

Preliminary data, not for publication or official use

# Project Tycho

*Preliminary data for the state of Connecticut*

## Tycho database beta test version

*The data presented in this report are of preliminary nature and should not be used for publication or other types of official use*

June 2011

University of Pittsburgh Graduate School of Public Health



University of Pittsburgh



## Table of Content

<b>Acknowledgements</b>	<b>page 3</b>
<b>Contact information</b>	<b>page 4</b>
<b>Introduction</b>	<b>page 5</b>
<b>Vision and goal</b>	<b>page 7</b>
<b>Collaboration</b>	<b>page 8</b>
<b>Summary of methods</b>	<b>page 9</b>
<b>Overview of data available for the state of Connecticut</b>	<b>page 12</b>
<b>Disease specific data for Connecticut</b>	<b>page 19</b>
- AIDS	page 19
- Chlamydia	page 20
- Cryptosporidiosis	page 21
- Diphtheria	page 22
- Escherichia Coli	page 24
- Giardiasis	page 25
- Gonorrhoea	page 26
- Haemophilus Influenzae	page 27
- Hepatitis	page 28
- Influenza	page 30
- Legionellosis	page 32
- Lyme disease	page 33
- Malaria	page 34
- Measles	page 36
- Meningitis	page 39
- Meningococcal disease	page 40
- Mumps	page 42
- Poliomyelitis	page 44
- Rabies in animals	page 45
- Rubella	page 46
- Salmonellosis	page 47
- Scarlet fever	page 48
- Shigellosis	page 51
- Streptococcal sore throat	page 52
- Syphilis	page 53
- Toxic Shock Syndrome	page 54
- Tuberculosis	page 55
- Typhoid fever	page 58
- Whooping cough	page 61
<b>Diseases with fragmented data</b>	<b>page 63</b>

## Acknowledgements

The development of this project would not have been possible without data collected every week by staff from city-, county-, state- and federal public health departments. We would like to acknowledge their invaluable contribution to the overall body of surveillance data that were digitized during this project. Particularly, we acknowledge staff from the Council of State and Territorial Epidemiologists (CSTE) and the US Centers for Disease Control for their overall efforts in disease surveillance and their particular input in this project.

We also acknowledge the help of many contributors that have contributed to the digitization and processing of weekly surveillance data included in this database. Digital Divide Data: Linda Thomas, Eric Gold, Heak Hok, Villa Kem and Hetel Patel. University of Pittsburgh Graduate School of Public Health: Steve Wisniewsky, Heather Eng, Sharon Crow, David Galloway, Jia Xu and Raaka Kumbhakar. University of Pittsburgh School of Information Sciences: Vladimir Zadorozhny, Ying-Feng Hsu, David Paul Koenig , Fatimah Ridwan and Divyasheel Sharma. Johns Hopkins Bloomberg School of Public Health: Derek Cummings. Pittsburgh Supercomputing Center: David Graham and Andrew Adams.

This project was made possible with funding from the Bill & Melinda Gates Foundation through the Vaccine Modeling Initiative (grant ID 49276).

The Tycho development team at the University of Pittsburgh Graduate School of Public Health:

Willem G. van Panhuis

John Grefenstette

Shawn T. Brown

Su Yon Jung

Nian Shong Chok

Anne Cross

Donald S. Burke

## Contact information

The Tycho database is currently being tested and a login account will be required to use the database and website ([www.tycho.pitt.edu](http://www.tycho.pitt.edu)). All State Epidemiologists, CDC partners, other federal agencies and our research collaborators will be provided with user accounts. User accounts can also be provided upon specific request. Note that the quality of the data during the testing phase cannot be guaranteed to be sufficient for publication or official use.

**Point of contact for the Tycho project:**

Dr. Willem G. van Panhuis, MD PhD  
Graduate School of Public Health  
University of Pittsburgh  
130 DeSoto Street  
704 Parran Hall  
Pittsburgh PA, 15261  
Email: [wav10@pitt.edu](mailto:wav10@pitt.edu)  
Tel: 412-624-7693

## Introduction

**Project Tycho** started in 2009 as part of the Vaccine Modeling Initiative (VMI) at the University of Pittsburgh Graduate School of Public Health, funded by the Bill & Melinda Gates foundation. This project aims to contribute to the availability of public health data for research and policy making. The vision for this project as described in the next section is a central access point for public health data of a detailed spatial and temporal scale for all countries. Currently, the Tycho database contains 10 million records that each represent a weekly report from a location for a specific disease. These records were extracted from weekly US Nationally Notifiable Disease Surveillance Reports between 1888 and 2009 (6300 reports) using double data entry (200 million keystrokes). The current database includes a total of 90 million reported cases and 4 million reported deaths due to notifiable diseases in the United States for the last 122 years.

