



Institute for Health Metrics and Evaluation

GBD Results Tool: GBD 2019 Codebook and Data Guide

Date of release: October 15, 2020

Date of update: June 29, 2022

Summary:

The Global Burden of Disease Study 2019 (GBD 2019), coordinated by the Institute for Health Metrics and Evaluation (IHME), estimated the burden of diseases, injuries, and risk factors for 204 countries and territories and selected subnational locations.

This set of files contain the following for GBD 2019 results in the GBD Results Tool:

- **Codebook** [machine-actionable]
- **The GBD 2019 Cause Hierarchy**
- **The GBD 2019 Risk, Etiology, Impairment, and Injury n-code (REI) Hierarchy**
- **Locations hierarchies**
 - The GBD Location Hierarchy file contains only GBD locations, including subnational locations for which results were released at the time of the study's publication.
 - The All Location Hierarchies file also includes hierarchies for other regions for which estimates were produced, such as WHO and World Bank regions.
- **Contexts by Age**
- **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

Contributing organizations:

- Global Burden of Disease Collaborative Network

Funders:

- Bill and Melinda Gates Foundation (BMGF)

File Inventory

File Name	Description	Version Date
IHME_GBD_2019_CODEBOOK_Y2022M06D29.CSV	Codebook [machine-actionable]	June 29, 2022
IHME_GBD_2017_CAUSE_HIERARCHY_Y2020M11D25.XLSX	GBD 2019 Cause Hierarchy	November 25, 2020
IHME_GBD_2017_REI_HIERARCHY_Y2020M10D15.XLSX	GBD 2019 Risk, Impairment, Etiology, and Injury n-code (REI) Hierarchy	October 15, 2020
IHME_GBD_2019_GBD_LOCATION_HIERARCHY_Y2022M06D29.XLSX	GBD 2019 - GBD Location Hierarchy	June 29, 2022
IHME_GBD_2019_ALL_LOCATION_HIERARCHIES_Y2022M06D29.XLSX	GBD 2019 - GBD and Custom Region hierarchies	June 29, 2022
IHME_GBD_2019_CONTEXTS_BY_AGE_Y2020M12D18.XLSX	Contexts by Age	December 18, 2020
IHME_GBD_2017_MEASURE_METRIC_DEFINITIONS_Y2021M01D19.XLSX	Definitions for each metric, by measure, for metrics in the GBD Results Tool	January 19, 2021
IHME_GBD_2019_INFO_SHEET_Y2022M06D29.PDF	Codebook and Data Guide Information Sheet	June 29, 2022

Data Files Information

Data Update

The GBD Cause Hierarchy was updated on November 25, 2020 with four new causes:

1029	Total cancers
1026	Total burden related to hepatitis B
1027	Total burden related to hepatitis C
1028	Total burden related to Non-alcoholic fatty liver disease (NAFLD)

Etiology results were added to the GBD Results tool on December 18, 2020. Subnational results for South Africa were added to the tool on January 20, 2022. Subnational results for Ethiopia were added on March 13, 2022. Subnational results for Iran were added on April 6, 2022. Subnational results for Pakistan were added on April 29, 2022. Subnational results for Norway were added on June 29, 2022.

Codebook

The Codebook file is a machine-actionable GBD 2019 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 Location Hierarchy** file contain GBD 2019 reporting hierarchies with values in the order they appear in online tools, such as the GBD Results Tool and GBD Compare.

The **All Location Hierarchies** file contains the GBD 2019 reporting hierarchy and a number of other hierarchies which will allow users of GBD 2019 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 location hierarchies files:

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

Location Set Version IDs

Location Set ID	Location Set
1073	GBD 2019 Reporting
587	African Union
592	Commonwealth
591	European Union
593	Four World Regions
588	G20
586	Nordic Region
589	OECD Countries
750	WHO region
596	World Bank Income Levels
595	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.
Location 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 2019 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.
Parent Name	This indicates the name of the parent value. This variable is only in the Cause and REI hierarchies.
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.
YLD Only	Indicates causes for which only nonfatal estimates (years lived with disability (YLDs)) were produced.

Contexts by Age

This Excel file provides the full set of age groups used, cumulatively, in results available from the tool, and the context each age group appears in. Filter on contexts like Cause, Risk, Life expectancy, Population, etc. to see the ages available for each.

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:

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.