



[www.tycho.pitt.edu](http://www.tycho.pitt.edu)

Wilbert van Panhuis, MD PhD  
[wilbert.van.panhuis@pitt.edu](mailto:wilbert.van.panhuis@pitt.edu)

# Open standardized data, faster discovery

## Unusable data



## Easy-to-use data

Project Tycho® data show dengue epidemic linked to El Niño effect

Project Tycho is a project of the University of Pittsburgh to enhance the availability and use of public health data for science and policy making. Currently, the Project Tycho database includes data from all weekly notifiable disease reports for the United States dating back to 1950. These data are freely available for publicly interested. Additional U.S. and international data will be released in the near future. Please register here to start using these data for your project!

Level 1 data: These data have been verified for immediate use and include a larger number of diseases, locations, and years. These data are not complete because standardization is ongoing.

Level 2 data: These are standardized data for immediate use and include a larger number of diseases, locations, and years. These data are not complete because standardization is ongoing.

Level 3 data: These data have not been standardized at all and cannot be used for analysis. These data are provided upon request.

Current version: 1.0.0

Indicators: cases, deaths, and summary statistics

Locations: 50 states, 50 territories, 3143 counties, and 3226 cities, towns, and parishes

Year: 1950 to 2014

[www.tycho.pitt.edu](http://www.tycho.pitt.edu)

## Research

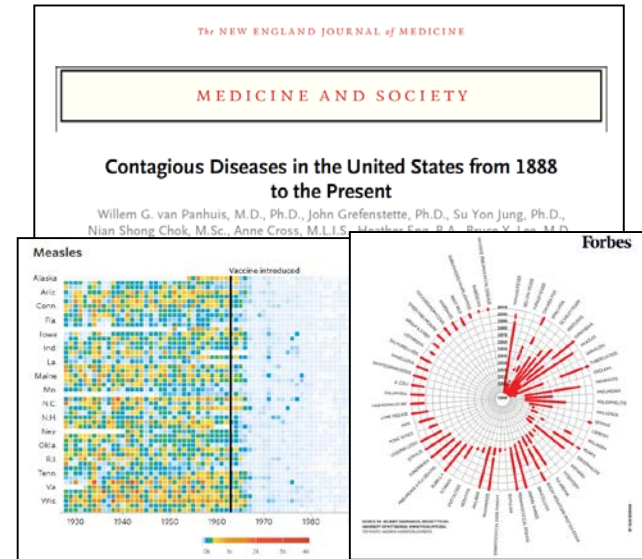
- Digitization
- Extraction
- Standardization
- Integration
- Annotation

Country/Territory	Week 37	Confirmed cases*	Incidence rate†	Deaths	Population ‡
Algeria	Week 37	0	0.0	0	31,200
Anguilla	Week 37	0	0.0	0	68,000
Antigua and Barbuda	Week 37	0	0.0	0	90,000
Aruba	Week 37	0	0.0	0	100,000
Bahamas	Week 37	0	0.0	0	340,000
Belize	Week 37	0	0.0	0	380,000
Bermuda	Week 37	0	0.0	0	240,000
Bonaire	Week 37	0	0.0	0	18,000
Brazil	Week 37	20	0.0	0	190,000,000
Canada	Week 37	199	0.2	0	32,000,000
Cuba	Week 37	0	0.0	0	11,000,000
Curaçao	Week 37	0	0.0	0	400,000
Dominican Republic	Week 37	0	0.0	0	7,000,000
Ecuador	Week 37	0	0.0	0	13,000,000
El Salvador	Week 37	0	0.0	0	5,000,000
French Guiana	Week 37	0	0.0	0	600,000
Guatemala	Week 37	0	0.0	0	14,000,000
Honduras	Week 37	0	0.0	0	6,000,000
Jamaica	Week 37	0	0.0	0	2,500,000
Paraguay	Week 37	0	0.0	0	6,500,000
Peru	Week 37	0	0.0	0	28,000,000
Puerto Rico	Week 37	0	0.0	0	3,500,000
Suriname	Week 37	0	0.0	0	450,000
Turkey	Week 37	0	0.0	0	72,000,000
Uruguay	Week 37	0	0.0	0	3,500,000
Venezuela	Week 37	0	0.0	0	24,000,000

TABLE II. (Part 6) Provisional cases of selected notifiable diseases, United States, week ending March 9, 2013 (WEEK 10)\*

Reporting Area	Legionellosis			Listeria monocytogenes			Malaria			
	Current week	Previous 52 weeks	Case/100,000	Current week	Previous 52 weeks	Case/100,000	Current week	Previous 52 weeks	Case/100,000	
UNITED STATES	27	62	133	34	407	27	435	1,433	1,034	2,538
NEW ENGLAND	5	19	38	22	114	407	59	179	2	7
Midwest	1	7	14	4	19	182	3	200	0	2
South	0	3	2	1	16	69	28	74	0	2
West	0	1	1	1	10	22	1	11	0	4
Arizona	0	1	1	1	10	22	1	11	0	4
California	0	1	1	1	10	22	1	11	0	4
Colorado	0	1	1	1	10	22	1	11	0	4
Hawaii	0	1	1	1	10	22	1	11	0	4
Idaho	0	1	1	1	10	22	1	11	0	4
Montana	0	1	1	1	10	22	1	11	0	4
Nevada	0	1	1	1	10	22	1	11	0	4
New Mexico	0	1	1	1	10	22	1	11	0	4
Oregon	0	1	1	1	10	22	1	11	0	4
Utah	0	1	1	1	10	22	1	11	0	4
Washington	0	1	1	1	10	22	1	11	0	4
Wyoming	0	1	1	1	10	22	1	11	0	4
ALABAMA	1	12	14	102	22	197	457	1,240	1	4
ALASKA	0	0	0	0	0	0	0	0	0	0
ARIZONA	2	12	5	19	10	166	144	442	0	2
ARKANSAS	0	6	13	14	23	18	16	232	192	140
CALIFORNIA	1	2	13	14	1	12	1	12	1	3
COLORADO	4	19	28	34	6	46	40	379	633	1
CONNECTICUT	0	13	14	67	96	23	168	39	143	2
DELAWARE	0	2	3	1	2	13	2	8	1	3
FLORIDA	1	2	6	26	33	1	6	5	0	2
GEORGIA	2	9	14	14	2	8	5	6	0	3
IDAHO	3	5	13	18	1	4	8	6	0	1
ILLINOIS	1	7	4	17	138	15	136	0	0	2
INDIANA	3	10	13	24	2	18	2	5	1	3
IOWA	0	2	6	1	18	1	3	1	0	2
KANSAS	0	2	1	1	0	2	1	0	0	1

US data:  
1888-present  
WHO dengue  
data: 1952-2010



Wall street journal Forbes magazine

# Project Tycho version 1

University of Pittsburgh

Home About Data Resources FAQ Contact Us

**PROJECT TYCHO**  
DATA FOR HEALTH

Project Tycho™ makes data available for use by others

**Level 1 data**  
These data have been custom tailored for specific projects. These are the most complete data but for a limited number of diseases, locations, and years.  
Current version: 1.0.0  
Indicators: cases, incidence rates  
Diseases: measles, polio, mumps, rubella, hepatitis A, diphtheria, and whooping cough  
Locations: all 50 states and 122 cities (diphtheria only)  
Years: varying between 1916 and 2009  
[Go to level 1 data](#)

**Level 2 data**  
These data have been cleaned for immediate use, including a large number of diseases, locations, and years. Data for some diseases, locations, and years have not been cleaned yet, resulting in gaps.  
Current version: 1.0.0  
Indicators: cases and deaths  
Diseases: 47 contagious diseases  
Locations: all 50 states, 6 territories, and 1282 cities  
Years: varying between 1888 to 2013  
[Go to level 2 data](#)