Project **Tycho** is named after the Danish nobleman Tycho Brahe (1546 – 1601), who meticulously collected astronomical data. After Brahe's death, his assistant Kepler used these data to develop area laws of planetary motion.

## Background

***"No health department, State or local, can effectively prevent or control diseases without knowledge of when, where and under what conditions cases are occurring"***

(Quote stated above weekly surveillance reports published in the Public Health Reports, 1888-1951)

This quote has accompanied weekly reports on the prevalence of notifiable diseases in the United States as long as these were published in the Public Health Reports, emphasizing the importance of data for disease control policy making. In the current day and age, this statement has lost none of its significance. On the contrary, as infectious diseases continue to pose challenges to the global public health system in an increasingly interconnected global society. The public health system has responded to this challenge by evermore sophisticated disease surveillance systems. The availability of internet and information technology has facilitated collection of detailed data on infectious disease incidence, the spread of pathogens, disease determinants, health behavior, etc. New technologies such as remote sensing and mobile reporting systems have advanced measurement of these factors in both high and low resource settings. The enormous wealth of public health information that is currently being accrued also poses its own challenges for data storage, management, preservation and dissemination. Public health agencies will need to allocate resources and develop capacity to specifically address these issues.

Although the importance of public health data collection has been widely recognized, the dissemination of data has been neglected for many reasons. The above quote does not only imply data collection but also its use and dissemination as knowledge on disease occurrence can only be derived from data after processing and analysis. Not only has disease surveillance become more sophisticated, analytical tools have changed as well and now include computational models to assess disease transmission and to evaluate disease control options. These methods have introduced the opportunity for data intensive studies of fine spatial-temporal patterns using dynamic models. Disease surveillance data at such scale are often not available for analysis for multiple reasons. First, surveillance data are not collected for research purposes but for disease monitoring and planning. Secondly, public health staff often lack time or capacity to prepare these data for analysis. Third, surveillance methodology often lacks transparency and standardization, reducing options for scientific analyses. Fourth, disease surveillance data are often

# Preliminary data, not for publication or official use

kept confidential and ownership or data use requirements are often unclear. Finally, archiving and preservation of surveillance data are often not a priority and disaggregated records may get lost, losing the opportunity to ever use these data for analysis.

The Tycho project aims at overcoming some of these challenges by providing open access to large public health datasets of great spatial-temporal detail. Open access to these data will provide great opportunities for scientific analysis and better informed policy making on disease control.

## **The Tycho database and website**

As described in the methods section, a large heterogeneity in reporting practice was observed over the past 122 years of weekly disease surveillance and substantial efforts were made to clean and standardize these data. The current online database includes all cleaned and standardized data in a format that allows comparisons across time and space. It does not include parts of the data that require additional cleaning and research such as classification of disease subcategories.

Each record in the database provides a reported number of cases or deaths due to a specific disease for a specific location, for a certain week. The database can be searched online ([www.tycho.pitt.edu](http://www.tycho.pitt.edu)) by selecting a dataset, an aggregation method and an output format. Data can be viewed in tables, graphics and maps and these can be downloaded for further use.

## **Data availability**

The data provided by the Tycho database reflect the weekly US Nationally Notifiable Disease Surveillance System between 1888 and 2009. It includes all diseases that were reported as part of this system, but it does not yet include records from other surveillance systems such as the annual system or disease specific subsystem (such as the influenza surveillance system).

The available data also reflects changes in reporting practices over time. This explains why city or state level reports are only available for a certain time period (1888-1953 for cities and 1927-2009 for states) and why morbidity and mortality reports are available for different time periods. We are currently working on inclusion of additional data as described in the section on collaboration.

