	432 (0.49%)	328 (0.51%)
Multiply by 100	0 (0.00%)	1 (0.00%)	0 (0.00%)	2 (0.00%)	3 (0.00%)	0 (0.00%)
Multiply by 10	55 (0.01%)	33 (0.02%)	13 (0.01%)	36 (0.03%)	39 (0.04%)	33 (0.05%)
No Action Needed	461123 (96.61%)	165405 (97.53%)	95244 (97.07%)	123241 (97.57%)	85824 (97.77%)	62628 (97.69%)
Divide by 10	11505 (2.41%)	2853 (1.68%)	2022 (2.06%)	2212 (1.75%)	1348 (1.54%)	1001 (1.56%)
Divide by 100	143 (0.03%)	39 (0.02%)	20 (0.02%)	21 (0.02%)	15 (0.02%)	22 (0.03%)
Divide by 1000	2393 (0.50%)	642 (0.38%)	393 (0.40%)	255 (0.20%)	124 (0.14%)	96 (0.15%)

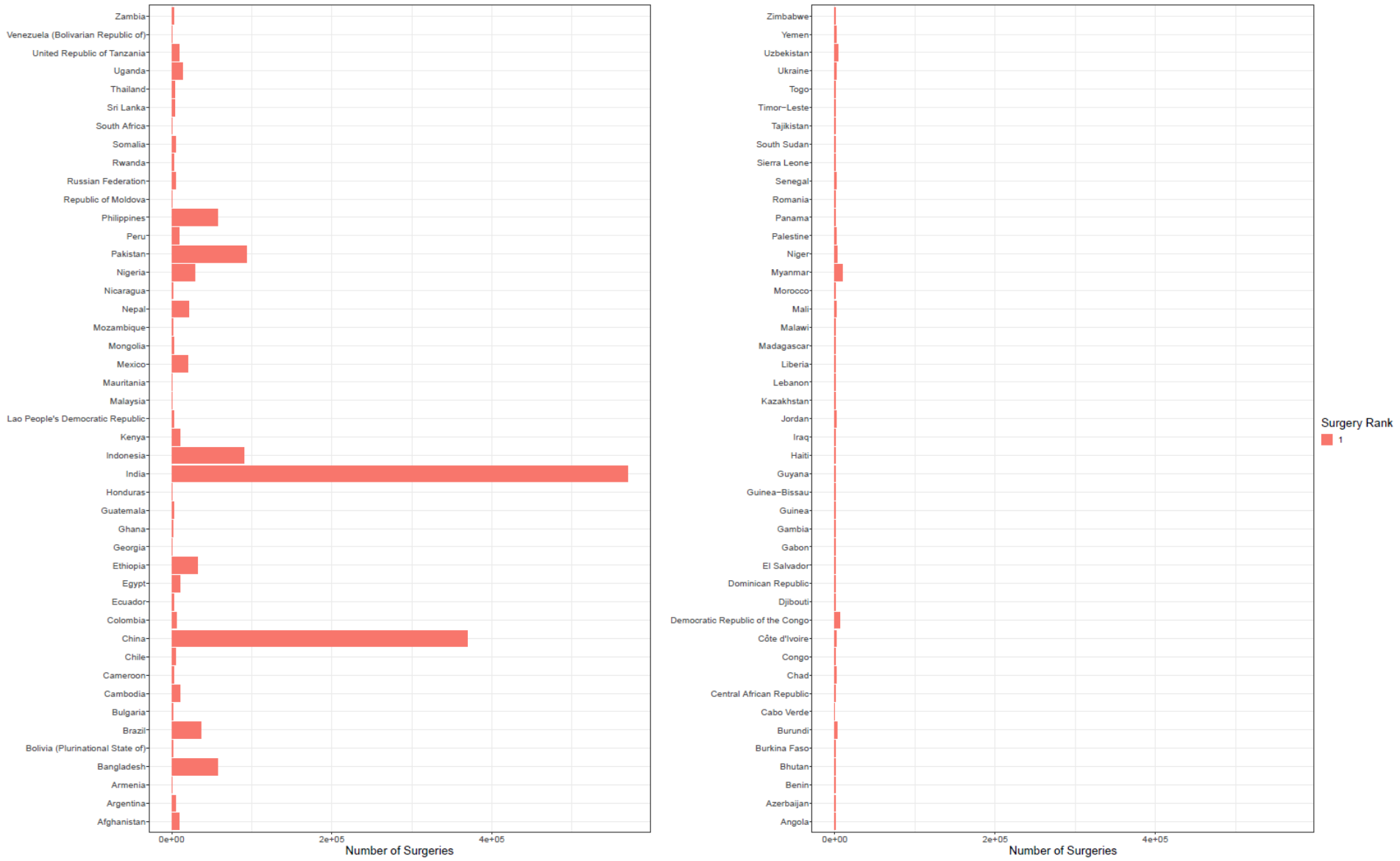
Appendix Table 8: Weight correction actions (top) and counts/ percentages where that action was taken (bottom)

