



# Institute for Health Metrics and Evaluation

## Data Release Information Sheet

### ***Data Summary***

Dataset name: Low- and Middle-Income Country Under-5 Diarrhea Incidence, Prevalence, and Mortality Geospatial Estimates 2000-2017

Date of release: May 6, 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 2000-2017. These estimates were produced using geo-positioned data from 466 household surveys. Survey sources used include the Demographic and Health Survey (DHS) and UNICEF Multiple Indicator Cluster Survey (MICS) series, and other country-specific surveys.

This dataset includes the following:

- GeoTIFF raster files for pixel-level estimates of under-5 diarrhea prevalence, incidence, and diarrhea-related mortality
- CSV files of aggregated under-5 diarrhea prevalence, incidence, and diarrhea-related mortality for each country at zero, first and second administrative divisions
- Code files used to generate the estimates

#### [Get Data Files](#)

#### Relevant publications and visualizations:

- Reiner RC, Local Burden of Disease Diarrhoea Collaborators. Mapping geographic inequalities in childhood diarrhoeal morbidity and mortality in low- income and middle-income countries, 2000–2017: analysis for the Global Burden of Disease Study 2017. *The Lancet*. 6 May 2020.
- [Local Burden of Disease – Diarrhea](#)

## 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). Low- and Middle-Income Country Under-5 Diarrhea Incidence, Prevalence, and Mortality Geospatial Estimates 2000-2017. 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>\_<LEVEL\_OF\_AGGREGATION>.CSV

(Example: IHME\_LMIC\_DIARRHEA\_2000\_2017\_PREV\_ADMIN\_1\_Y2020M02D29.CSV)

- **Measure:** Incidence, Prevalence, Mortality (Diarrhea)
- **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 2000-2017
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
age_group_name	Age Group Name	Age group estimated. Possible values: Under 5

Variable	Variable Label	Variable Definition
sex_id	Sex ID	Unique numeric identifier for the sex generated and stored in an IHME database of data dimensions. Possible values: 3
sex	Sex	Sex estimated: Possible values: Both
measure	Measure	The measure (indicator) estimated. Possible values: <ul style="list-style-type: none"> <li>• Incidence</li> <li>• Mortality</li> <li>• Prevalence</li> </ul>
metric	Metric	Metric/unit of measure for the estimate. Values: Counts
mean	Mean	Mean posterior population-weighted estimate for the administrative unit
upper	Upper Confidence Interval	97.5% population-weighted posterior quantile estimate for the administrative unit
lower	Lower Confidence Interval	2.5% 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\\_LMIC\\_DIARRHEA\\_2000\\_2017\\_CODEBOOK\\_ADMIN\\_0\\_Y2020M01D14.CSV](#)

Variable names, labels, and value encoding for admin 1 files can be found in the machine-actionable codebook file [IHME\\_LMIC\\_DIARRHEA\\_2000\\_2017\\_CODEBOOK\\_ADMIN\\_1\\_Y2020M01D14.CSV](#)

Variable names, labels, and value encoding for admin 2 files can be found in the machine-actionable [IHME\\_LMIC\\_DIARRHEA\\_2000\\_2017\\_CODEBOOK\\_ADMIN\\_2\\_Y2020M01D14.CSV](#)

## GeoTIFF Raster Files for Pixel-level Estimates of Childhood Diarrhea

Stored in files named <MEASURE>\_<METRIC>\_<STAT>\_<YEAR>.TIF

(Example: IHME\_LMIC\_DIARRHEA\_2000\_2017\_PREV\_CT\_MEAN\_2003\_Y2020M02D29.TIF)

- **Measure:** Incidence, Prevalence, Mortality (Diarrhea)
- **Metric:** Counts (CT) or Rates (RT)
- **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 2017, 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 2000-2017. The first layer corresponds to 2000, the 18th layer corresponds to 2017.

## **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\_LMIC\_DIARRHEA\_2000\_2017\_DATA\_INPUT\_SOURCES\_Y2020M01D14.CSV**

## ***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](mailto:data@healthdata.org)

[www.healthdata.org](http://www.healthdata.org)

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