Data Release Information Sheet

**Data Summary**

**Dataset name:** Global Exclusive Breastfeeding Prevalence Geospatial Estimates 2000-2019

**Date of release:** June 3, 2021

**Summary:** Annual estimates were produced for exclusive breastfeeding prevalence among infants under 6 months of age at the 5x5 km-level for 94 low- and middle-income countries (LMICs) between 2000 and 2019. These estimates were produced using a geo-positioned dataset created from 394 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. 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 of exclusive breastfeeding prevalence among infants under 6 months of age for 94 LMICs
- CSV files of aggregated for 195 countries at the national level, 94 LMICs plus GBD subnational locations at the first-level administrative divisions, and 94 LMICs at the second-level administrative divisions
- Code files used to generate the estimates

**Relevant publications and visualizations:**


**Acknowledgements**

**Contributing organizations:**

- Institute for Health Metrics and Evaluation (IHME)

**Funders:**

- Bill and Melinda Gates Foundation (BMGF)

**Suggested Citation:**

Data Files Information

CSV files of Aggregated Estimates of Exclusive Breastfeeding

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

(Example: IHME_GLOBAL_EBF_2000_2019_EBF_PREV_ADMIN_0_Y2020M08D31.CSV)

- **Measure**: EBF Prevalence
- **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.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Variable Label</th>
<th>Variable Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADM0_CODE</td>
<td>GADM Admin 0 Code</td>
<td>GADM code identifying the administrative unit</td>
</tr>
<tr>
<td>ADM0_NAME</td>
<td>Admin 0 Name</td>
<td>Zero level administrative unit (Country) name as found in the GADM shapefile</td>
</tr>
<tr>
<td>ADM1_CODE</td>
<td>GADM Admin 1 Code</td>
<td>GADM code identifying the administrative unit (Only in the admin1 files)</td>
</tr>
<tr>
<td>ADM1_NAME</td>
<td>Admin 1 Name</td>
<td>First level administrative unit name as found in the GADM shapefile</td>
</tr>
<tr>
<td>ADM2_CODE</td>
<td>GADM Admin 2 Code</td>
<td>GADM code identifying the administrative unit (Only in the admin2 files)</td>
</tr>
<tr>
<td>ADM2_NAME</td>
<td>Admin 2 Name</td>
<td>Second level administrative unit name as found in the GADM shapefile (Only in the admin2 files)</td>
</tr>
<tr>
<td>year</td>
<td>Year</td>
<td>Time period of estimate. Possible values: years in the range 2000-2019</td>
</tr>
<tr>
<td>age_group_id</td>
<td>Age Group ID</td>
<td>Unique numeric identifier for the age group generated and stored in an IHME database of data dimensions. Possible values: 390</td>
</tr>
<tr>
<td>age_group_name</td>
<td>Age Group Name</td>
<td>Age group estimated. Possible values: &lt;6 months</td>
</tr>
<tr>
<td>sex_id</td>
<td>Sex ID</td>
<td>Unique numeric identifier for the sex generated and stored in an IHME database of data dimensions. Possible values: 3</td>
</tr>
<tr>
<td>sex</td>
<td>Sex</td>
<td>Sex estimated: Possible values: Both</td>
</tr>
<tr>
<td>measure</td>
<td>Measure</td>
<td>The measure (indicator) estimated. Possible values:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- EBF prevalence</td>
</tr>
<tr>
<td>metric</td>
<td>Metric</td>
<td>Metric/unit of measure for the estimate. Values: Percent</td>
</tr>
<tr>
<td>mean</td>
<td>Mean</td>
<td>Mean posterior population-weighted estimate for the administrative unit</td>
</tr>
<tr>
<td>lower</td>
<td>Lower Confidence Interval</td>
<td>2.5% population-weighted posterior quantile estimate for the administrative unit</td>
</tr>
<tr>
<td>upper</td>
<td>Upper Confidence Interval</td>
<td>97.5% population-weighted posterior quantile estimate for the administrative unit</td>
</tr>
</tbody>
</table>
**Codebooks**

Variable names, labels, and value encoding for admin 0 files can be found in the machine-actionable codebook file `IHME_GLOBAL_EBF_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_EBF_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_EBF_2000_2019_CODEBOOK_ADMIN_2_Y2020M08D31.CSV`

**GeoTIFF Raster Files for Pixel-level Estimates of Exclusive Breastfeeding**

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


- **Measure**: EBF prevalence
- **Metric**: Percent
- **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).

**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)](http://www.healthdata.org/about/terms-and-conditions), a statement that promotes best practices in reporting health estimates.

`IHME_GLOBAL_EBF_2000_2019_DATA_INPUT_SOURCES_SURVEY_Y2020M08D31.XLSX`

**Additional Information**

**Terms and Conditions**

[http://www.healthdata.org/about/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.