Step 6: Z score Calculation and Trimming

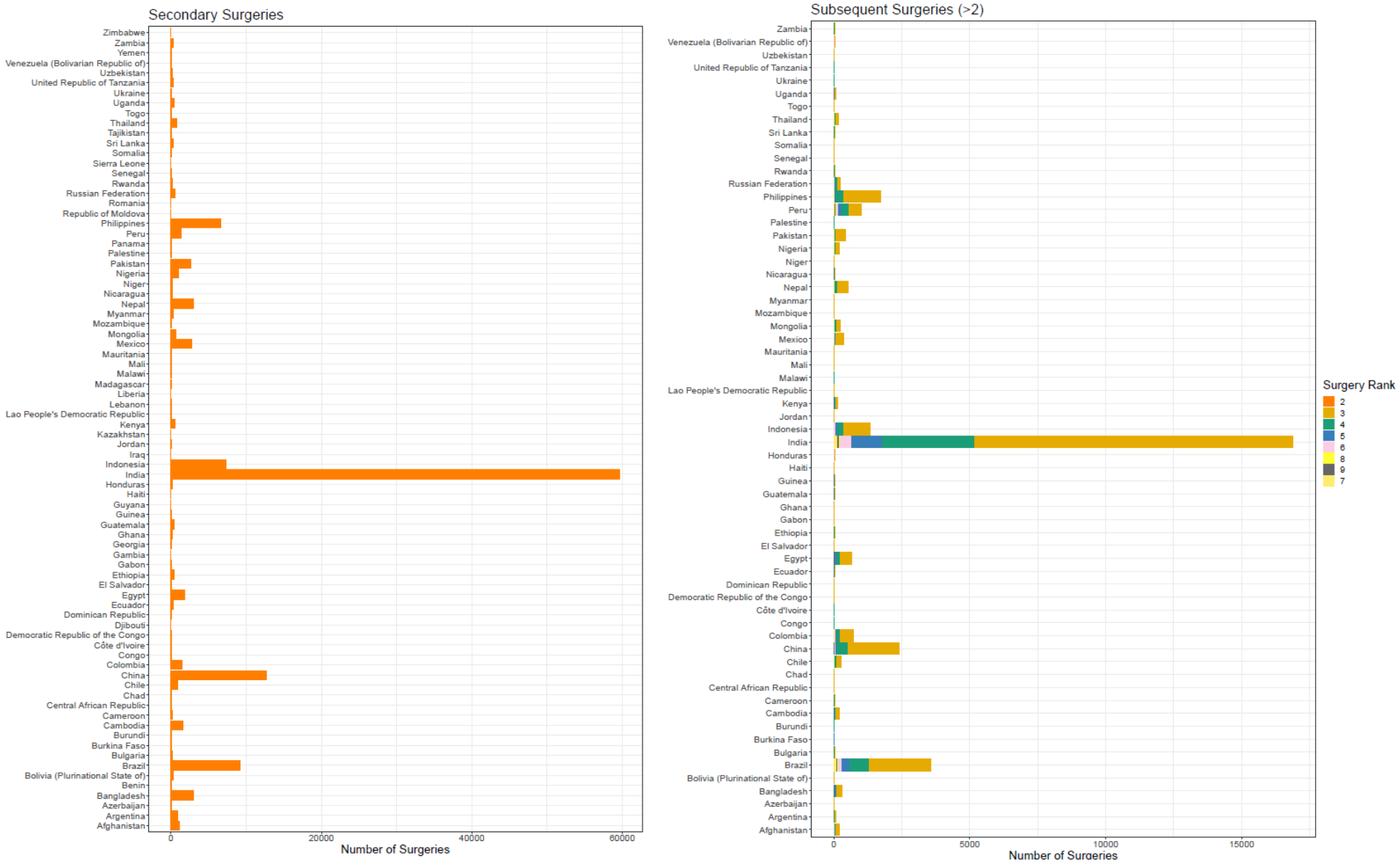
Weight-for-age Z scores (WAZ) and height-for-age Z scores (HAZ) were calculated for every individual using the 2006 WHO Growth Standards[17] and the adjusted age and weight values. As can be seen in Appendix Figure 16, this still resulted in a number of observations with implausible Z score measurements. The largest portion of these, which can also be seen in Appendix 4 Figure 5, is the apparently common practice of a child's birth weight being entered instead of a measured weight at the time of evaluation. We applied criteria of +/- 6 Z scores and removed from the dataset any observations outside that range (illustrated in Appendix Figure 17). We developed an algorithm for GBD that would allow for the very plausible situation where a child actually does have anthropometrics outside that range, but since this requires simultaneous evaluation of height, weight, and age, it could not be applied to this dataset, where a majority of individuals only had weight measurements.