**Level 3 data**  
These data have not been cleaned and cannot yet be used for analysis. We are continuing data cleaning and would welcome help and support for this. Raw data are provided upon request.  
Current version: 1.0.0  
Indicators: cases, deaths, and summary statistics  
Diseases: 56 contagious diseases and 72 subcategories  
Locations: 10 regions, 50 states, 6 territories, 3165 counties, and 2906 cities, towns, and parishes  
Years: varying between 1887 to 2013  
[Level 3 data request form](#)

Project Tycho™ is funded by the Bill & Melinda Gates Foundation and the National Institutes of Health  
© 2013 Public Health Dynamics Laboratory, University of Pittsburgh

The NEW ENGLAND JOURNAL of MEDICINE

MEDICINE AND SOCIETY 2013

## Contagious Diseases in the United States from 1888 to the Present

Willem G. van Panhuis, M.D., Ph.D., John Grefenstette, Ph.D., Su Yon Jung, Ph.D., Nian Shong Chok, M.Sc., Anne Cross, M.L.I.S., Heather Eng, B.A., Bruce Y. Lee, M.D., Vladimir Zadorozhny, Ph.D., Shawn Brown, Ph.D., Derek Cummings, Ph.D., M.P.H., and Donald S. Burke, M.D.

*Weekly US surveillance data: 1888-2014*

### Level 1 data

- Data from NEJM paper
- Imputed
- Population and rates
- 8 conditions
- 1916-2009
- GUI

### Level 2 data

- All “consistent” data
- 50 conditions
- 1888-2014
- GUI or API

### Level 3 data

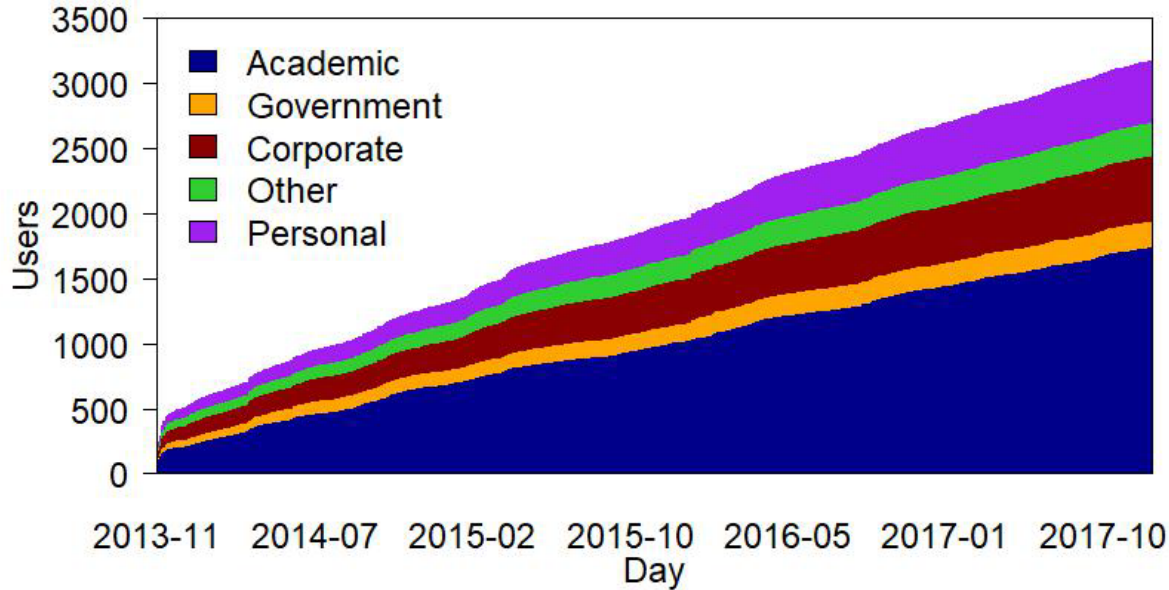
- Everything
- Duplicates, unknown condition, etc.
- 58 conditions, many subcategories
- 1887-2014
- Upon request

### In development

- Dengue from WHO for 99 countries
- Dengue from 8 SE Asia countries
- Chikungunya data from PAHO
- GUI

