



Data Release Information Sheet

Data Summary

Dataset name: Global Under-5 Diarrhea Incidence, Prevalence, and Mortality Geospatial Estimates 2000-2019

Date of release: August 31, 2020

Summary:

Annual estimates were produced for the prevalence and incidence of diarrhea and diarrhea-related mortality among children younger than 5 years of age at the 5x5 km-level for 94 low- and middle-income countries (LMICs) between 2000-2019. These estimates were produced using geo-positioned data from 466 household surveys, including the Demographic and Health Survey (DHS), Multiple Indicator Cluster Survey (MICS), and other country-specific surveys. Countries and subnational units outside of these 94 LMICs were supplemented with GBD results.

This dataset includes the following:

- GeoTIFF raster files for pixel-level estimates for 94 LMICs
- CSV files of aggregated estimates for 195 countries at the national level, 94 LMICs plus GBD subnational locations at the admin 1 level, and 94 LMICs at the admin 2 level
- Code files used to generate the estimates

[Get Data Files](#)

Acknowledgements

Contributing organizations:

- Institute for Health Metrics and Evaluation (IHME)

Funders:

- Bill and Melinda Gates Foundation (BMGF)

Suggested Citation:

Institute for Health Metrics and Evaluation (IHME). Global Under-5 Diarrhea Incidence, Prevalence, and Mortality Geospatial Estimates 2000-2019. Seattle, United States of America: Institute for Health Metrics and Evaluation (IHME), 2020.

Data Files Information

CSV files of Aggregated Estimates of Childhood Diarrhea

Stored in files named <MEASURE>_<AGE_GROUP>_<SEX>_<LEVEL_OF_AGGREGATION>.CSV

(Example: IHME_GLOBAL_DIARRHEA_2000_2019_PREV_A1_S1_ADMIN_1_Y2020M08D31.CSV)

- **Measure:** Incidence, Prevalence, Mortality
- **Age group:** GBD age group IDs (1-5): Under-5 (1), Early Neonatal (2), Late Neonatal (3), Post Neonatal (4), and 1 to 4 (5)
- **Sex:** GBD sex IDs (1-3): Males (1), Females (2), Both (3)
- **Level of aggregation:** admin0, admin1, or admin2, corresponding to first and second administrative level areas as defined in the Database of Global Administrative Areas (GADM) 2019 shapefiles, with adjustments made in some countries. Each row in each table is unique by administrative unit and year.

Variable	Variable Label	Variable Definition
ADM0_CODE	GADM Admin 0 Code	GADM code identifying the administrative unit
ADM0_NAME	Admin 0 Name	Zero level administrative unit (Country) name as found in the GADM shapefile
ADM1_CODE	GADM Admin 1 Code	GADM code identifying the administrative unit (Only in the admin1 files)
ADM1_NAME	Admin 1 Name	First level administrative unit name as found in the GADM shapefile
ADM2_CODE	GADM Admin 2 Code	GADM code identifying the administrative unit (Only in the admin2 files)
ADM2_NAME	Admin 2 Name	Second level administrative unit name as found in the GADM shapefile (Only in the admin2 files)
year	Year	Time period of estimate. Possible values: years in the range 2000-2019
age_group_id	Age Group ID	Unique numeric identifier for the age group generated and stored in an IHME database of data dimensions. Possible values: 1,2,3,4,5
age_group_name	Age Group Name	Age group estimated. Possible values: Under-5, early neonatal, late neonatal, post neonatal, and 1 to 4
sex_id	Sex ID	Unique numeric identifier for the sex generated and stored in an IHME database of data dimensions. Possible values: 1,2,3
sex	Sex	Sex estimated: Possible values: Male, Female, Goth

Variable	Variable Label	Variable Definition
measure	Measure	The measure (indicator) estimated. Possible values: <ul style="list-style-type: none"> • Incidence • Mortality • Prevalence
metric	Metric	Metric/unit of measure for the estimate. Values: Rate
mean	Mean	Mean posterior population-weighted estimate for the administrative unit
lower	Lower Confidence Interval	2.5% population-weighted posterior quantile estimate for the administrative unit
upper	Upper Confidence Interval	97.5% population-weighted posterior quantile estimate for the administrative unit

Codebooks

Variable names, labels, and value encoding for admin 0 files can be found in the machine-actionable codebook file [IHME_GLOBAL_DIARRHEA_2000_2019_CODEBOOK_ADMIN_0_Y2020M08D31.CSV](#)

Variable names, labels, and value encoding for admin 1 files can be found in the machine-actionable codebook file [IHME_GLOBAL_DIARRHEA_2000_2019_CODEBOOK_ADMIN_1_Y2020M08D31.CSV](#)

Variable names, labels, and value encoding for admin 2 files can be found in the machine-actionable [IHME_GLOBAL_DIARRHEA_2000_2019_CODEBOOK_ADMIN_2_Y2020M08D31.CSV](#)

GeoTIFF Raster Files for Pixel-level Estimates of Childhood Diarrhea

Stored in files named <MEASURE>_<METRIC>_<AGE_GROUP>_<SEX>_<STAT>.TIF

(Example: [IHME_GLOBAL_DIARRHEA_2000_2019_PREV_RT_A1_S1_MEAN_Y2020M08D31.TIF](#))

- **Measure:** Incidence, Prevalence, Mortality (Diarrhea)
- **Metric:** Rate (RT)
- **Age group:** GBD age group IDs (1-5): Under-5 (1), Early Neonatal (2), Late Neonatal (3), Post Neonatal (4), and 1 to 4 (5)
- **Sex:** GBD sex IDs (1-3): Males (1), Females (2), Both (3)
- **Stat:** mean, upper, or lower summary statistics from the predictive posterior distribution at each pixel. Lower and upper correspond to 2.5% and 97.5% quantiles
- **Year:** From 2000 to 2019, corresponding to the time period of the estimate

Note that rasters mask (i.e., have NA values) for lakes and areas with low population (10 people per 1km and classified as barren/sparsely vegetated). Rasters are stacked, with a layer for each year from 2000-2019. The first layer corresponds to 2000, the 20th layer corresponds to 2019.

Data Input Sources

This file contains relevant metadata about the input sources as suggested in the [Guidelines for Accurate and Transparent Health Estimates Reporting \(GATHER\)](#), a statement that promotes best practices in reporting health estimates.

IHME_GLOBAL_DIARRHEA_2000_2019_DATA_INPUT_SOURCES_Y2020M08D31.XLSX

Additional Information

Terms and Conditions

<http://www.healthdata.org/about/terms-and-conditions>

Contact information

To request further information about this dataset, please contact IHME:

Institute for Health Metrics and Evaluation

2301 Fifth Ave., Suite 600

Seattle, WA 98121

USA

Telephone: +1-206-897-2800

Fax: +1-206-897-2899

Email: data@healthdata.org

www.healthdata.org

These files may be updated periodically, so we appreciate hearing feedback or additional information about how these data are being used.