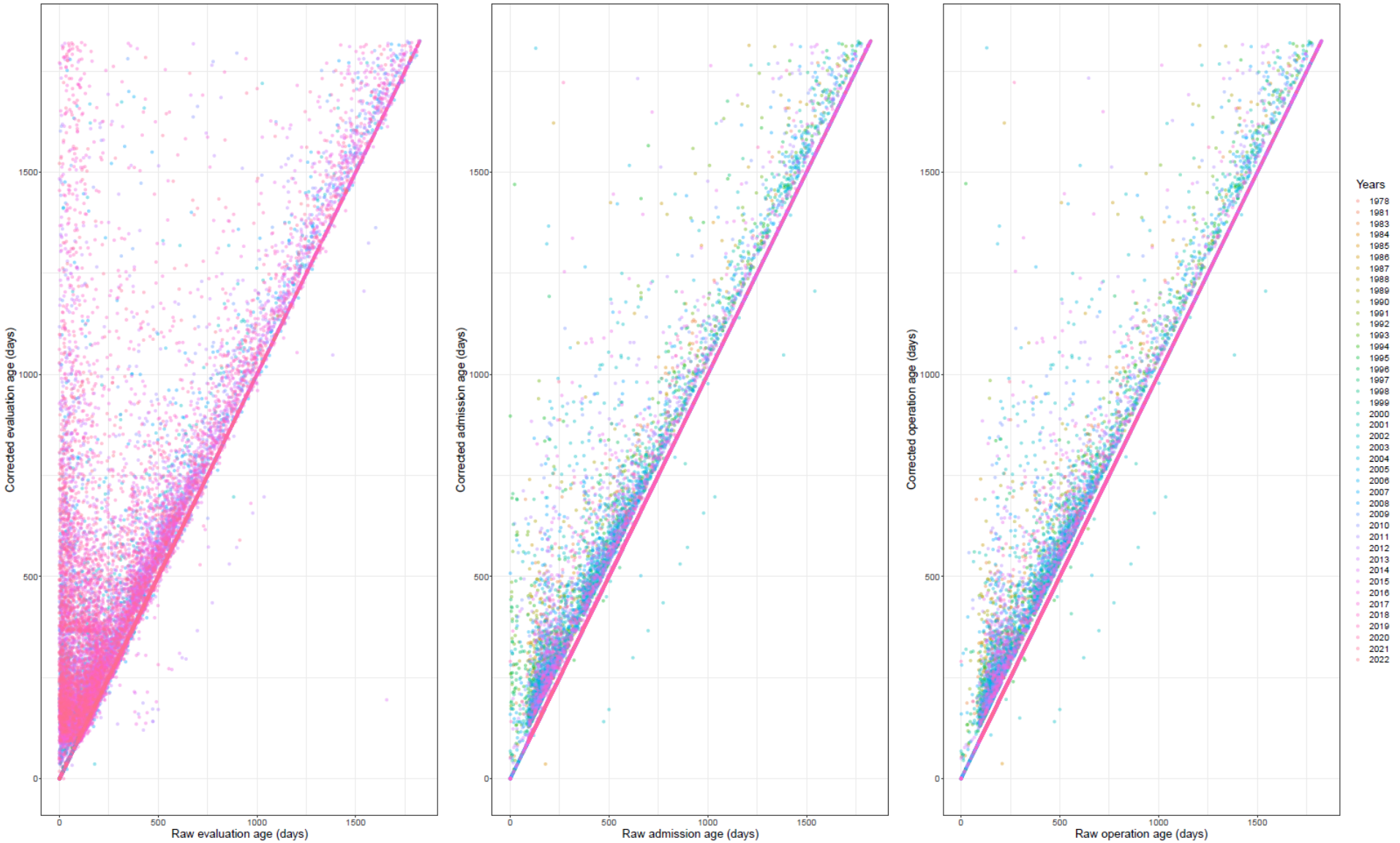
Appendix Figure 10: Evaluation age, admission age, and operation age versus evaluation weight and height for Brazil, Males, under 5 years, all years: Raw data.



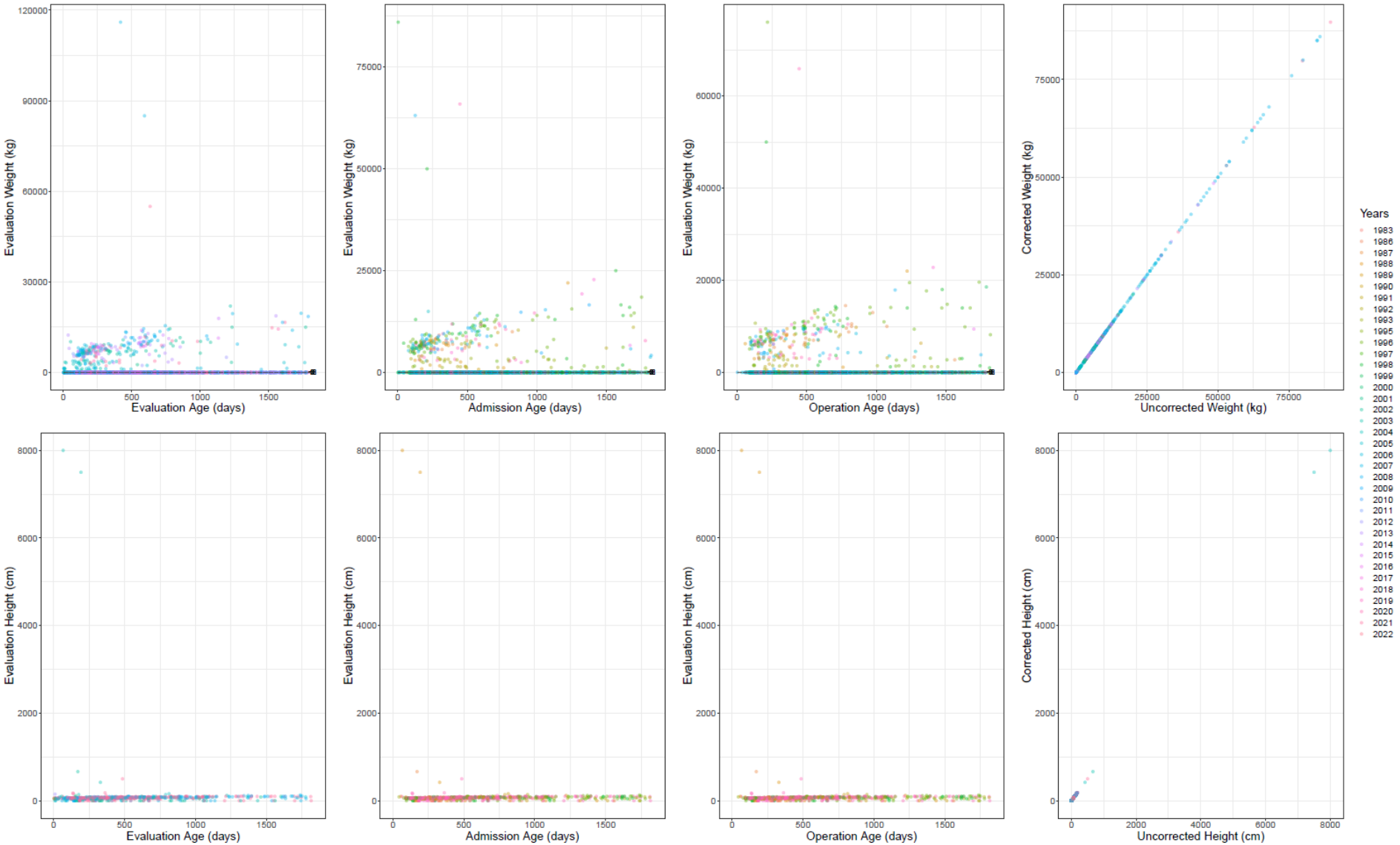
Appendix Figure 11: Number of **primary** surgeries by location across all years, under 5 years, both sexes.