# Reuse of Project Tycho v1 data

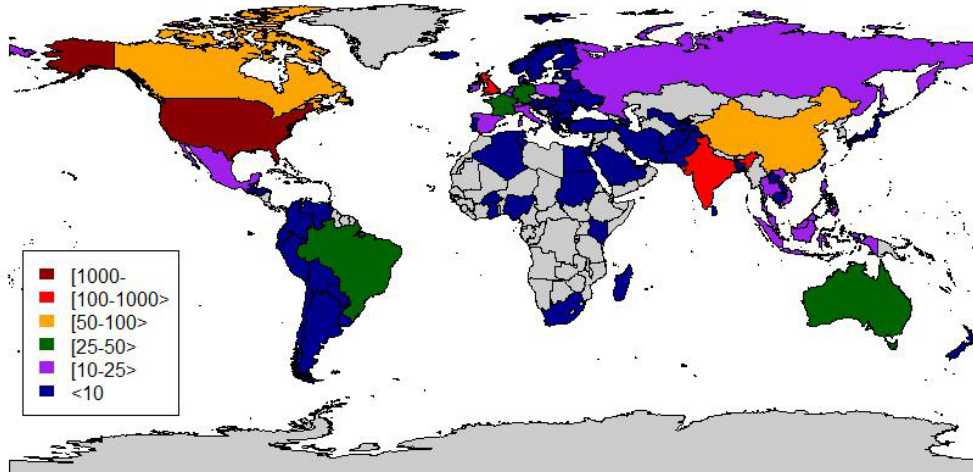
Cumulative number of registered users



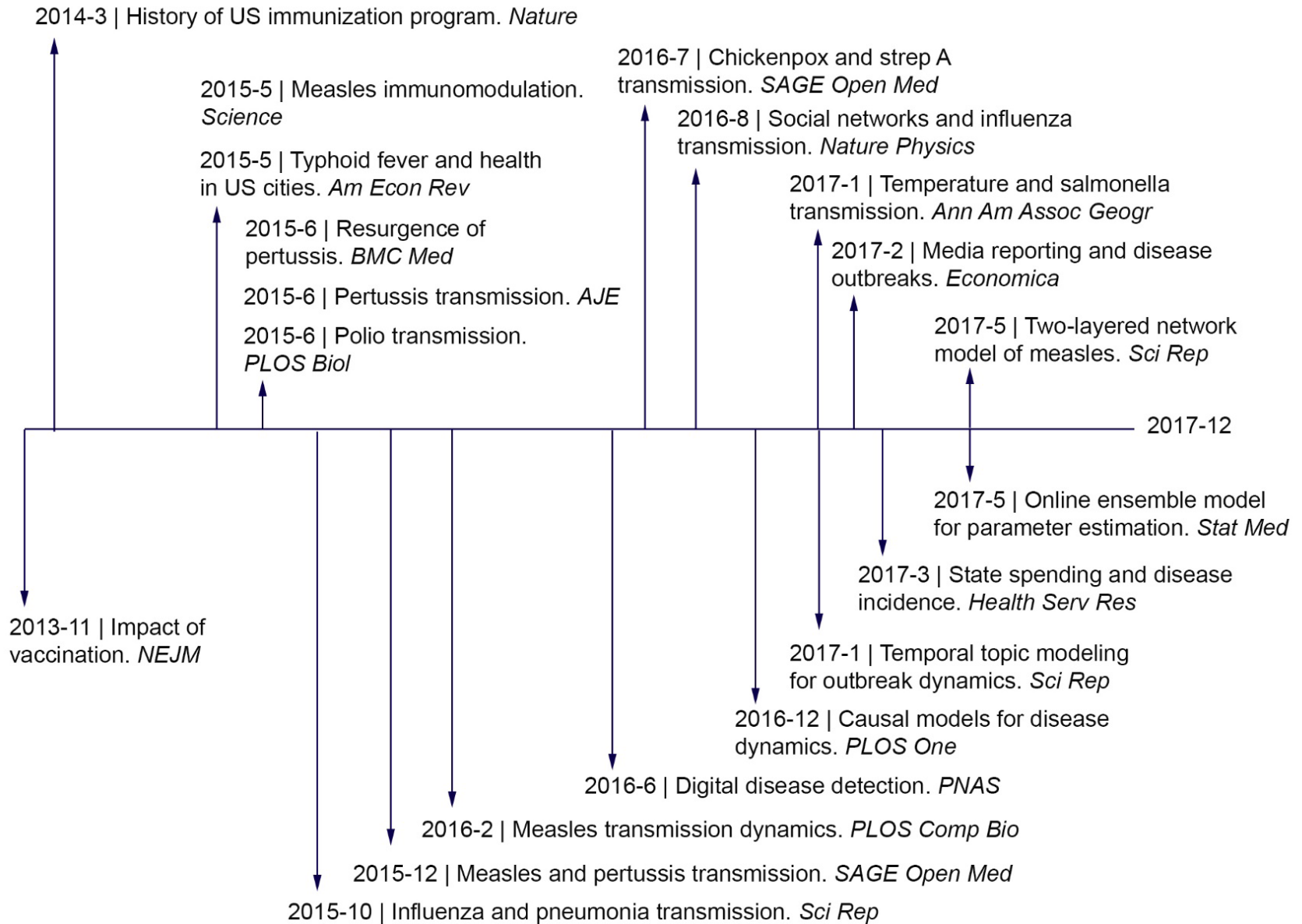
**3500+** users  
**7000** downloads  
**1 million** API calls  
**272,000** page visits

**50 Creative works by others:**

- 18 peer-reviewed papers
- 9 other papers
- 4 theses
- 19 visualizations/  
software applications
  - Infographics in newspapers
  - R packages
  - Code on Github
  - Visualization applications



# Scientific papers using v1 data



# New in Project Tycho 2.0

---

## NEW DATA ADDED

- US data updated until 31 December 2017
- Integration of dengue surveillance data with US data (no longer “in development”)
  - Annual/monthly aggregate counts from WHO online databases and from 8 partner countries in SE Asia

## FINDABLE, INTEROPERABLE, REUSABLE (FAIR)

- Data grouped in 360 datasets, one for each condition-country combination
  - Digital object identifiers and metadata files
- Extensive standardization
  - Standards for condition, pathogen, location
  - Standard data format for all Project Tycho data

# SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES

- » Research data
- » Publication characteristics

## Comment: The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson *et al.*<sup>#</sup>

**Findable:** Digital object identifier, data indexed in online catalogues

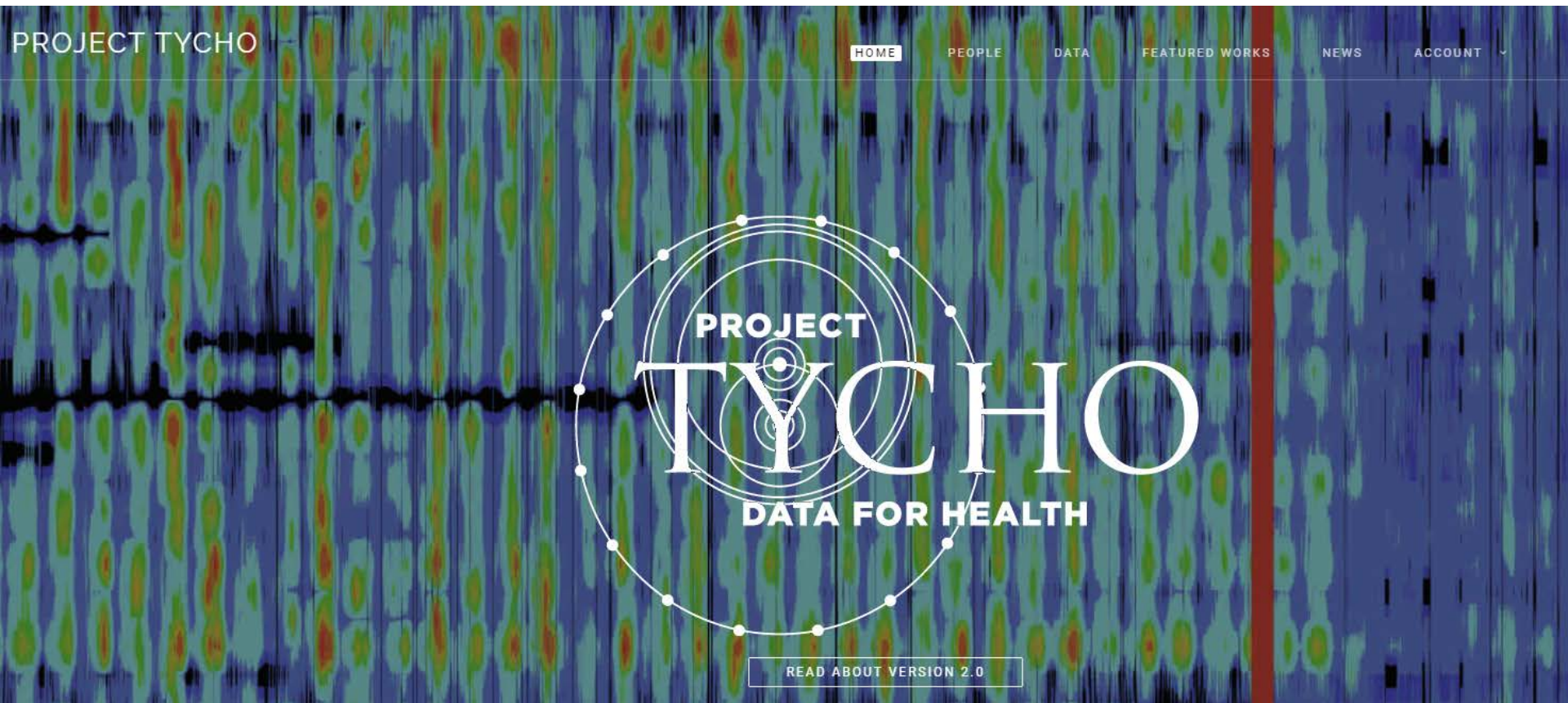
**Accessible:** Process for data retrieval clear and no proprietary software

**Interoperable:** Standard data format, standard ontologies/vocabularies used to describe the data and metadata

**Reusable:** Rich metadata to help determine reuse options

FAIR -> Machine-interpretable -> Faster discovery

New website: [www.tycho.pitt.edu](http://www.tycho.pitt.edu)





# Continued access to previous versions

---

## PREVIOUS VERSIONS

DOWNLOAD A FLAT FILE OF A PREVIOUS VERSION

VERSION 1.0.0

### Level 1 Data

Used for peer-reviewed research

8 diseases

50 US states and 122 US cities

1916 to 2009

DOWNLOAD

VERSION 1.1.0

### Level 2 Data

Not peer-reviewed

50 diseases

50 US states and 1284 US cities

1888 to 2014

DOWNLOAD

*Not yet included in version 2.0!*

*More and better in version 2.0*

# Three ways to access the data

---

## ACCESS DATA

EXPLORE AND DOWNLOAD DATA

### DATASETS

## Pre-Compiled

All data for one condition in one country  
compiled in one dataset

Select from 360 datasets: 92 conditions for the  
US and dengue-related conditions for countries  
around the world

Each dataset has a digital object identifier and  
standard citation

START NOW!