## **Testing and release of the Tycho database**

The Tycho database is currently in a beta testing phase that will start during the 2011 annual meeting of the Council of State and Territorial Epidemiologists (CSTE) in Pittsburgh June 12-16. This phase will end with a release to the general public in the fall of 2011. During the testing phase, invited users will be provided with login accounts that will enable full use of all features of the website and database. Invited users will include all state epidemiologists, partners in the CDC and other federal agencies as well as research collaborators. During this phase, data cleaning and standardization will continue. The quality of the data will not be sufficient for publication or official use during this phase.

## **Future developments**

After optimization of the database and website during the testing phase, the first version of the Tycho database will be released to the general public. After that, efforts to clean and standardize the data will continue and more data will be released in subsequent versions. In addition, ongoing and new collaborations for inclusion of new data will allow gradual expansion of the dataset over time and increased opportunities for data driven analysis and policy making.

## Vision and goal

The vision for project Tycho is based on the value of open access to detailed, disaggregated public health data for scientific analysis and policy making. Starting with 122 years of weekly US surveillance data, it is our hope that this example will convince public health authorities worldwide of the value of this resource and will stimulate similar contributions from other countries.

In the past, data sharing and archiving have not been a priority in public health or among health sciences in general. Continuous and diligent public health data collection has led to a large pool of disaggregated data scattered over the world including the internet, ranging from paper archives to basements of local health departments. The need for open access to a central repository of public health data has been recognized and both benefits and challenges of open access to data in general have been well described. Open access to data could lead to increased accountability, transparency, innovation, collaboration, cost-effectiveness, replication of results, the development of new methods and insights, and a narrower data access gap between low- and higher income countries. Ultimately, this would result in the advancement of science and technology to the benefit of all.

Challenges to open access include inconsistent formats, lack of annotation or metadata, lack of information technology (IT) capacity, a lack of incentive, and cultural norms around data ownership. Open access to public health data in particular is confronted by decentralized archiving of disease reports, lack of time and capacity for data management, degraded physical format of records and archives, and governmental level legal constraints. Multiple principles and requirements for a central open access repository for public health data have been described and include (1) capacity building for data management in low- and middle income countries, (2) feedback loops to data contributors, (3) sustainability, (4) common standards, (5) interoperability, and (6) user-friendliness.

The ultimate goal of the Tycho project is to provide a central global public health data access point. Historical as well as current public health data are of great value if archived and accessible for research and analysis. Open access will enable the use of analytical capacity from around the globe which will lead to new discoveries of disease patterns and control policies. Furthermore, a central data access point will facilitate data archiving and preservation into the future which will be an increasing need in a data rich public health environment.

## Collaboration

The application of the surveillance data from the Tycho database as well as the inclusion of new data will require a wide range of collaborations with public health offices in the US and abroad. The Tycho project has currently made the first step of digitizing 122 years of weekly US surveillance reports that had already been published. The next step will be the completion of the existing dataset as well as the inclusion of new data from unpublished sources.

This report provides an overview of data available from the Tycho database for the state of Connecticut. This includes all data that has ever been published at state or city level for Connecticut in the weekly US Nationally Notifiable Disease Surveillance System. As you will find, the data availability will vary greatly over time, between city and state level reports, between diseases and between morbidity vs. mortality reports.

The Tycho development team would look forward to work with state health departments in the US to:

1. Better understand the current Tycho data for each state (eg. by collecting historic documentation)
2. Provide better usability and applications of current Tycho data for each state by continued joint development of website and database features.
3. Collect and/or digitize official, confirmed data to validate the current (preliminary) weekly data.
4. For each state, collect and/or digitize new data that has not yet been included. For example all city level reports discontinued in 1953 and completion of city level data until 2009 would greatly increase opportunities for analysis of disease patterns and trends.
5. Provide support to state health offices to manage, preserve and provide access to public health data.

## Summary of methods

Detailed documentation on the methods used for the Tycho database has been provided on the website ([www.tycho.pitt.edu](http://www.tycho.pitt.edu)). This section describes these methods in short.

### Data collection

Weekly reports that contain tables on the occurrence of nationally notifiable diseases have been published since 1888 by public health authorities at the Federal level in various journals. Table 1 provides the list of publications and the responsible agency since 1888. All weekly nationally notifiable disease reports between 1888 and 1951 could be retrieved from the PubMed Central repository of the National Library of Medicine <sup>1</sup>. For 1995 to 2009, these reports could be retrieved from the MMWR digital archive on the CDC website <sup>2</sup>. Most weekly reports between 1952 and 1995 could be retrieved from the HathiTrust Digital Library <sup>3 4</sup>, but many could not be found and had to be copied from hard copies of MMWR issues in the University of Pittsburgh library.