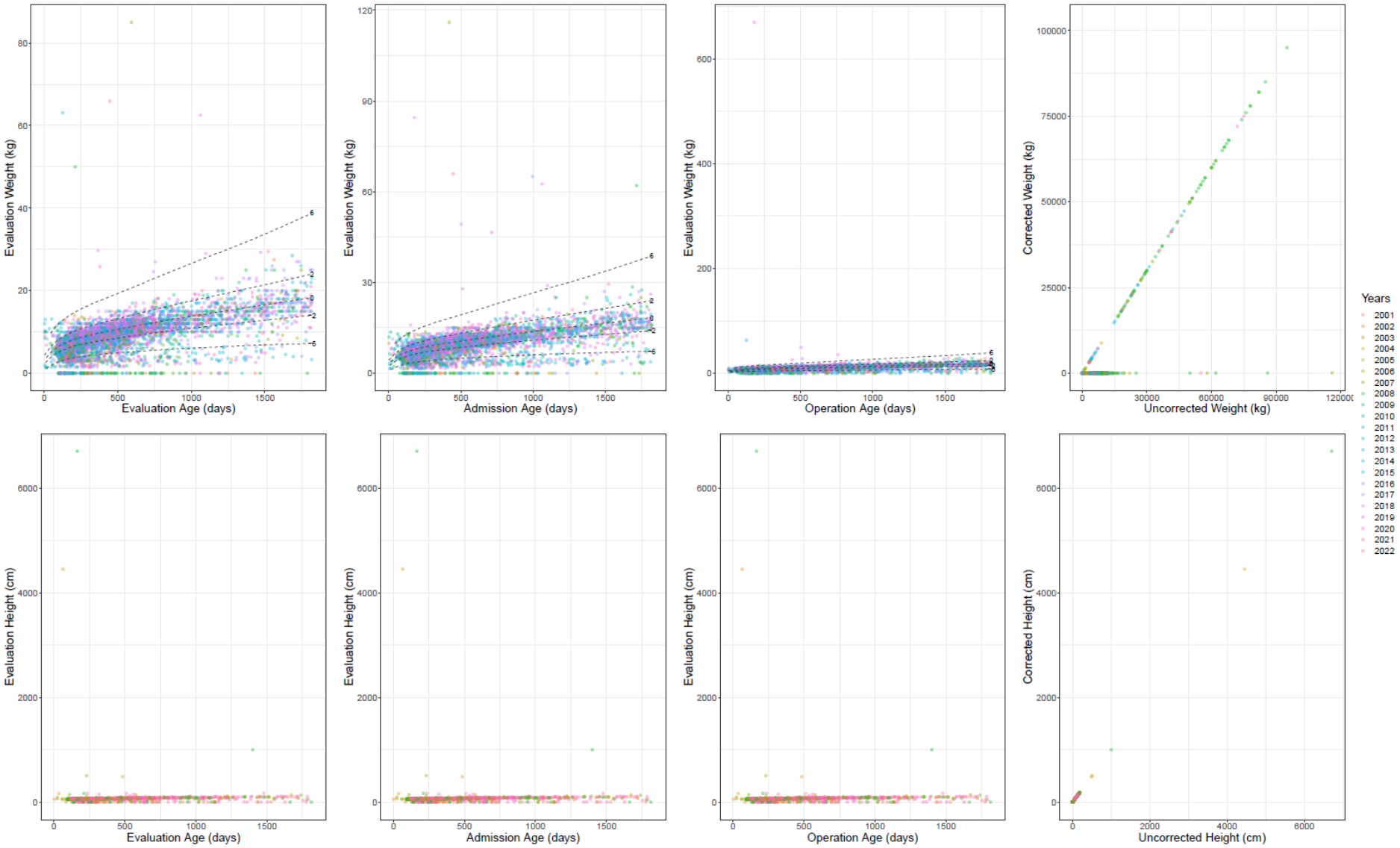
Appendix Figure 12: Number of **non-primary** surgeries by rank order and location across all years, under 5 years, both sexes.