### DATA QUERY

## Compile Your Own

Search data for any condition and location

Visualize and compare data for multiple  
conditions or locations

Download selected data in Project Tycho  
standard format

START NOW!

### DATA API

## Application Programming Interface

Enables data access for computer programs

Search data for any possible combination of  
countries, conditions, and other variables

Download selected data in Project Tycho  
standard format

START NOW!

# Pre-compiled US datasets: 92

## PRE-COMPILED US DATASETS

DATASETS FOR 92 CONDITIONS IN THE UNITED STATES

Show  entries

Search:

Country name	Condition	DOI	# Records	Action
United States of America	Amebic dysentery	10.25337/T7/ptycho.v2.0/US.387754006	11708	<a href="#">↓</a>
United States of America	Anthrax	10.25337/T7/ptycho.v2.0/US.409498004	7100	<a href="#">↓</a>
United States of America	Babesiosis	10.25337/T7/ptycho.v2.0/US.21061004	3300	<a href="#">↓</a>
United States of America	Bacillary dysentery	10.25337/T7/ptycho.v2.0/US.274081004	11094	<a href="#">↓</a>
United States of America	Brucellosis	10.25337/T7/ptycho.v2.0/US.75702008	20002	<a href="#">↓</a>
United States of America	Chlamydia trachomatis infection	10.25337/T7/ptycho.v2.0/US.240589008	7049	<a href="#">↓</a>
United States of America	Chlamydial infection	10.25337/T7/ptycho.v2.0/US.105629000	74097	<a href="#">↓</a>
United States of America	Cholera	10.25337/T7/ptycho.v2.0/US.63650001	101	<a href="#">↓</a>
United States of America	Coccidioidomycosis	10.25337/T7/ptycho.v2.0/US.60826002	10678	<a href="#">↓</a>
United States of America	Cryptosporidiosis	10.25337/T7/ptycho.v2.0/US.240370009	54082	<a href="#">↓</a>

Showing 1 to 10 of 92 entries

< Previous  2 3 4 5 ... 10 Next >

# Pre-compiled dengue datasets: 271

## PRE-COMPILED DENGUE DATASETS

DATASETS FOR DENGUE-RELATED CONDITIONS FOR 99 COUNTRIES

Show  entries

Search:

Country name	Condition	DOI	# Records	Action
American Samoa	Dengue	10.25337/T7/ptycho.v2.0/AS.38362002	49	<a href="#">Download</a>
American Samoa	Dengue hemorrhagic fever	10.25337/T7/ptycho.v2.0/AS.20927009	3	<a href="#">Download</a>
American Samoa	Dengue without warning signs	10.25337/T7/ptycho.v2.0/AS.722862003	18	<a href="#">Download</a>
Anguilla	Dengue	10.25337/T7/ptycho.v2.0/AI.38362002	61	<a href="#">Download</a>
Anguilla	Dengue hemorrhagic fever	10.25337/T7/ptycho.v2.0/AI.20927009	19	<a href="#">Download</a>
Anguilla	Dengue without warning signs	10.25337/T7/ptycho.v2.0/AI.722862003	24	<a href="#">Download</a>
Antigua and Barbuda	Dengue	10.25337/T7/ptycho.v2.0/AG.38362002	69	<a href="#">Download</a>
Antigua and Barbuda	Dengue hemorrhagic fever	10.25337/T7/ptycho.v2.0/AG.20927009	18	<a href="#">Download</a>
Antigua and Barbuda	Dengue without warning signs	10.25337/T7/ptycho.v2.0/AG.722862003	34	<a href="#">Download</a>
Argentina	Dengue	10.25337/T7/ptycho.v2.0/AR.38362002	129	<a href="#">Download</a>

Showing 1 to 10 of 271 entries

< Previous

1

2

3

4

5

...

28

Next >

# Dataset landing page

---



## DOI

10.25337/T7/ptycho.v2.0/KH.38362002



## DOWNLOAD LINKS

Dataset contains: 8889 records  
Download data and readme in Project Tycho  
format: [KH.38362002.zip](#)  
Download metadata in DATS format:  
[KH.38362002.json](#)  
Download metadata in DataCite XML format:  
[KH.38362002.xml](#)



## CONDITION AND PATHOGEN INFORMATION

SNOMED-CT name: Dengue  
SNOMED-CT code: [38362002](#)  
Pathogen name: Dengue virus  
Pathogen code: [12637](#)



## SPACE AND TIME INFORMATION

Locations included: [CAMBODIA](#)  
Apollo Location Service: [CAMBODIA](#)  
Earliest year: 1980  
Latest year: 2011



## CITATION

Van Panhuis W., Cross A., Burke D., Choisy M., Counts of Dengue reported in  
CAMBODIA: 1980-2011 (version 2.0, April 1, 2018): Project Tycho data release,  
DOI: [10.25337/T7/ptycho.v2.0/KH.38362002](#)

# Custom-compiled datasets (GUI)

**RECORDS SELECTED**  
0

**SERIES SELECTED**  
0

Select values for each of the variables shown below; hold Control (on Windows) or ⌘ (on Mac) to select multiple values for a variable.

**CONDITION:**  
Congenital syphilis : 3927  
Cryptosporidiosis : 54082  
Dengue : 83892  
Dengue hemorrhagic fever : 10798  
Dengue without warning signs : 16331  
Diphtheria : 370739  
Disease caused by West Nile virus : 611  
Disorder of nervous system caused by West Nile virus : 11758  
Dysentery : 10444  
Encephalitis : 6736  
Encephalitis lethargica : 29711  
Giardiasis : 49482  
Gonorrhea : 163739  
Haemophilus influenzae infection : 59450  
Haemophilus influenzae type b infection : 919  
Hepatitis non-A non-B : 34073

**COUNTRIES:**  
AMERICAN SAMOA : 70  
ANGUILLA : 104  
ANTIGUA AND BARBUDA : 121  
ARGENTINA : 244  
ARUBA : 49  
AUSTRALIA : 72  
BAHAMAS : 122  
BANGLADESH : 67  
BARBADOS : 117  
BELIZE : 104  
BERMUDA : 72  
BHUTAN : 46  
BOLIVIA (PLURINATIONAL STATE OF) : 120  
BRAZIL : 2068  
BRUNEI DARUSSALAM : 31  
CAMBODIA : 11932  
CANADA : 6

**PATHOGEN:**  
Anaplasma phagocytophilum : 8464  
Babesiidae : 3300  
Bacillus anthracis : 7100  
Borrelia : 51254

**PLACE OF ACQUISITION:**  
Abroad : 17043  
Domestic : 23331  
NA : 5636050

**SOURCE NAME:**  
Cambodia Dengue Surveillance System : 3677  
Laos Dengue Surveillance System : 2627  
Malaysia Dengue Surveillance System : 1944  
Pan American Health Organization Website : 1058