Table 1, publications and responsible Federal agencies for nationally notifiable disease reports

Time period	Publication title	Responsible federal agency
1888-1889	Weekly Abstract of Sanitary Reports	US Marine Hospital Service
1890-1895	Abstract of Sanitary Reports	US Marine Hospital Service
1896-1901	Public Health Reports	US Marine Hospital Service
1902-1911	Public Health Reports	US Public Health and Marine Hospital Service
1912-1951	Public Health Reports	US Public Health Service
1952-1960	Morbidity and Mortality Weekly Report	National Office of Vital Statistics, US Public Health Service
1961-1969	Morbidity and Mortality Weekly Report	Communicable Disease Center
1970-1991	Morbidity and Mortality Weekly Report	Center for Disease Control
1992-2009	Morbidity and Mortality Weekly Report	Centers for Disease Control and Prevention

### Inclusion criteria

Weekly reports of each year were reviewed systematically to assess the diseases reported. We included all tables that provided disease specific information by week for US cities, townships, counties or states. Tables that provided summary or aggregated information by month, year or at the national level were not included. Similarly, tables that did not contain disease specific information (such as all cause mortality) were not included.

### Data entry

Weekly reports were downloaded or scanned as PDF files and selected tables with notifiable disease reports were entered into computer spreadsheets in a highly standardized fashion using double data entry. During the second round of data entry, operators could not see what had been entered in the first round and could not continue if the system detected a discrepancy between the second and first entry

<sup>1</sup> <http://www.pubmedcentral.nih.gov/tocrender.fcgi?journal=333&action=archive>

<sup>2</sup> <http://www.cdc.gov/mmwr>

<sup>3</sup> <http://catalog.hathitrust.org/Record/003910026>

<sup>4</sup> <http://catalog.hathitrust.org/Record/003843660>

# Preliminary data, not for publication or official use

for a specific value. Such discrepancies could only be resolved by checking the PDF file and try again or discussion with the group leader and verification of a value in the source documents.

## **Quality control for data entry**

The accuracy of data entry was checked at various levels. First, completeness of data was verified by comparing the content of entered data with PDF sources files. Secondly, accuracy of data entry was verified by multiple rounds of comparing random samples of entered files with PDF source files. Thirdly, data formatting was verified by various checks to ensure appropriate formatting for data loading.

## **Data loading and standardization**

All data was entered in Excel spreadsheets and various components of these spreadsheets were loaded in data files. Table titles, column headers, place names and reported numbers were loaded in separate files. These files were used to extract information on each reported number, including:

1. the disease reported
2. the disease subcategory reported
3. cases or deaths reported
4. the reporting location (name, state and type of location)
5. the time period for which a number was reported
6. the date of publication of the original weekly report associated with a reported number

## **Integration**

All reported numbers and extracted information was integrated in one database with one record per reported number and associated information.

## **Post-processing quality control**

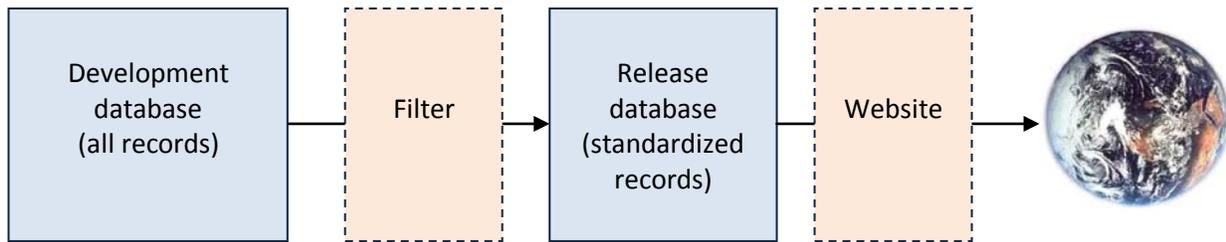
After integration of all data in one database, checks were performed to detect duplicate reports and data inconsistencies. Duplicate records were removed and inconsistencies resolved by verification with original PDF source files.

