GBD Results Tool: GBD 2017 Codebook and Data Guide

Summary:

The Global Burden of Disease Study 2017 (GBD 2017), coordinated by the Institute for Health Metrics and Evaluation (IHME), estimated the burden of diseases, injuries, and risk factors for 195 countries and territories, and at the subnational level for a subset of countries.

This set of files contain the following for GBD 2017 results in the GBD Results Tool:

- **Codebook** [machine-actionable]
- **The GBD 2017 Cause Hierarchy**
- **The GBD 2017 Risk, Etiology, Impairment, and Injury n-code (REI) Hierarchy**
- **Locations hierarchies**
    - The GBD Locations Hierarchy file contains only GBD locations, including subnational locations for which results were released at the time of the study's publication.
    - The All Locations Hierarchies file also includes hierarchies for other regions for which estimates were produced, such as WHO and World Bank regions.
- **Results by Measure**
- **Metrics by Measure Definitions**

The hierarchy files allow users to filter for sets of values by level or parent category, including cause or risk group, GBD super region or region, or custom region.

Acknowledgements

File Inventory

File Name	Description	Version Date
IHME_GBD_2017_CODEBOOK_COLLAB_TOOL_Y2018M11D17.CSV	Codebook [machine-actionable]	November 18, 2018
IHME_GBD_2017_CAUSE_HIERARCHY_Y2018M11D18.XLSX	GBD 2017 Cause Hierarchy	November 18, 2018
IHME_GBD_2017_REI_HIERARCHY_Y2018M11D18.XLSX	GBD 2017 Risk, Impairment, Etiology, and Injury n-code (REI) Hierarchy	November 18, 2018
IHME_GBD_2017_GBD_LOCATIONS_HIERARCHY_Y2018M11D18.XLSX	GBD 2017 - GBD Locations Hierarchy	November 18, 2018
IHME_GBD_2017_ALL_LOCATIONS_HIERARCHIES_Y2018M11D18.XLSX	GBD 2017 - GBD and Custom Region hierarchies	November 18, 2018
IHME_GBD_2017_RESULTS_BY_MEASURE_Y2018M11D18.XLSX	Results Available by Measure: Causes, Risks, Etiologies, Impairments, Injuries by Nature, Sequelae [level 5]	November 18, 2018
IHME_GBD_2017_MEASURE_METRIC_DEFINITIONS_Y2018M11D18.XLSX	Definitions for each metric, by measure, for metrics in the GBD Results Tool	November 18, 2018
IHME_GBD_2017_CODEBOOK_INFO_SHEET_Y2018M11D08.PDF	Codebook Information Sheet	November 18, 2018

Data Files Information

Codebook

The Codebook file is a machine-actionable GBD 2017 codebook for results from the GBD Results Tool. It includes variables, labels, and value coding (with text names and numeric identifiers for each value) for dimensions of the results data.

Hierarchy Files

The **Cause Hierarchy** file**, REI Hierarchy** file (risk factor, etiology, impairment, and injury n-code), and **GBD Locations Hierarchy** file contain GBD 2017 reporting hierarchies with values in the order they appear in online tools, such as the GBD Results Tool and GBD Compare.

The **All Locations Hierarchies** file contains the GBD 2017 reporting hierarchy and a number of other hierarchies which will allow users of GBD 2017 results to aggregate results by location in various ways (by GBD regions, World Bank regions, OECD countries, African Union countries, etc.).

To filter for the child values of a given parent value in any hierarchy file (e.g., all countries in the GBD region "South Asia," or all causes under "Chronic respiratory diseases"), filter on the parent ID in the parent_id column.

To view each hierarchy in its proper nested order (the order they appear in online tools), sort by the sort_order column.

For the Cause Hierarchy file, users can also filter by cause outline value (e.g., all causes in "A" or "A.1"), or for causes for which either only years of life lost (YLLs) or years lived with disability (YLDs) results were produced.

For the All Locations Hierarchies file, filter first by location_set_version_id, and then by parent ID and/or sort order.

Locations and Levels: In the locations hierarchies files:

- Level 3 = country or territory
- Levels 4 and 5 = subnational units

Location Set Version IDs

Location Set ID	Location Set
423	GBD 2017 Reporting
385	African Union
380	Commonwealth
379	European Union
381	Four World Regions
384	G20
386	Nordic Region
383	OECD Countries
377	WHO region
382	World Bank Income Levels
378	World Bank Regions

Select Variable Definitions

Variable	Definition
cause_id	Cause IDs for cause variables come from an IHME database that creates and stores unique numeric identifiers for each GBD cause.
cause_outline	This outline represents the hierarchy of causes and the depth of each cause in the hierarchy by alpha-numeric code for the current round of GBD.
level	This indicates a location value's level in a location hierarchy.
locations_id	Locations IDs for geographic variables come from an IHME database that creates and stores unique numeric identifiers for locations of various location types.

location_set_version_id	IDs for GBD 2017 location sets (including the GBD locations and custom region locations) come from an IHME database that creates and stores unique numeric identifiers for location set versions.
parent_id	This indicates the parent ID for a value's parent in the hierarchy. For example, all 31 Mexican states and 1 federal district (Mexico City) have a parent ID of 130, the location ID for the country of Mexico.
rei_id	REI IDs for risk, etiology, impairment, and injury n-code (injury by nature) variables come from an IHME database that creates and stores unique numeric identifiers for each risk, etiology, impairment, and injury.
rei_type	This indicates whether the value is a risk, etiology, impairment, or injury.
sort_order	This can be used to sort a set of values in a particular order (not alphabetical, but the order in which the values appear in online tools, paper tables, etc.).
yll_only	Indicates causes for which only mortality estimates (deaths, years of life lost (YLLs)) were produced.
ylds_oly	Indicates causes for which only nonfatal estimates (years lived with disability (YLDs) were produced.

Results by Measure

This Excel file contains a set of sheets that display the measures (deaths, DALYs, incidence, etc.) for which estimates are provided for each GBD cause, risk, etiology, impairment, injury by nature, and sequela (only level 5 sequelae (or "sequelae aggregates") from the GBD 2017 Cause List and Sequelae Hierarchy).

Filter on cause, risk, etc., or a measure (deaths, YLDs, incidence, etc.), to see available results.

Additionally, the sheets display which causes each risk, etiology, impairment, injury by nature, and sequela are associated with. For example:

- In the "Risk" sheet, filter for "Ambient particulate matter pollution" in the "Risks" column and see the 18 causes in that cause-risk pair
- In the "Impairment" sheet, filter for "Blindness" in the "Impairments" column and see the 34 causes associated with the Blindness impairment

Metrics by Measure Definitions

This file (also included as a tab in the "Results by Measure" file) provides definitions for each metric, by measure, contained in the GBD Results Tool.

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
www.healthdata.org

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