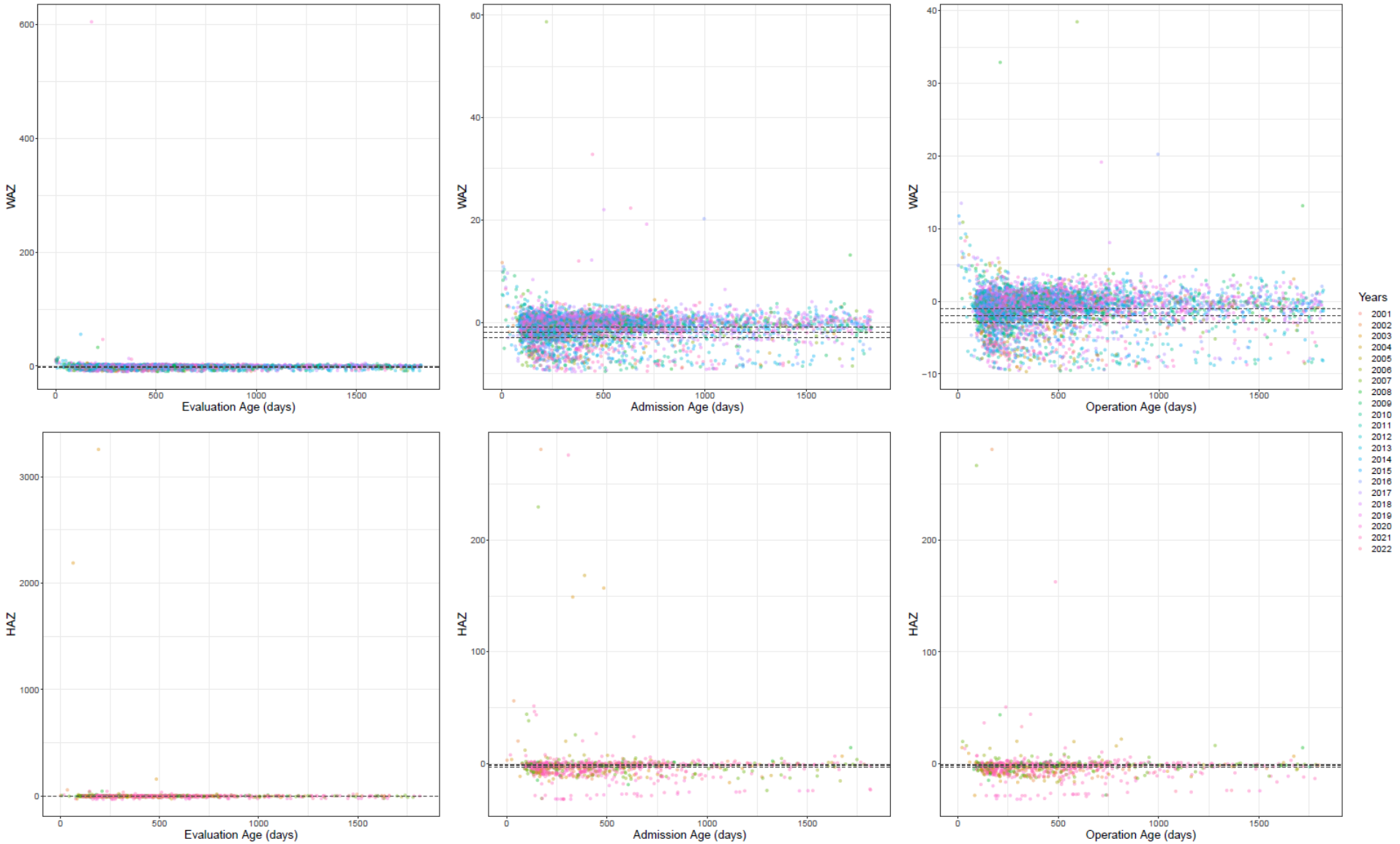
Appendix Figure 13: Raw versus corrected evaluation, admission, and surgery age for Brazil, Males, under 5 years, all years



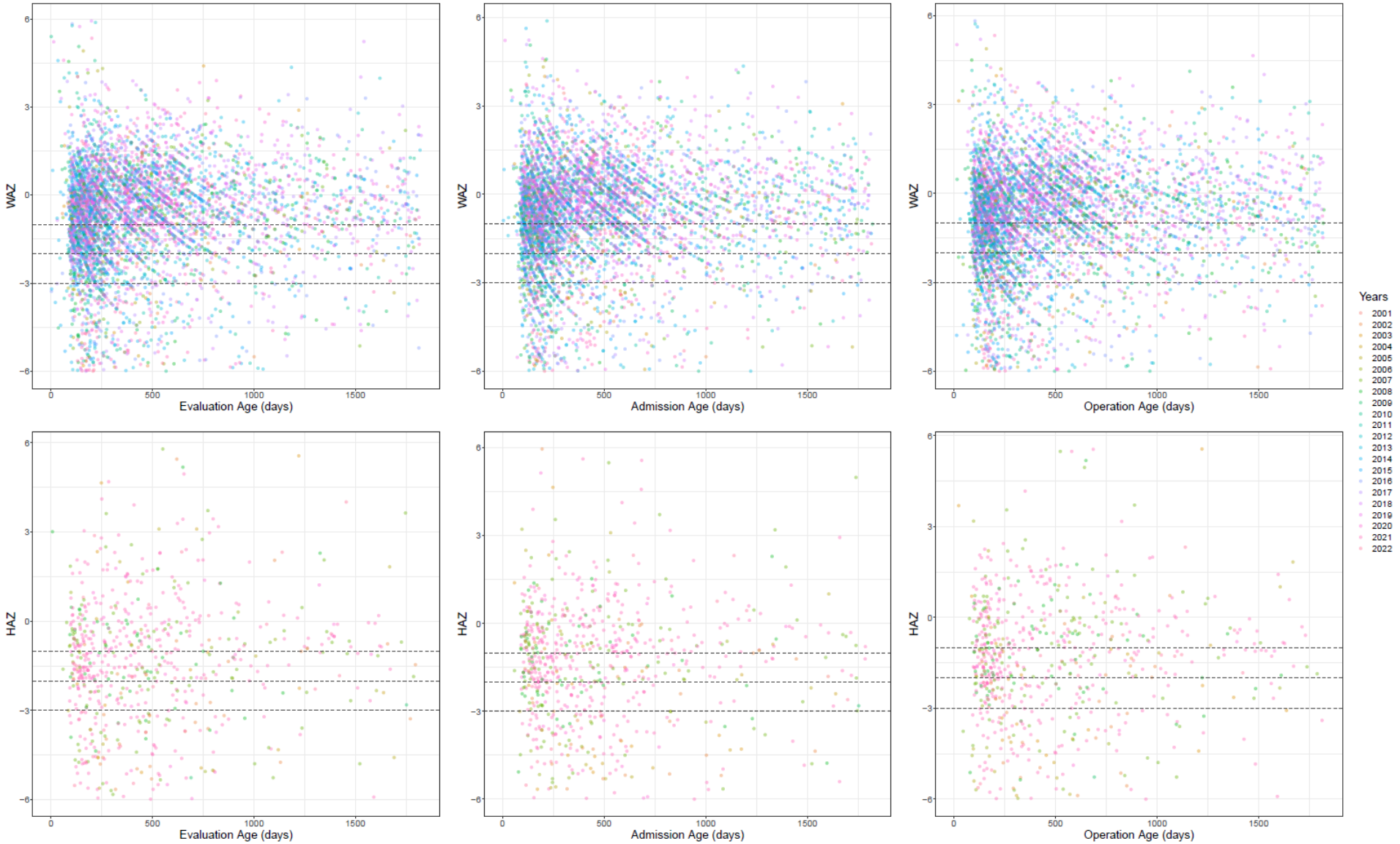
Appendix Figure 14: Evaluation age, admission age, and operation age versus evaluation weight and height for Brazil, Males, under 5 years, all years: **After date corrections**.