## **Data filtering**

The digitized version of all historical weekly US nationally notifiable disease surveillance records is a very heterogeneous dataset. It took a substantial number of data processing protocols to standardize time and space variables as well as disease names. All standardized records have been separated from non-standard records by a filter. All remaining heterogeneity in non-standard records is inherent to the surveillance system and can only be standardized after further analysis will have been completed (eg. remaining heterogeneity in reporting periods for reports before 1953 and standardization of disease subcategories). The largest proportion of data (>4 million records) has been standardized however and has been made available in the current testing version.

## **Data visualization**

Figures in this reports were made with the R system, version 2.9.2 and the maps were generated by the GAIA platform developed at the University of Pittsburgh Graduate School of Public Health in collaboration with the Pittsburgh Supercomputing Center (PSC). See <http://midas-pitt.psc.edu/gaia> for more information.



**Figure 1, Schematic of the Tycho data architecture.** Access to a consistent, standardized subset of the digitized US weekly surveillance reports will be provided for beta testing through a website that will allow querying and downloading of data. Data from the development database will be pushed through a filter to feed the Tycho database.

The current Tycho database will provide the reported number of cases or deaths reported by city or state health authorities to the federal health agency for all weeks between 1888 and 2009. Currently, only standardized, consistent data are being released for testing, according to the following criteria:

1. Only weekly reports are provided (reports for other periods such as 10 days, 2 weeks, 1 month, 1 year) are currently not included due to lack of comparability with the weekly reports.
2. Only reports that were published in the same year as the reporting period were included (this excludes updates or comparisons that were published more than a year after the original reporting period)
3. Only numbers were included for which information about the location, time period and disease could be extracted.
4. Only numbers or time periods for which no disease subcategories were reported were included.

Each of the steps described above have been described in detailed protocols that are posted in the documentation part of the Tycho website ([www.tycho.pitt.edu](http://www.tycho.pitt.edu)).

### **Website**

The Tycho website features simple and advanced searches of the data, visualizations and downloading of machine readable files. A dataset can be selected by specifying: 1) the disease of interest, 2) the location(s) of interest and 3) the time period. In a second step, the type of aggregation and length of increments can be specified and in a third step, the type of output (table, graphic or map) can be selected. After output has been generated, data can be downloaded.

As mentioned above, data availability varies largely by disease, time period and location. We aimed to restrict selection options by data availability to avoid "no results available."

Access to the most of the website will be password protected during the testing phase. For invited users after login, full functionality of the database and website will be available. During this phase however, the data quality will not be sufficient for publication or official use.

## Overview of data available for the state of Connecticut

In this section, an overview of data available for the state of Connecticut will be provided. As described earlier, data availability depends entirely on historical reporting practices in the weekly US Nationally Notifiable Disease Surveillance System between 1888 and 2009. For this report, we only included morbidity reports (cases) to improve the format and limit the size of this report. As mentioned earlier, the analysis and standardization of has not been completed yet for all records in the database and preliminary data are provided here that may not yet be available in the online database.

This section will provide a general overview of data availability. The next sections will provide disease specific data from state and city level reports of Connecticut. Data for a maximum of 3 major cities were provided depending on availability (city data was only reported until 1953). Some summary data will be provided at the end of this report, for a subset of diseases for which data was fragmented over time.

Table 1 lists the number of weekly state or city reports that are available for Connecticut per each disease and subcategory.

**Table 1, Number of weekly state or city reports per disease and subcategory**

<b>Disease</b>	<b>City</b>	<b>State</b>
Aids	-	948
Anthrax	-	143
Brucellosis [undulant fever]	1	293
Chickenpox [varicella]	452	514
Chlamydia	-	646
Cryptosporidiosis	-	399
Dengue	8	-
Diphtheria	2130	1163
Dysentery		
<i>Amebic</i>	-	235
<i>Bacillary</i>	-	220
<i>Unspecified</i>	-	196
Encephalitis		
<i>Lethargic</i>	206	-
<i>Post infectious</i>	1	178
<i>Primary [infectious] including unspecified</i>	349	891
Escherichia coli		
<i>EHEC 0157</i>	-	170
<i>EHEC non-0157</i>	-	151
<i>O157:H7 NETSS</i>	-	243
<i>O157:H7 PHLIS</i>	-	184
<i>STEC</i>	-	158
Giardiasis	-	358