**ADMIN LEVEL 1:**  
No options available based on your set filters

**ADMIN LEVEL 2:**  
No options available based on your set filters

**ADMIN LEVEL 3:**  
No options available based on your set filters

**CITY:**  
No options available based on your set filters

**INTERVAL TYPE:**  
Month : 93208  
Week : 5575117  
Year : 7863

**CUMULATIVE:**  
No : 3836899  
Yes : 1839525

**FATALITIES:**  
No : 4859865  
Yes : 816559

**AGE RANGE:**  
0-130 : 5641972  
0-14.99 : 94  
0-19.99 : 6048  
0-4.99 : 25688

**START YEAR:**  
1888  
1889  
1890  
1891

**END YEAR:**  
1888  
1889  
1890  
1891

**DIAGNOSTIC CERTAINTY:**  
Definite : 10438  
NA : 5654828  
Probable diagnosis : 11158

**PLOT PREVIEW:**  
To preview a plot please select:  

- At least 1 condition
- At least 1 source name
- The type of interval
- The type of cumulative count
- The records selected is greater than 0

**RESET ALL FILTERS** | **DOWNLOAD** | **DETAILED PLOT**

1. Explore data content
2. Download a custom datasets
3. Visualize data time series

# GUI – Explore data content

## CONDITION:

Acquired immune deficiency syndrome : 11480  
Active tuberculosis : 11480  
Acute hepatitis C : 671  
Acute nonparalytic poliomyelitis : 194  
Acute paralytic poliomyelitis : 194  
Acute poliomyelitis : 137581  
Acute type A viral hepatitis : 4105  
Acute type B viral hepatitis : 4127  
Amebic dysentery : 11708  
Anthrax : 7100  
Aseptic meningitis : 41828  
Babesiosis : 3300  
Bacillary dysentery : 11094  
Brucellosis : 20002  
Campylobacteriosis : 4486  
Chlamydia trachomatis infection : 74097  
Chlamydia infection : 74097

## COUNTRIES:

BRAZIL : 2068  
BRUNEI DARUSSALAM : 31  
CAMBODIA : 11932  
CANADA : 6  
CAYMAN ISLANDS : 75  
CHILE : 40  
CHINA : 7060  
COLOMBIA : 159  
COOK ISLANDS : 74  
COSTA RICA : 97  
CUBA : 831  
CURAÇAO : 25  
DOMINICA : 116  
DOMINICAN REPUBLIC : 149  
ECUADOR : 101  
EL SALVADOR : 5184  
FIJI : 102

# GUI – Explore data content

## CONDITION:

Dengue : 8889  
Dengue hemorrhagic fever : 1543  
Dengue without warning signs : 1500

## COUNTRIES:

BOLIVIA (FEDERATIONAL STATE OF) : 120  
BRAZIL : 2068  
BRUNEI DARUSSALAM : 31  
**CAMBODIA : 94 [11838 in admin level below]**  
CANADA : 6  
CAYMAN ISLANDS : 75  
CHILE : 40  
CHINA : 7060  
COLOMBIA : 159  
COOK ISLANDS : 74  
COSTA RICA : 97  
CUBA : 831  
CURAÇAO : 25  
DOMINICA : 116  
DOMINICAN REPUBLIC : 149  
ECUADOR : 101  
EL SALVADOR : 5184  
FIJI : 102

## SOURCE NAME:

Cambodia Dengue Surveillance System : 3677  
World Health Organization DengueNet Database :  
World Health Organization Western Pacific Region  
World Health Organization Western Pacific Region

## ADMIN LEVEL 1:

BÂNTÉAY MÉANCHEY : 507  
BĂTDÂMBÂNG : 515  
KAÔH KÔNG : 506  
KRÂCHÉH : 487  
KRÔNG KÊP : 405



# GUI – Explore data content

## CONDITION:

Dengue : 8889  
Dengue hemorrhagic fever : 1543  
Dengue without warning signs : 1500

## COUNTRIES:

BOLIVIA (FEDERATIONAL STATE OF) : 120  
BRAZIL : 2068  
BRUNEI DARUSSALAM : 31  
**CAMBODIA : 94 [11838 in admin level below]**  
CANADA : 6  
CAYMAN ISLANDS : 75  
CHILE : 40  
CHINA : 7060  
COLOMBIA : 159  
COOK ISLANDS : 74  
COSTA RICA : 97  
CUBA : 831  
CURAÇAO : 25  
DOMINICA : 116  
DOMINICAN REPUBLIC : 149  
ECUADOR : 101  
EL SALVADOR : 5184  
FIJI : 102

## SOURCE NAME:

Cambodia Dengue Surveillance System : 3677  
World Health Organization DengueNet Database :  
World Health Organization Western Pacific Region  
World Health Organization Western Pacific Region

## ADMIN LEVEL 1:

**BÂNTÉAY MÉANCHEY : 507**  
BÂTDĂMBĂNG : 515  
KAÔH KŌNG : 506  
KRÂCHÉH : 487  
KRŌNG KĒP : 405

# GUI – Explore data content

COUNTRIES:

CAMBODIA : 0 [507 in admin level below]

PATHOGEN:

Dengue virus : 507

PLACE OF ACQUISITION:

NA : 507

SUBPOPULATION:

None specified : 507

DIAGNOSTIC CERTAINTY:

NA : 507

FATALITIES:

No : 399  
Yes : 108

AGE RANGE:

0-130 : 507

START YEAR:

1998  
1999  
2000  
2001  
2002

END YEAR:

1998  
1999  
2000  
2001  
2002

ADMIN LEVEL 1:

BĂNTÉAY MÉANCHEY : 507

BĂTDĂMBĂNG : 515

KAÔH KÔNG : 506

KRĂCHÉH : 487

KRĂNG KĒP : 495

ADMIN LEVEL 2:

No options available based on your set filters

PLOT PREVIEW:

To preview a plot please select:

At least 1 condition

At least 1 source name

The type of interval

The type of cumulative count

# GUI – Download custom data

<p>CONDITION:</p> <ul style="list-style-type: none"><li>Meningococcal meningitis : 126493</li><li>Mumps : 165242</li><li>Murine typhus : 2786</li><li>Ornithosis : 824</li><li>Pellagra : 29738</li><li>Pertussis : 297794</li><li>Pneumonia : 198788</li><li>Post-infectious encephalitis : 12918</li><li>Primary encephalitis : 26152</li><li>Rocky Mountain spotted fever : 71024</li><li>Rubella : 72676</li><li>Salmonella infection : 71824</li><li>Scarlet fever : 351882</li><li>Shigellosis : 63790</li><li>Smallpox : 260066</li><li>Smallpox without rash : 19</li><li>Spotted fever group rickettsial disease :</li></ul>	<p>COUNTRIES:</p> <ul style="list-style-type: none"><li>UNITED STATES OF AMERICA : 0 [71024 i</li></ul>
<p>SOURCE NAME:</p> <ul style="list-style-type: none"><li>US Nationally Notifiable Disease Surveilla</li></ul>	<p>ADMIN LEVEL 1:</p> <p>No options available based on your set fil</p>
<p>INTERVAL TYPE:</p> <ul style="list-style-type: none"><li>Week : 71021</li></ul>	<p>ADMIN LEVEL 2:</p> <p>No options available based on your set fil</p>

# GUI – Download custom data

**CONDITION:**

- Pneumonia : 5423
- Post-infectious encephalitis : 509
- Primary encephalitis : 1062
- Rocky Mountain spotted fever : 4615
- Rubella : 2129
- Salmonella infection : 3224
- Scarlet fever : 10613
- Shigellosis : 3050
- Smallpox : 10088
- Spotted fever group rickettsial disease : 369
- Streptococcal sore throat : 157
- Syphilis : 2
- Tetanus : 1337
- Toxic shock syndrome : 346
- Tuberculosis : 10532
- Tularemia : 2919

**COUNTRIES:**

UNITED STATES OF AMERICA : 0 [4615 in admin level below]

