Data Release Information Sheet

Data Summary


Date of release: May 15, 2019

Summary:
Estimates were produced for HIV prevalence among adults ages 15-49 and the corresponding number of people living with HIV (PLHIV) at the 5x5 km-level in 47 countries in Africa between 2000-2017. These estimates were produced using data on HIV and geographical locations from seroprevalence surveys and sentinel surveillance of antenatal care clinics.

This dataset includes the following:

- GeoTIFF raster files for pixel-level estimates of HIV prevalence, PLHIV, and eight covariates constructed for the analysis: prevalence of male circumcision, self-reported STI symptoms, marriage or living with a partner as married, one’s current partner living elsewhere, condom use at last sexual encounter, reported intercourse among young adults, and multiple partners in the last year for men and for women
- CSV files of aggregated HIV prevalence and PLHIV estimates for each country at the zero, first, and second administrative divisions
- Code files used to generate the estimates

Relevant publications and visualizations:

- Local Burden of Disease - HIV

Acknowledgements

Contributing organizations:

- Institute for Health Metrics and Evaluation (IHME)
Funders:
- Bill and Melinda Gates Foundation (BMGF)
- National Institutes of Health (NIH)

Suggested Citation:

**Data Files Information**

**CSV files of Aggregated Estimates of Educational Attainment**
Stored in files named HIV_<MEASURE>_LEVEL_OF_AGGREGATION>.CSV
(Example: IHME_AFRICA_HIV_2000_2017_HIV_PREV_ADMIN_1_Y2019M03D15.CSV)

- **Measure**: HIV prevalence (HIV_PREV) or People living with HIV (HIV_COUNTS)
- **Level of aggregation**: admin0, admin1, or admin2, corresponding to first and second administrative level areas as defined in the 2014_2015 FAO Global Administrative Unit Layers (GAUL) shapefiles, with minor adjustments made where names were missing in the original shapefile. 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>GAUL Admin 0 Code</td>
<td>GAUL 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 GAUL shapefile</td>
</tr>
<tr>
<td>ADM1_CODE</td>
<td>GAUL Admin 1 Code</td>
<td>GAUL 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 GAUL shapefile</td>
</tr>
<tr>
<td>ADM2_CODE</td>
<td>GAUL Admin 2 Code</td>
<td>GAUL 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 GAUL 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 2000-2017</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: 24</td>
</tr>
<tr>
<td>age_group_name</td>
<td>Age Group Name</td>
<td>Age group estimated. Possible values: 15-49 years</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:</td>
</tr>
<tr>
<td>Variable</td>
<td>Variable Label</td>
<td>Variable Definition</td>
</tr>
<tr>
<td>----------</td>
<td>---------------</td>
<td>---------------------</td>
</tr>
</tbody>
</table>
| measure  | Measure       | The measure (indicator) estimated. Possible values:  
|          |               | - HIV prevalence  
|          |               | - People living with HIV |
| metric   | Metric        | Metric/unit of measure for the estimate. Values:  
|          |               | - Percent  
|          |               | - Count (Number) |
| 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_AFRICA_HIV_2000_2017_CODEBOOK_ADMIN_0_Y2019M03D15.CSV`

Variable names, labels, and value encoding for admin 1 files can be found in the machine-actionable codebook file `IHME_AFRICA_HIV_2000_2017_CODEBOOK_ADMIN_1_Y2019M03D15.CSV`

Variable names, labels, and value encoding for admin 2 files can be found in the machine-actionable file `IHME_AFRICA_HIV_2000_2017_CODEBOOK_ADMIN_2_Y2019M03D15.CSV`

**GeoTIFF Raster Files for Pixel-level Estimates of Years and Disparity**

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


- **Indicator:** HIV prevalence, People living with HIV, or one of the 8 HIV covariate measures:
  1. Prevalence of male circumcision (all forms) `[MALE_CIRCUMCISION_PREV_MEAN ]`
  2. Prevalence of self-reported STI symptoms `[STI_SYMPTOMS_PREV_MEAN]`
  3. Prevalence of marriage or living with a partner as married `[IN_UNION_PREV_MEAN]`
  4. Prevalence of one’s current partner living elsewhere `[PARTNER_AWAY_PREV_MEAN]`
  5. Prevalence of condom use at last sexual encounter `[CONDOM_LAST_TIME_PREV_MEAN]`
  6. Prevalence of reporting ever had intercourse among young adults `[HAD_INTERCOURSE_PREV_MEAN]`
  7. Prevalence of multiple partners in the last year for men `[PARTNERS_YEAR_MN_PREV_MEAN]`
  8. Prevalence of multiple partners in the last year for women `[PARTNERS_YEAR_WN_PREV_MEAN]`

- **Measure:** Prevalence (%), or counts (people living with HIV)
• **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).

**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:

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](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.