Institute for Health Metrics and Evaluation

Data Release Information Sheet

Data Summary

Dataset name: Africa and Yemen Onchocerciasis Prevalence Geospatial Estimates 2000-2018

Date of release: September 7, 2022

Summary:

Estimates were produced for onchocerciasis all-age microfiladermia (positive skin snip) prevalence at the 5x5 km-level in 34 endemic countries across Africa, plus Yemen, annually between 2000 and 2018. These estimates were produced using reported data on onchocerciasis prevalence from endemicity mapping surveys, surveillance during elimination programs, and other sources.

This dataset includes the following:

- GeoTIFF raster files for pixel-level estimates of onchocerciasis prevalence
- CSV files of aggregated estimates of onchocerciasis prevalence for each country at the national level, and for first and second administrative divisions
- Code files used to generate the estimates

Relevant publications and visualizations:

- Schmidt CA, Cromwell EA, Hill E, Donkers KM, Schipp MF, Johnson KB, et al. The prevalence of onchocerciasis in Africa and Yemen, 2000–2018: a geospatial analysis. *BMC Medicine*. 7 September 2022.

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). Africa and Yemen Onchocerciasis Prevalence Geospatial Estimates 2000-2018. Seattle, United States: Institute for Health Metrics and Evaluation (IHME), 2022.

Data Files Information

CSV files of Aggregated Onchocerciasis Prevalence Estimates

Stored in files named ONCHO_<MEASURE>_<LEVEL_OF_AGGREGATION>.CSV
(Example: IHME_AFRICA_ONCHO_2000_2018_PREV_ADMIN_1_Y2020M08D18.CSV)

- **Measure**: Onchocerciasis prevalence (PREV)
- **Level of aggregation**: admin0, admin1, or admin2, corresponding to national, first and second administrative level areas as defined in a modified version of the Database of Global Administrative Areas (GADM) 2020_05_21 shapefile. 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 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
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

Variable	Variable Label	Variable Definition
measure	Measure	The measure (indicator) of the estimate
metric	Metric	Metric/unit of measure for the estimate
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

Codebooks

Variable names, labels, and value encoding for admin 0 files can be found in the machine-actionable codebook file **IHME_AFRICA_ONCHO_2000_2018_CODEBOOK_ADMIN_0_Y2020M08D18.CSV**

Variable names, labels, and value encoding for admin 1 files can be found in the machine-actionable codebook file **IHME_AFRICA_ONCHO_2000_2018_CODEBOOK_ADMIN_1_ Y2020M08D18.CSV**

Variable names, labels, and value encoding for admin 2 files can be found in the machine-actionable codebook file **IHME_AFRICA_ONCHO_2000_2018_CODEBOOK_ADMIN_2_ Y2020M08D18.CSV**

GeoTIFF Raster Files for Pixel-level Onchocerciasis Prevalence Estimates

Stored in files named <MEASURE>_<STAT>_<YEAR>_<DATE>.TIF
(Example: IHME_AFRICA_ONCHO_2000_2018_PREV_MEAN_2015_Y2020M08D18.TIF)

- **Measure:** Onchocerciasis prevalence (PREV)
- **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^2 and classified as barren/sparsely 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_AFRICA_ONCHO_2000_2018_DATA_INPUT_SOURCES_ Y2020M08D18.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:

Address:
Institute for Health Metrics and Evaluation
Population Health Building/Hans Rosling Center
3980 15th Ave. NE, Seattle, WA 98195 USA
UW Campus Box #351615

Telephone: +1-206-897-2800

Fax: +1-206-897-2899

Email: data@healthdata.org

Website: www.healthdata.org

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