Appendix Figure 15: Evaluation age, admission age, and operation age versus evaluation weight and height for Brazil, Males, under 5 years, all years: **After weight corrections**.



Appendix Figure 16: Evaluation age, admission age, and operation age versus evaluation weight-for-age and height-for age Z scores (WAZ, HAZ) for Brazil, Males, under 5 years, all years.



Appendix Figure 17: Evaluation age, admission age, and operation age versus evaluation weight-for-age and height-for age Z scores (WAZ, HAZ) for Brazil, Males, under 5 years, all years: After dropping observation >6 and < -6 Z scores.

Modeling CGF Rates in Smile Train Patients

After all processing was completed at the individual level for Smile Train data, we calculated the prevalence of moderate underweight condition (<-2 WAZ) for every country in the dataset by age, sex, and year. These prevalence rates were then paired with corresponding GBD 2020 estimates of moderate underweight prevalence to calculate a prevalence rate ratio (PRR) as follows:

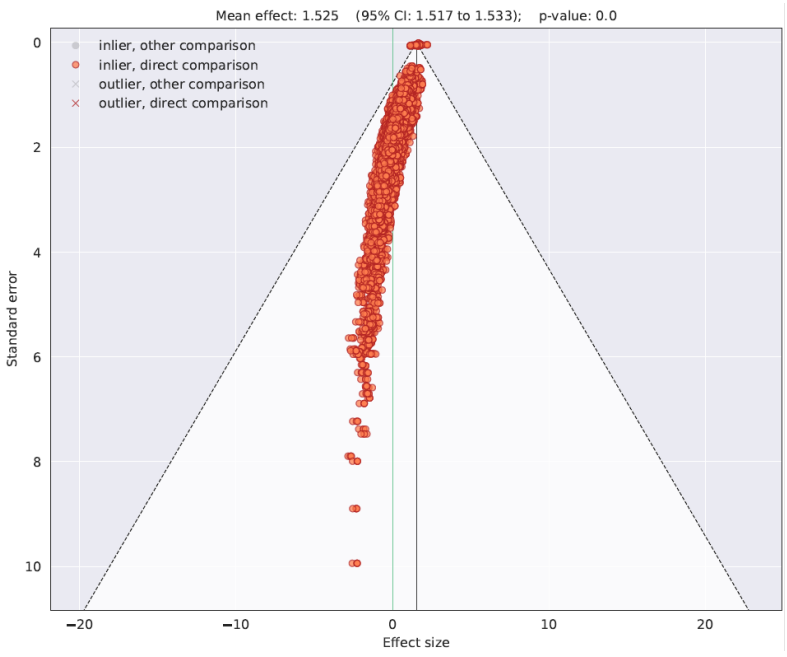
$$PRR = \frac{Underweight\ prevalence\ in\ Smiletrain}{Underweight\ prevalence\ in\ GBD\ 2020}$$

Uncertainty in PRR was calculated from the mean and uncertainty in each quantity using the Wilson method.[39]

Meta-regression of CGF: Cleft Measures

We assessed global trends in PRR using an advanced meta-analytic tool called Meta-Regression Bayesian Regularized Trimmed (MR-BRT),[22] which was developed by IHME as a flexible, mixed-effects model that can be used to synthesize data from multiple sources in a network meta-regression. This approach incorporates prior knowledge in the form of regularization or constraints in the optimization and optional automatic likelihood-based outlier detection and allows the user to statistically evaluate drivers of residual heterogeneity (e.g., study design, quality criteria) and constitutional factors that would affect the relationship (e.g., age, sex, sociodemographic factors, or other predictive covariates).

