



# Institute for Health Metrics and Evaluation

## Data Release Information Sheet

### ***Data Summary***

Dataset name: Database of Geopositioned Middle East Respiratory Syndrome Coronavirus Occurrences

Date of release: July 23, 2019

Summary:

This dataset contains 882 unique occurrences of Middle East Respiratory Syndrome Coronavirus (MERS-CoV) extracted from 217 reports published between October 2012 and February 2018. Occurrences of MERS-CoV among humans, mammals, and environmental sources were extracted and geopositioned to the highest resolution possible (up to 5x5km). Special attention was given to spillover events (i.e., humans becoming infected from mammals). Thus, MERS-CoV occurrences meeting the criteria for index or unspecified cases are assigned their own rows; alternatively, multiple mammal, import, or secondary cases are collapsed to one row in the dataset.

### **Acknowledgements**

Contributing organizations:

- Institute for Health Metrics and Evaluation (IHME)

Funders:

- Bill and Melinda Gates Foundation (BMGF)

Suggested Citation:

Institute for Health Metrics and Evaluation (IHME). Geopositioned Middle East Respiratory Syndrome Coronavirus Occurrences Database 1983-2017. Seattle, WA, USA: Institute for Health Metrics and Evaluation (IHME), 2019.

## File Inventory

File Name	Description	Version Date
IHME_MERS_COV_DATABASE_1983_2017_Y2019M07D19.CSV	Geopositioned MERS-CoV Occurrences Database	July 23, 2019
IHME_MERS_COV_DATABASE_1983_2017_INPUT_SOURCES_Y2019M07D19.PDF	Data Input Sources	July 23, 2019
IHME_MERS_COV_DATABASE_1983_2017_INFO_SHEET_Y2019M07D19.PDF	Data Release Information Sheet	July 23, 2019

## Data Files Information

Variable	Variable Label	Variable Definition
nid	NID	A unique identifier for the <a href="#">Global Health Data Exchanged (GHDx)</a> catalog record representing each publication from which data were extracted. The GHDx can be searched by NID to locate the corresponding record
title	Title	Title of the publication
author	Author	Article's author(s)
doi	DOI	Article's DOI
abstract	Abstract	Article's abstract, if available
source_title	Source Title (Journal)	Journal in which the article was published
year	Year	Article's publication year
source	Source (Database)	Database where article was found. Values: pubmed, web of science, scopus
pmid_if_applicable	PMID (if applicable)	PMID if the article is from PubMed
full_text_link_if_included	Full text link (if included)	Link to the full text, if available
file_id	File ID	Reference to pdf in format FirstAuthor_Year (e.g. Smith_2017)

<b>Variable</b>	<b>Variable Label</b>	<b>Variable Definition</b>
occ_id	Occurrence ID	Unique identifier assigned to each occurrence of MERS-CoV. A single pdf may represent more than one occurrence. Each row will have its own occ_id, starting at 1 and numbered consecutively to 882.
organism_type	Organism type	What type of organism tested positive for MERS-CoV. Values: human, mammal, environmental
organism_specific	Organism specific	Specifies the exact organism that tested positive for MERS-CoV. Latin Binomial Names used
pathogen	Pathogen	Name the pathogen identified (e.g. MERS-CoV, Bat Coronaviruses, and other MERS-CoV-like pathogens)
pathogen_note	Pathogen note	Miscellaneous notes regarding pathogen

Variable	Variable Label	Variable Definition
patient_type	Patient type	<p>Patient type values: index, unspecified, secondary, mammal, import, absent</p> <p><u>Index</u>: Any human infection of MERS-CoV resulting after direct contact with an animal and no reported contact with a confirmed MERS-CoV case or healthcare setting.</p> <p><u>Unspecified</u>: Cases that lacked sufficient epidemiological evidence to classify them as any other status (e.g. serosurvey studies)</p> <p><u>NA</u>: Non-applicable field; case was not a patient (e.g. mammal)</p> <p><u>Secondary</u>: Defined as any cases arising from existing human infections. Cases reported after the index case can be assumed to be secondary cases unless accompanied by specific details of likely independent exposure to an animal reservoir.</p> <p><u>Import</u>: Cases that were brought into a non-endemic country after transmission occurred elsewhere</p> <p><u>Absent</u>: Suspected case(s) ultimately confirmed negative for MERS-CoV</p>
transmission_route	Transmission route	<p>Transmission route values: zoonotic, direct, unspecified</p> <p><u>Zoonotic</u>: Transmission occurred directly from an animal</p> <p><u>Direct</u>: Only relevant for human-to-human transmission</p> <p><u>Unspecified</u>: Lacked sufficient epidemiological evidence to classify as any other status</p>
clinical	Clinical	Describes whether the MERS-CoV case was symptomatic, asymptomatic, or suspected
diagnostic	Diagnostic	Describes the class of diagnostic method that was used. PCR, serology, or reported