**SOURCE NAME:**

US Nationally Notifiable Disease Surveillance System : 4615

**ADMIN LEVEL 1:**

- ALABAMA : 2384 [4 in admin level below]
- ALASKA : 29
- ARIZONA : 873
- ARKANSAS : 2227
- CALIFORNIA : 1665

**INTERVAL TYPE:**

Week : 4615

**ADMIN LEVEL 2:**

No options available based on your set filters

**CUMULATIVE:**

No : 1193  
Yes : 3422

**CITY:**

No options available based on your set filters

# GUI – Visualize data

---

## PLOT PREVIEW:

To preview a plot please select:

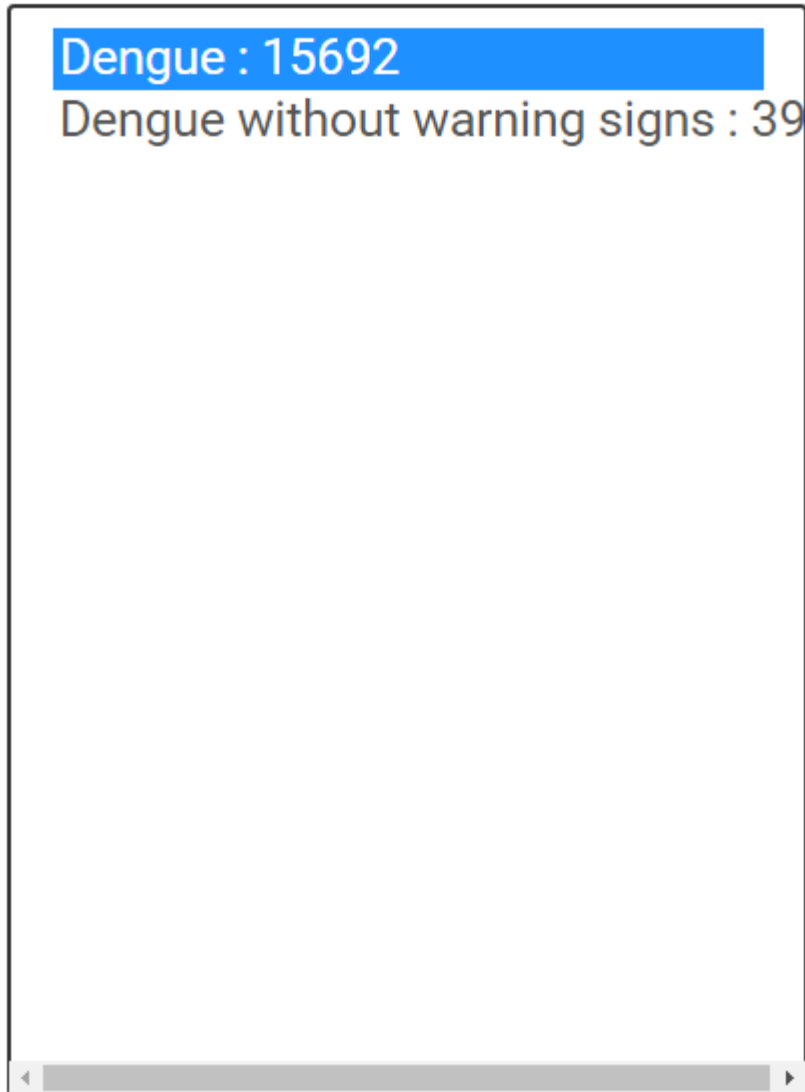
At least 1 condition	At least 1 condition
At least 1 source name	At least 1 source
The type of interval	Type of interval: weekly, monthly, yearly
The type of cumulative count	Cumulative or not
The records selected is greater than 0	Number of records > 0

Need to select sufficient information to get a time series of equivalent counts, e.g, not mixing up different age groups, cumulative and non-cumulative counts, etc.

# GUI – Visualize data

---

## CONDITION:



## COUNTRIES:



# GUI – Visualize data

---

SOURCE NAME:

Thailand Dengue Surveillance System : 64

INTERVAL TYPE:

Month : 648

CUMULATIVE:

No : 648

ADMIN LEVEL 1:

PHUKET : 216  
PRACHIN BURI : 216  
PRACHUAP KHIRI KHAN : 216  
RANONG : 216  
RATCHABURI : 216

ADMIN LEVEL 2:

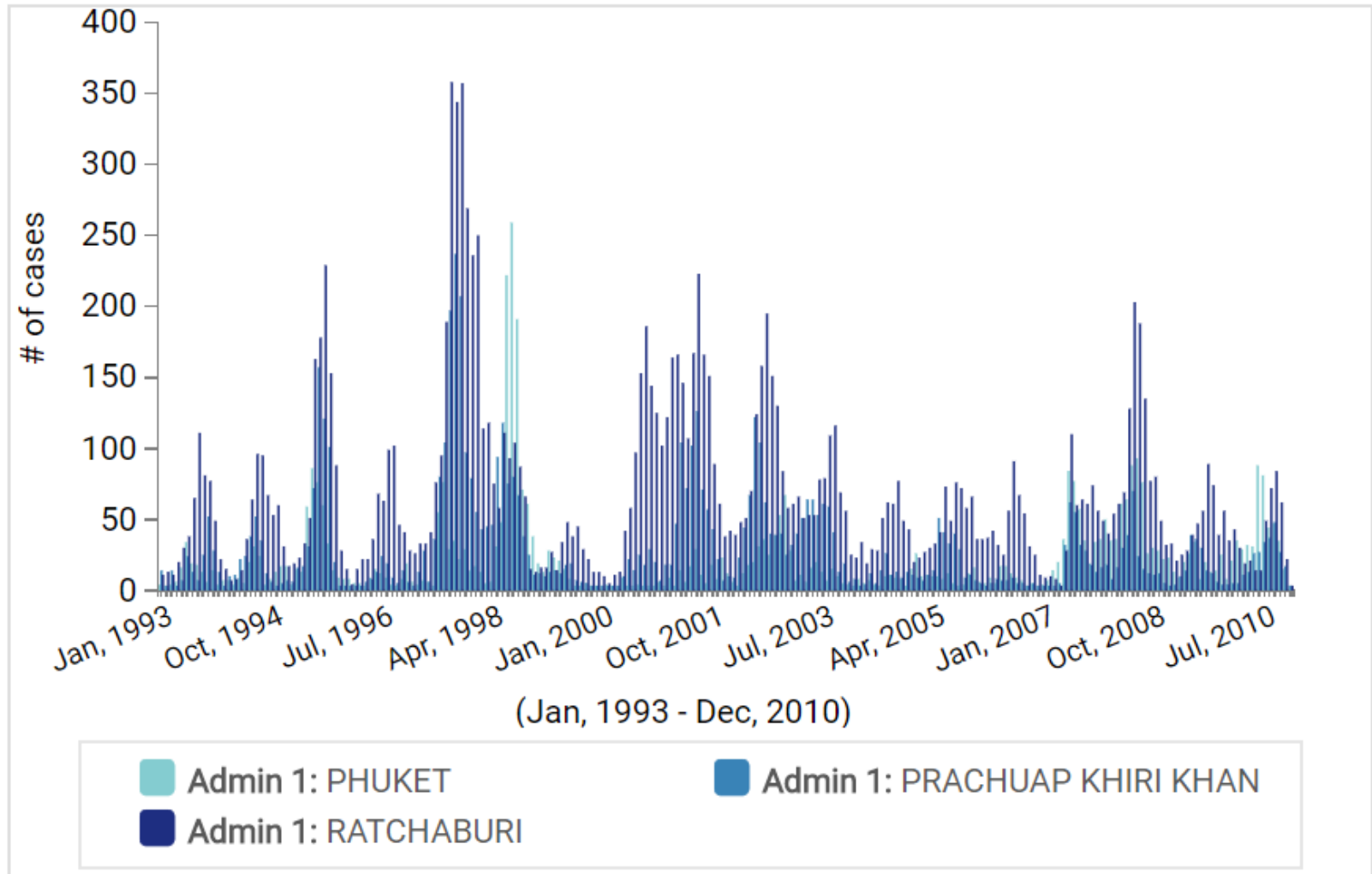
No options available based on your set filter