# Preliminary data, not for publication or official use

**Table 1, Number of weekly state or city reports per disease and subcategory, *continued***

<b>Disease</b>	<b>City</b>	<b>State</b>
Gonorrhea		
<i>Civilian</i>	-	758
<i>Unspecified</i>	-	1045
Haemophilus influenzae		
<i>Age &lt;5 non-serotype B</i>	-	129
<i>Age &lt;5 unknown serotype</i>	-	17
<i>All ages all serotypes</i>	-	514
Hepatitis		
<i>Acute type A</i>	-	375
<i>Acute type B</i>	-	298
<i>Acute type C</i>	-	31
<i>All types, &lt;20 years</i>	-	193
<i>All types, &gt;=20 years</i>	-	99
<i>All types, all ages</i>	35	690
<i>Type A [infectious]</i>	1	1621
<i>Type B [serum]</i>	1	1363
<i>Type NA NB [including C]</i>	-	260
<i>Type unspecified</i>	1	316
Influenza	879	925
Legionellosis	-	614
Leprosy	-	305
Listeriosis	-	227
Lyme disease	-	679
Malaria		
<i>Civilian</i>	-	10
<i>Unspecified</i>	2	1626
Measles		
<i>Imported</i>	-	413
<i>Indigenous</i>	-	498
<i>Unspecified</i>	2080	2875
Meningitis		
<i>Aseptic</i>	1	569
<i>Meningococcus</i>	583	1212
<i>Unspecified</i>	319	5
Meningococcal disease		
<i>All serogroups</i>	-	45
<i>Invasive all serogroups</i>	-	181
<i>Invasive serogroup unknown</i>	-	45
<i>Serogroup A C Y and W-135</i>	-	39
<i>Serogroup B</i>	-	2
<i>Serogroup unspecified</i>	25	2286