We tested a number of different model formulations, starting with a pooled meta-analysis that included no trimming of observations and no differential effects for age, sex, or location. This analysis, as shown in Appendix Figure 18, found a pooled natural log-transformed PRR effect size of 1.525 (95% confidence interval [CI] 1.517–1.533). Exponentiating this value would suggest that the PRR globally is 4.6 (4.55–4.63).

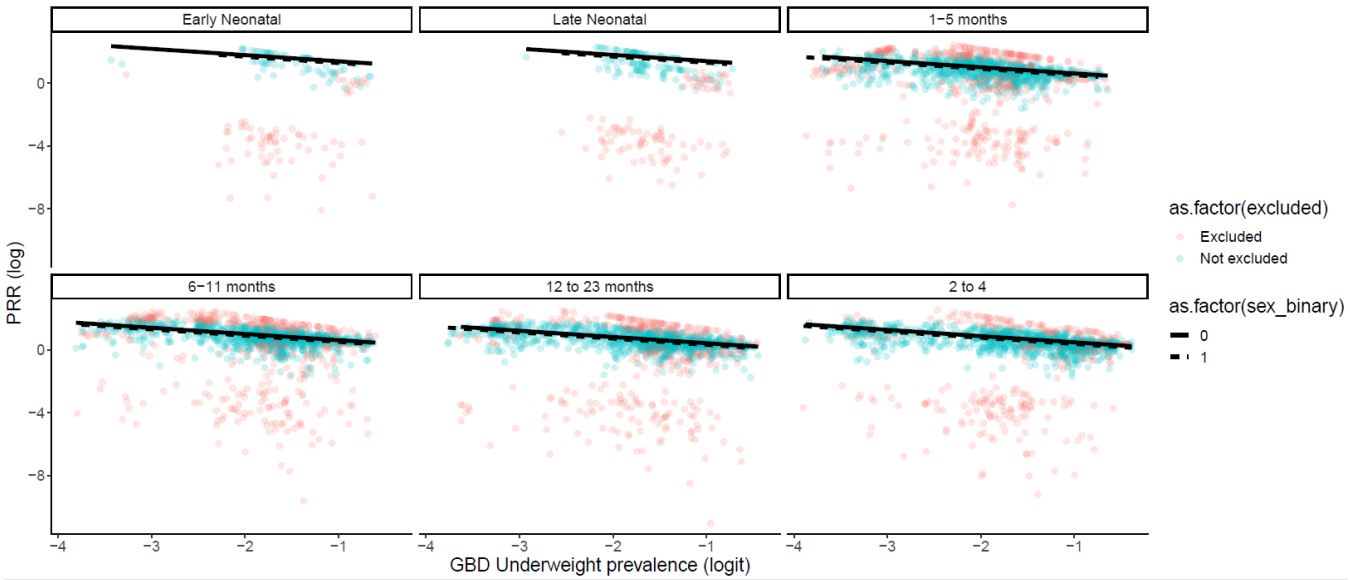


Appendix Figure 18: Meta-analyzed LN(PRR) globally for underweight in cleft: underweight in general population

However, because this dataset contains observations from across a wide range of settings, a pooled global PRR would likely overestimate the true status of underweight in cleft and not reflect local considerations. We therefore evaluated a number of patterns of differences in the PRR data in developing the final model. First, we tested and found significant sex differences in PRR across geographies, with males having higher PRR than females. We also found largely consistent age patterns, with higher PRR in the youngest age groups (<28 days)—although the sample sizes were comparatively small in these age groups—and the lowest PRR in the 12–23 months and 2–4 years age groups. We also tested a number of potential predictive covariates of PRR including Healthcare Access and Quality (HAQ) Index,[23] Sociodemographic Index (SDI),[24] and the Universal Health Coverage (UHC) Index,[25] all of which are composite indicators representing different components of societal development, access to preventive care, or medical treatment. A fourth covariate we tested, GBD underweight prevalence, was found to have the strongest relationship with PRR and one where the effect was linear.

Our final model meta-regression contained the following characteristics: the reference demographic was 2–4 year old males with fixed effects on sex, age group, and a dose response linear fixed effect on logit-transformed prevalence of underweight in the general population and 30% trimming. The final results of this model can be seen in Appendix Figure 19 and the predictions from this model for all locations can be viewed in Appendix 3. Figure 1. The observed residuals between predictions and input data were consistently quite small without directional bias so no additional location effects were added.

Covariate	Age group name	Beta	Gamma	Exponentiated beta
Alternate_age_1	Early Neonatal (0-6 days)	0.888		2.430
Alternate_age_2	Late Neonatal (7-27 days)	0.901		2.462
Alternate_age_3	1–5 months	0.114		1.121
Alternate_age_4	6–11 months	0.136		1.146
Alternate_age_5	12 to 23 months	−0.057		0.945
Reference	2 to 4 years	0.093	0	1.097
Female sex (male = ref)		−0.396		0.673
Logit(gbd_underweight_prev)		−0.105		0.900



Appendix Figure 19: Meta-analyzed LN(PRR) with sex and age effects, 30% trim, and linear logit(GBD underweight)

Assessing the Burden of Malnutrition Associated with Cleft Condition

GBD Inputs

We used multiple inputs from GBD, combined with PRR predictions, to estimate the total and excess burden of malnutrition in cleft. These included GBD estimates of prevalence of orofacial clefts, moderate underweight prevalence, and attributable deaths, YLLs, and YLDs to underweight from each of the conditions with causal relationships to underweight that have been included in GBD (measles, lower respiratory infections, diarrhea, and protein energy malnutrition). We also used GBD population estimates to be able to estimate both rates of outcomes as well as total counts.