CITY:

No options available based on your set filter

# GUI – Visualize data

PLOT PREVIEW:





# GUI – Visualize data

SOURCE NAME:

Thailand Dengue Surveillance System : 64

INTERVAL TYPE:

Month : 648

CUMULATIVE:

No : 648

ADMIN LEVEL 1:

SONGKHLA : 216  
SUKHOTHAI : 216  
SUPHAN BURI : 216  
SURAT THANI : 216  
SURI : 216

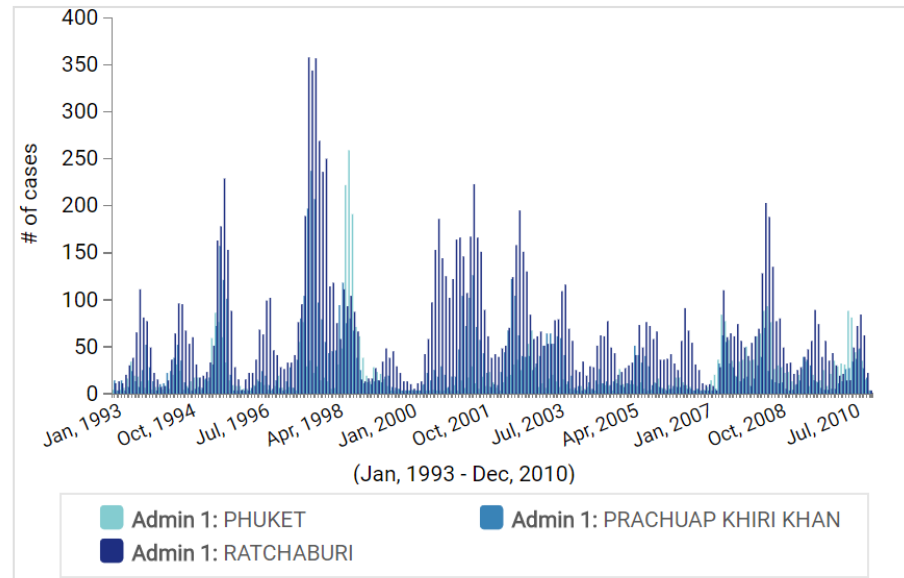
ADMIN LEVEL 2:

No options available based on your set filter

CITY:

No options available based on your set filter

PLOT PREVIEW:



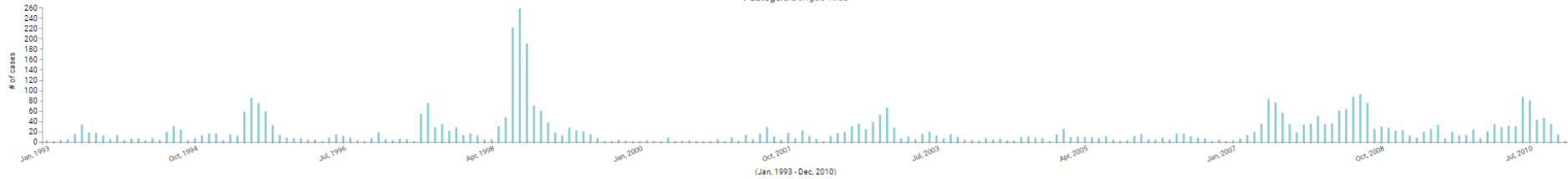
⌂ RESET ALL FILTERS

⬇️ DOWNLOAD

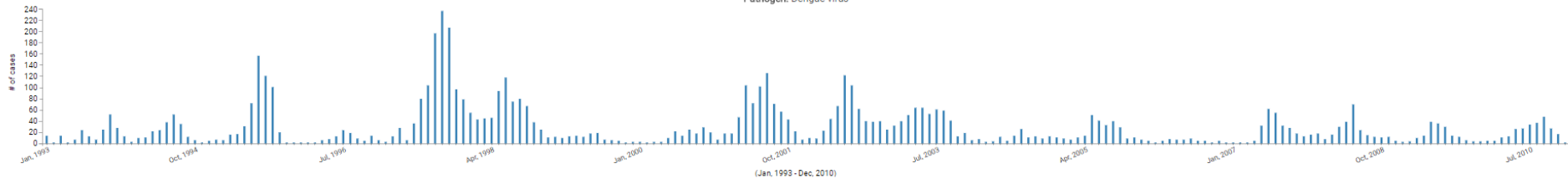
📊 DETAILED PLOT

# GUI – Visualize data

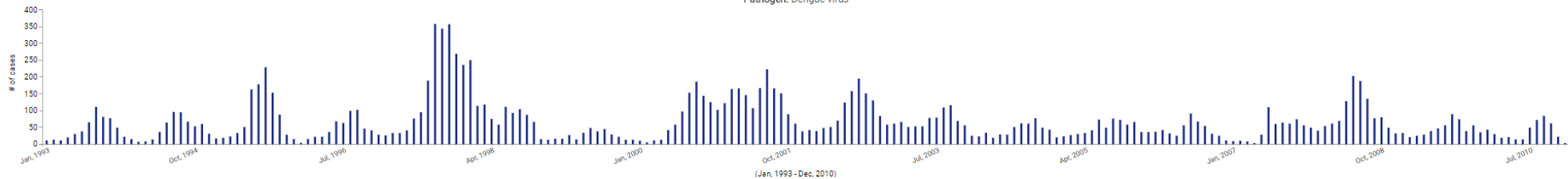
Country: THAILAND Admin 1: PHUKET Condition: Dengue Fatalities: No Diagnostic Certainty: NA Age Range: 0-130 Subpopulation: None specified  
Place of Acquisition: NA Source: Thailand Dengue Surveillance System Cumulative: No Interval Type: Month  
Pathogen: Dengue virus



Country: THAILAND Admin 1: PRACHUAP KHIRI KHAN Condition: Dengue Fatalities: No Diagnostic Certainty: NA Age Range: 0-130  
Subpopulation: None specified Place of Acquisition: NA Source: Thailand Dengue Surveillance System Cumulative: No Interval Type: Month  
Pathogen: Dengue virus



Country: THAILAND Admin 1: RATCHABURI Condition: Dengue Fatalities: No Diagnostic Certainty: NA Age Range: 0-130  
Subpopulation: None specified Place of Acquisition: NA Source: Thailand Dengue Surveillance System Cumulative: No Interval Type: Month  
Pathogen: Dengue virus

