



# Institute for Health Metrics and Evaluation

## Data Release Information Sheet

### ***Data Summary***

Dataset name: Global Lymphatic Filariasis Prevalence Geospatial Estimates 2000-2018

Date of release: August 19, 2020

#### Summary:

Estimates were produced for lymphatic filariasis (LF) all-age prevalence at the 5x5 km-level in endemic countries across Africa, Asia, and Hispaniola, annually between 2000 and 2018. Bayesian time series estimates were produced for 17 small area geographies in South America, the Indian Ocean, and Oceania. These estimates were produced using data on LF and geographical locations from endemicity mapping surveys, sentinel surveillance surveys, transmission assessment surveys (TAS), and other sources.

This dataset includes the following:

- GeoTIFF raster files for pixel-level estimates of LF prevalence rate, counts, and posterior probability that prevalence was lower than 2% in 2018
- CSV files of aggregated estimates of LF prevalence rate, count and posterior probability of prevalence below 2% (2018) for each country at the zero, first, and second administrative divisions
- Code files used to generate the estimates

#### Relevant publications and visualizations:

- LBD 2019 Neglected Tropical Diseases Collaborators. The changing global distribution of lymphatic filariasis, 2000–2018. *The Lancet Global Health*. 19 Aug 2020.
- [Local Burden of Disease – Lymphatic Filariasis](#)

### ***Acknowledgements***

#### Contributing organizations:

- Institute for Health Metrics and Evaluations (IHME)

#### Funders:

- Bill & Melinda Gates Foundation (BMGF)

### Suggested Citation:

Institute for Health Metrics and Evaluation (IHME). Global Lymphatic Filariasis Prevalence Geospatial Estimates 2000-2018. Seattle, United States: Institute for Health Metrics and Evaluation (IHME), 2020.

### **Data Files Information**

#### **CSV files of aggregated LF estimates**

Stored in files named LF\_<MEASURE>\_<LEVEL\_OF\_AGGREGATION>\_<DATE>.CSV  
(Example: IHME\_GLOBAL\_LF\_2000\_2018\_PREV\_ADMIN\_1\_Y2020M05D29.CSV)

- **Measure:** LF prevalence rate (PREV), counts (COUNT) or posterior probability of prevalence below 2% in 2018 (POSTPROB2).
- **Level of aggregation:** admin0, admin1, or admin2, corresponding to first and second administrative level areas as defined in a modified version of the Database of Global Administrative Areas (GADM) 2020\_02\_20 shapefiles, with adjustments made in some countries. Each row in each table is unique by administrative unit and year.
- **Date:** Date on which the file was produced, indicating year, month and day.

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 admin 1 files)
ADM1_NAME	Admin 1 Name	First level administrative unit name as found in the GADM shapefile
ADM2_CODE	GADM Admin 1 Code	GADM code identifying the administrative unit (only in the admin 2 files)
ADM2_NAME	Admin 1 Name	Second level administrative unit name as found in the GADM shapefile

<b>Variable</b>	<b>Variable Label</b>	<b>Variable Definition</b>
year	Year	Time period of the estimate. Possible values: years in the range 2000-2018
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: 22
age_group_name	Age Group Name	Age group estimated. Possible values: All Ages
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) of the estimate. Possible values: Prevalence
metric	Metric	Metric/unit of measure for the estimate. Possible values: Count, Rate, Prevalence < 2%
mean	Mean	Mean posterior population-weighted posterior estimate for the administrative unit
lower	95% uncertainty interval (lower bound)	2.5% population-weighted posterior quantile estimate for the administrative unit
upper	95% uncertainty interval (upper bound)	97.5% population-weighted posterior quantile estimate for the administrative unit
Posterior Probability	posterior_probability	Posterior probability that prevalence was lower than 2% (posterior probability files only).

## Codebooks

Variable names, labels, and value encoding for admin 0 files can be found in the machine-actionable codebook file

[IHME\\_GLOBAL\\_LF\\_2000\\_2018\\_CODEBOOK\\_ADMIN\\_0\\_Y2020M05D29.CSV](#)

Variable names, labels, and value encoding for admin 1 files can be found in the machine-actionable codebook file

[IHME\\_GLOBAL\\_LF\\_2000\\_2018\\_CODEBOOK\\_ADMIN\\_1\\_Y2020M05D29.CSV](#)

Variable names, labels, and value encoding for admin 2 files can be found in the machine-actionable codebook file

[IHME\\_GLOBAL\\_LF\\_2000\\_2018\\_CODEBOOK\\_ADMIN\\_2\\_Y2020M05D29.CSV](#)

## GeoTIFF Raster Files for Pixel-level Estimates

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

(Example: IHME\_GLOBAL\_LF\_2000\_2018\_PREV\_MEAN\_2015\_Y2020M05D29.TIF)

- **Indicators:** LF prevalence rate and counts for antigenaemia (ICT test), and posterior probability that ICT prevalence is lower than 2% in 2018.
- **Measure:** Prevalence rate, counts, or posterior probability.
- **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 2018, corresponding to the time period of the estimate.
- **Date:** Date on which the file was produced, indicating year, month and day.

Note that lakes and areas with low population density (10 people per 1 km<sup>2</sup> and classified as barren/sparse vegetated) are masked from summary estimates.

## 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\\_LF\\_2000\\_2018\\_DATA\\_INPUT\\_SOURCES\\_Y2020M05D29.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](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.