Variable	Variable Label	Variable Definition
diagnostic_note	Diagnostic note	More detailed information related to the specific test used (e.g., rk39, IgG, or IgM serology)
serosurvey	Serosurvey	Describes the context if serological testing was used. Values: diagnostic, exploratory  <u>Diagnostic</u> : testing of symptomatic patients  <u>Exploratory</u> : historic exposure determined among healthy asymptomatic individuals
country	Country	ISO3 code for country in which the case occurred
origin	Origin	Open-ended field to provide more details on the specific in-country location of MERS-CoV case
problem_geography	Problem Geography	This field was utilized if the MERS-CoV case was reported in a location that could cause uncertainty when determining exact geographic occurrence (e.g. hospital, abattoir).
lat	Latitude	Latitude measured in decimal degrees
long	Longitude	Longitude measured in decimal degrees
latlong_source	Lat/long source	The source from which latitude and longitude were derived
loc_confidence	Location confidence	States the level of confidence that researchers had when assigning a geographic location to the MERS-CoV case (good or bad). An answer of 'good' meant the article stated clearly that the case occurred in a specific geographic location and no assumptions were required on part of the researcher. An answer of 'bad' meant the article did not clearly state the specific geographic location of the MERS-CoV case, but the researcher was able to infer the location of occurrence. The field SITE_NOTES was utilized to detail the logic behind researchers' decisions when inference was required.

Variable	Variable Label	Variable Definition
shape_type	Shape type	The geographic shape type assigned to the MERS-CoV occurrence. Values: point, polygon
poly_type	Polygon type	If the MERS-CoV occurrence was assigned a shape_type of polygon, was it admin (GAUL), custom, or buffer?
buffer_radius	Buffer radius	If a MERS-CoV occurrence was assigned a buffer, what is the radius in km?
gaul_year_or_custom_shapefile	GAUL year or custom shapefile	If a custom polygon was created to geoposition the data, .shp files can be found here: <a href="https://cloud.ihme.washington.edu/index.php/s/DGoyKYqnbjG54F2/download">https://cloud.ihme.washington.edu/index.php/s/DGoyKYqnbjG54F2/download</a>
poly_id	Polygon ID	Within the poly_reference, what is the unique identifier for the necessary polygon?
poly_field	Polygon field	Which type of polygon was used to geo-position the occurrence? (e.g., if admin1 polygon was used, enter ADM1_CODE)
site_notes	Site notes	Miscellaneous notes regarding the site of occurrence
month_start	Month start	Month that the occurrence(s) began. If the article provided a specific month of illness onset, the month was assigned a number from 1-12 (1=January, 2=February, etc.). If the article did not provide a specific month of illness onset, then researchers assigned a value of 'NA'.
month_end	Month end	Month that the occurrence(s) ended, defined as the date a patient tested negative for MERS-CoV. If the article provided a specific month for recovery, the month was assigned a number from 1-12 (1=January, 2=February, etc.). If the article did not provide a specific month of symptom onset, then researchers assigned a value of 'NA'.

Variable	Variable Label	Variable Definition
year_start	Year start	Year that the occurrence(s) began. If the year of illness onset was not provided in the article, the IHME standard was used:  (year_start = publication year – 3)
year_end	Year end	Year that the occurrence(s) ended. If the article did not provide a specific year for recovery, the IHME standard was used:  (year_end = publication year – 1)
year_accuracy	Year accuracy	If years were reported, this field was assigned a value of '0'. If assumptions were required, this field was assigned a value of '1'.

### ***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](mailto: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.