Derivative Measure: Total CGF in Cleft, Excess CGF in Cleft, and Prevalence Rate Ratio

Prevalence rate ratio was estimated directly from the MR-BRT model described above. We calculated the total rate (and cases) of underweight in cleft using the following formula:

$$Cleft\_UW_{\boldsymbol{total}\_rate} = \boldsymbol{PRR} * GBD_{UW\_rate} * GBD_{cleft\_rate}$$

And the excess rate (and cases) of underweight in cleft using the following formula:

$$Cleft\_UW_{\boldsymbol{excess}\_rate} = (\boldsymbol{PRR} - \mathbf{1}) * GBD_{UW\_rate} * GBD_{cleft\_rate}$$

Estimating Health Consequences of Malnutrition in Those with Clefts

The fatal and nonfatal health consequences of underweight in cleft were derived from the relationships found in GBD between exposure and health outcomes. Whereas the GBD analysis discretely evaluated the exposure to mild (<-1), moderate (<-2), and severe (<-3) Z scores and corresponding risk of illness, we only estimated PRR for moderate underweight in this analysis due to considerations leading to potential stochastic variation that could cause concerns with internal-consistency of results. To ensure comparability with GBD results, we approximated the number of associated deaths, years of life lost (YLLs), and years lived with disability (YLDs) associated with underweight in cleft by first calculating the number of attributable deaths, YLLs, and YLDs per case of moderate underweight (<-2) from overall GBD results. This was completed separately for each of measles, LRI, diarrhea, and PEM and we also calculated this separately for each location, year, age group, and sex to ensure secular trends in this relationship carried through to estimates of burden associated with underweight in cleft. We then multiplied deaths/case, YLLs/case, and YLDs/case by the estimated total and excess number of underweight cases in cleft to arrive at a final estimate of attributable burden. Attributable burden counts by cause, location, age group, sex and year were aggregated to both sexes, under 5 years combined, and for geographic groupings including GBD regions, GBD super regions, and the global level. We present annual and cumulative attributable cases, deaths, YLLs, and YLDs from 2000 to 2020. All of the results of the analysis can be viewed in Appendix 3 Tables and Figures.

Appendix 2: References to Descriptions of GBD Modeling Tools

Here we provide references for accessing the methodological details for the modeling tools used to generate the GBD estimates (spatiotemporal Gaussian process regression [ST-GPR] and DisMod-MR 2.1) and to perform the de novo analysis completed from the Smile Express database (Meta-Regression Bayesian Regularized Trimmed [MR-BRT]). Methodological details for DisMod-MR 2.1 can be found in "Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019"[34] and its associated appendices. For details on ST-GPR, please see "Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019"[40] and its associated appendices. For a complete description of MR-BRT, please see "Trimmed Constrained Mixed Effects Models: Formulations and Algorithms."[22]

Appendix 3: List of Results Tables and Figures

Appendix 4: List of Figures of Harmonized Dataset

Appendix 5: Analytic Code

A full harmonized dataset of individual records has been stored on IHME limited use, password-protected, secure file storage and is available to Smile Train upon request.

References

26. Vos T. Supplement to GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Accessed May 13, 2022. https://www.thelancet.com/cms/10.1016/S0140-6736(20)30925-9/attachment/deb36c39-0e91-4057-9594-cc60654cf57f/mmc1.pdf

27. International Classification of Diseases (ICD). Accessed May 4, 2022. https://www.who.int/standards/classifications/classification-of-diseases

28. International Clearnighouse for Birth Defects: Surveillance and Research. Accessed May 13, 2022. http://www.icbdsr.org/

29. International Clearinghouse for Birth Defects Monitoring Systems . International Centre for Birth Defects, Programme WHG, Anomalies ER of C. *World Atlas for Birth Defects*. World Health Organization; 2003. Accessed May 13, 2022. https://apps.who.int/iris/handle/10665/42630

30. European Platform on Rare Disease Registration. Accessed May 13, 2022. https://eu-rd-platform.jrc.ec.europa.eu

31. Maternal and Child Health: China. Accessed May 13, 2022. https://www.unicef.cn/sites/unicef.org.china/files/2019-06/03EN-MCH%20Atlas%202018.pdf

32. National Birth Defects Prevention Network. Accessed May 13, 2022. https://www.nbdpn.org/

33. Birth Defects Registry of India. Accessed May 13, 2022. https://fcrf.org.in/bdri_abus.asp

34. Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*. 2020;396(10258):1204-1222. doi:10.1016/S0140-6736(20)30925-9

35. Global Fortification Data Exchange | GFDx – Providing actionable food fortification data all in one place. Accessed July 29, 2022. https://fortificationdata.org/

36. WHO Global Database on Child Growth and Malnutrition. Accessed April 26, 2022. https://www.who.int/teams/nutrition-and-food-safety/databases/nutgrowthdb

37. Cole TJ. The LMS method for constructing normalized growth standards. *Eur J Clin Nutr*. 1990;44(1):45-60.

38. Alvarez JL, Dent N, Browne L, Myatt M, Briend A. *Putting Child Kwashiorkor on the Map*.; 2016:55.

39. Wallis S. Binomial Confidence Intervals and Contingency Tests: Mathematical Fundamentals and the Evaluation of Alternative Methods. *J Quant Linguist*. 2013;20(3):178-208. doi:10.1080/09296174.2013.799918

40. Murray CJL, Aravkin AY, Zheng P, et al. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*. 2020;396(10258):1223-1249. doi:10.1016/S0140-6736(20)30752-2