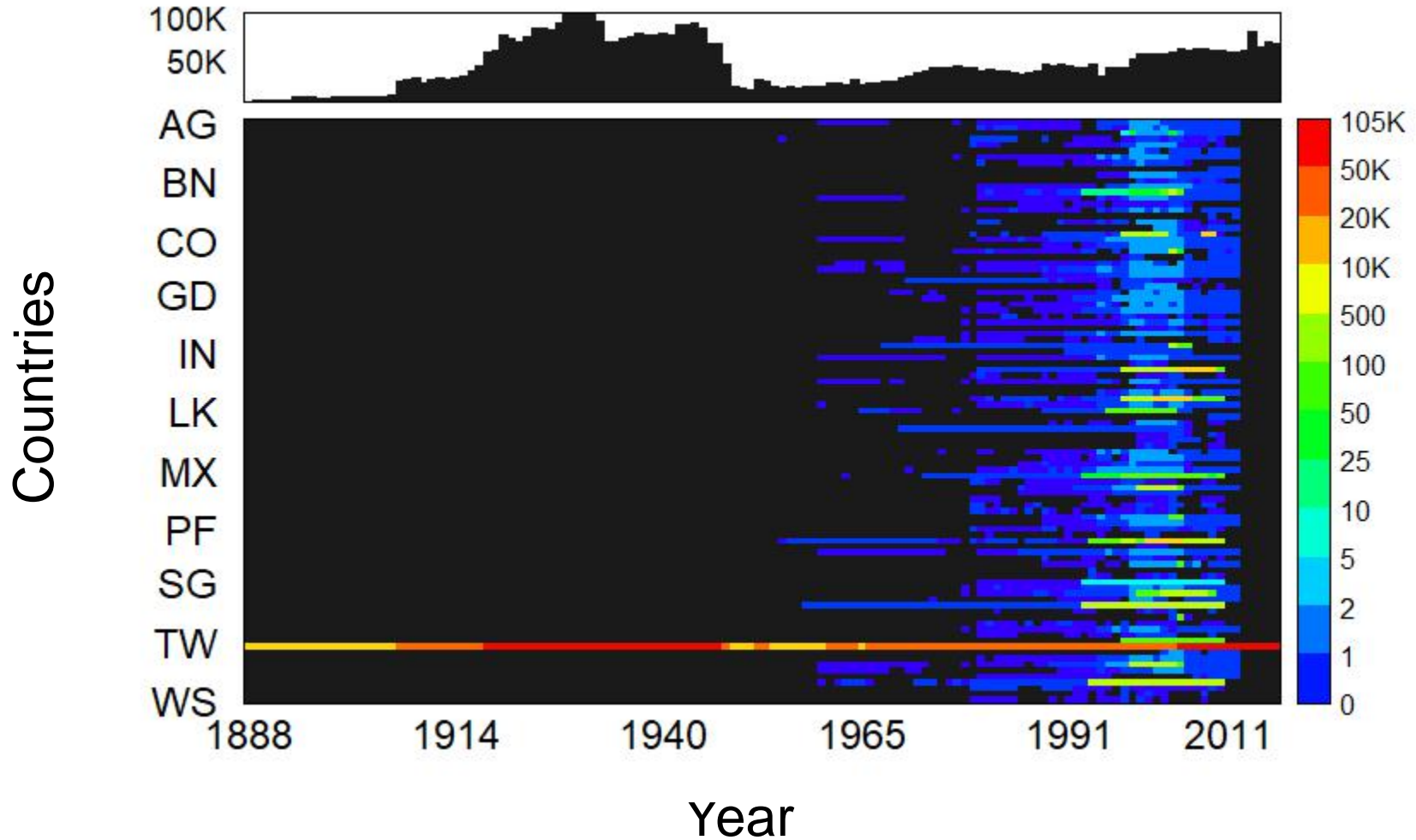


# Standard Project Tycho data format (CSV)

<b>ConditionName*</b>	CityName
<b>ConditionSNOMED*</b>	<b>PeriodStartDate*</b>
PathogenName	<b>PeriodEndDate*</b>
PathogenTaxonID	<b>PartOfCumulativeCountSeries*</b>
<b>Fatalities*</b>	<b>AgeRange*</b>
<b>CountryName*</b>	<b>Subpopulation*</b>
CountryISO	PlaceOfAquisition
Admin1Name	DiagnosisCertainty
Admin1ISO	<b>SourceName*</b>
Admin2Name	<b>CountValue*</b>



# Available data by country and year



# Next steps

---

- Expand data: online surveillance data, data from research/country partners, age-specific data, climate/demography/vaccination data
- Improve online graphical user interface: selecting data, visualization and analytics
- Integration of research data (level 1) with data from health agencies: Augmented datasets
  - Represent data collection/pre-processing methods
  - Data identifiers and metadata
  - User interface and download options
- Data contributions by others (repository function)
- Sustainable business model

# Acknowledgements

---



Donald Burke, MD  
Dean, Graduate  
School of Public  
Health



Mike Wagner, MD PHD  
Professor of Biomedical  
Informatics and Intelligent  
Systems



Anne Cross, MLIS  
Database manager



John Levander  
Software design,  
MIDAS ISG



Max Sibilla  
Software design,  
MIDAS ISG



Tejaswi Anataraju, MS  
Data programmer



Sharon Crow, MEd  
Project Coordinator

## Current Faculty, Staff, and Students

Marc Choisy, Institute de la Recherche pour le Développement; Derek Cummings, University of Florida; Ernesto Marques, University of Pittsburgh; Leah Goeke, Kelly Carey, Matt Loiacacono, Austin Chen, Lingshu Xue, Tenley Brownwright, Hannah Polglase, Christian Garcia, Angel Paternina, Danielle Seigneur, Jessica Salerno, Vivian Lin, Rhonda Toth, Scott Chadwell

## Partners

Laos Ministry of Health; Taiwan Ministry of Health; Thailand Ministry of Public Health; Cambodia Ministry of Health, Vietnam National Institute for Hygiene and Epidemiology; Philippines Ministry of Health; Singapore National Environment Agency; Malaysia Ministry of Health; Brazil Ministry of Health; U.S. Department of Health & Human Services; US Centers for Disease Control; World Health Organization;

## Funders

**BILL & MELINDA**  
**GATES** *foundation*

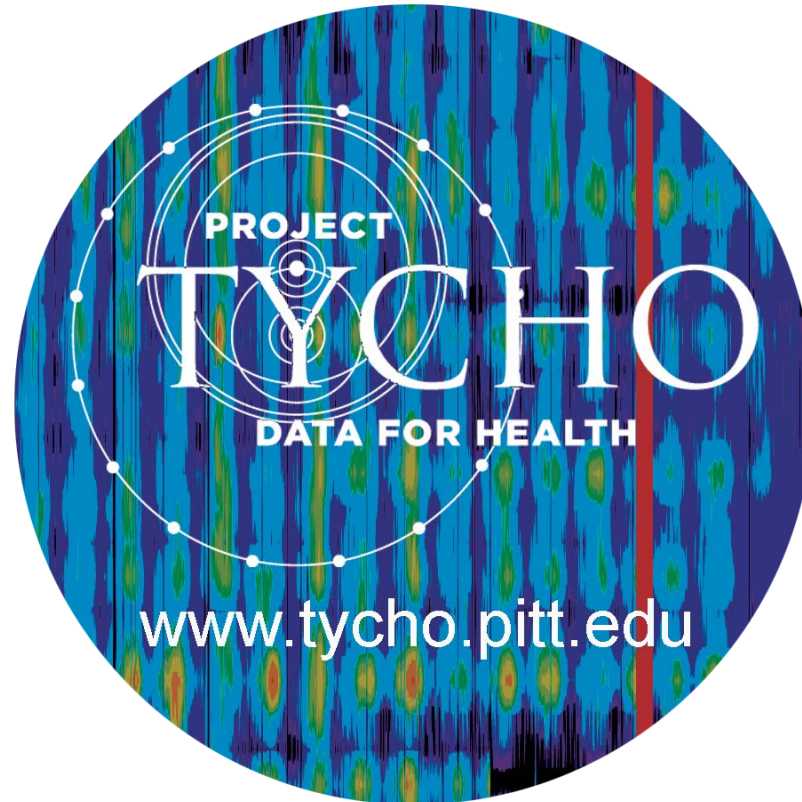


National Institute of  
General Medical Sciences



# Take some Project Tycho stickers!

---



Help improve the discovery of open  
disease data!