# Preliminary data, not for publication or official use

**Table 1, Number of weekly state or city reports per disease and subcategory, *continued***

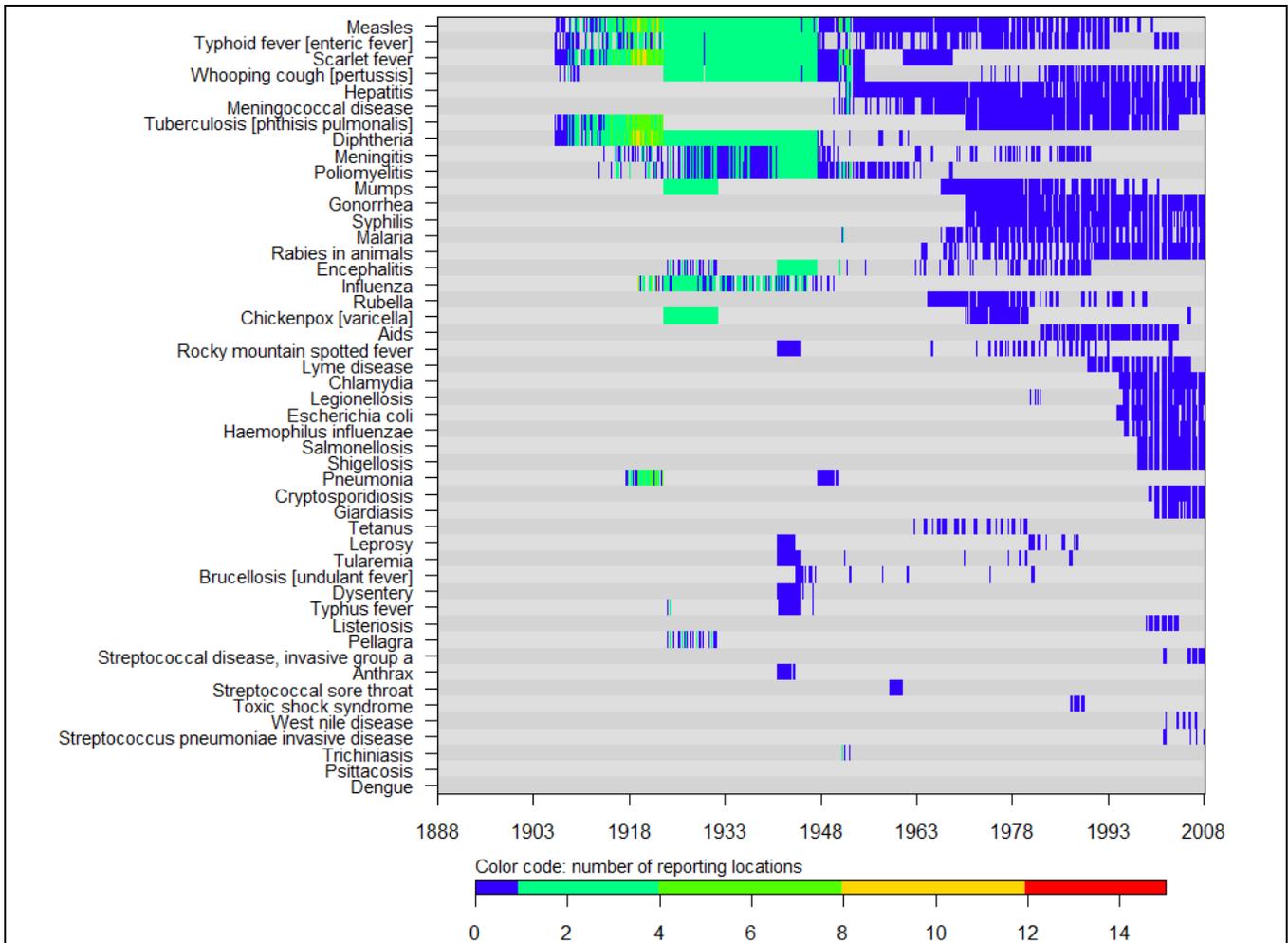
<b>Disease</b>	<b>City</b>	<b>State</b>
Mumps	451	1393
Pellagra	206	-
Pneumonia		
<i>Lobar</i>	26	-
<i>Unspecified</i>	255	171
Poliomyelitis		
<i>Non paralytic</i>	-	119
<i>Paralytic</i>	-	294
<i>Total</i>	847	1627
Psittacosis	-	24
Rabies in animals	2	1615
Rocky mountain spotted fever	-	722
Rubella	-	1167
Salmonellosis		
<i>NETSS</i>	-	101
<i>PHLIS</i>	-	125
<i>Unspecified</i>	-	380
Scarlet fever		
<i>Including streptococcal sore throat</i>	87	565
<i>Unspecified</i>	2122	1281
Shigellosis		
<i>NETSS</i>	-	97
<i>PHLIS</i>	-	126
<i>Unspecified</i>	-	374
Streptococcal disease, invasive group a	-	145
Streptococcal sore throat	-	101
Streptococcus pneumoniae invasive disease		
<i>Drug resistant &lt;5 years</i>	-	18
<i>Drug resistant A</i>	-	5
<i>Drug resistant all ages</i>	-	45
<i>Drug resistant B</i>	-	3
<i>Non drug resistant &lt;5 years</i>	-	18
Syphilis		
<i>Civilian primary and secondary</i>	-	737
<i>Congenital</i>	-	56
<i>Primary and secondary</i>	-	1051
Tetanus	-	332
Toxic shock syndrome	-	100
Trichiniasis	7	28

**Table 1, Number of weekly state or city reports per disease and subcategory, *continued***

<b>Disease</b>	<b>City</b>	<b>State</b>
Tuberculosis [phthisis pulmonalis]		
<i>New active</i>	-	144
<i>Unspecified</i>	846	1463
Tularemia	-	304
Typhoid fever [enteric fever]		
<i>Including paratyphoid fever</i>	339	584
<i>Unspecified</i>	1553	2143
Typhus fever		
<i>Endemic</i>	-	24
<i>Unspecified</i>	28	177
West nile disease		
<i>Neuroinvasive</i>	-	69
<i>Non-neuroinvasive</i>	-	69
Whooping cough [pertussis]	1368	2007

The diseases included in the weekly US Nationally Notifiable Disease Surveillance System varied largely over time and reflected the historical social-political priorities of each time period. Note that the diseases in the weekly system were a subset of all diseases included in the annual Notifiable Disease Surveillance System (that was not entered as part of this project).

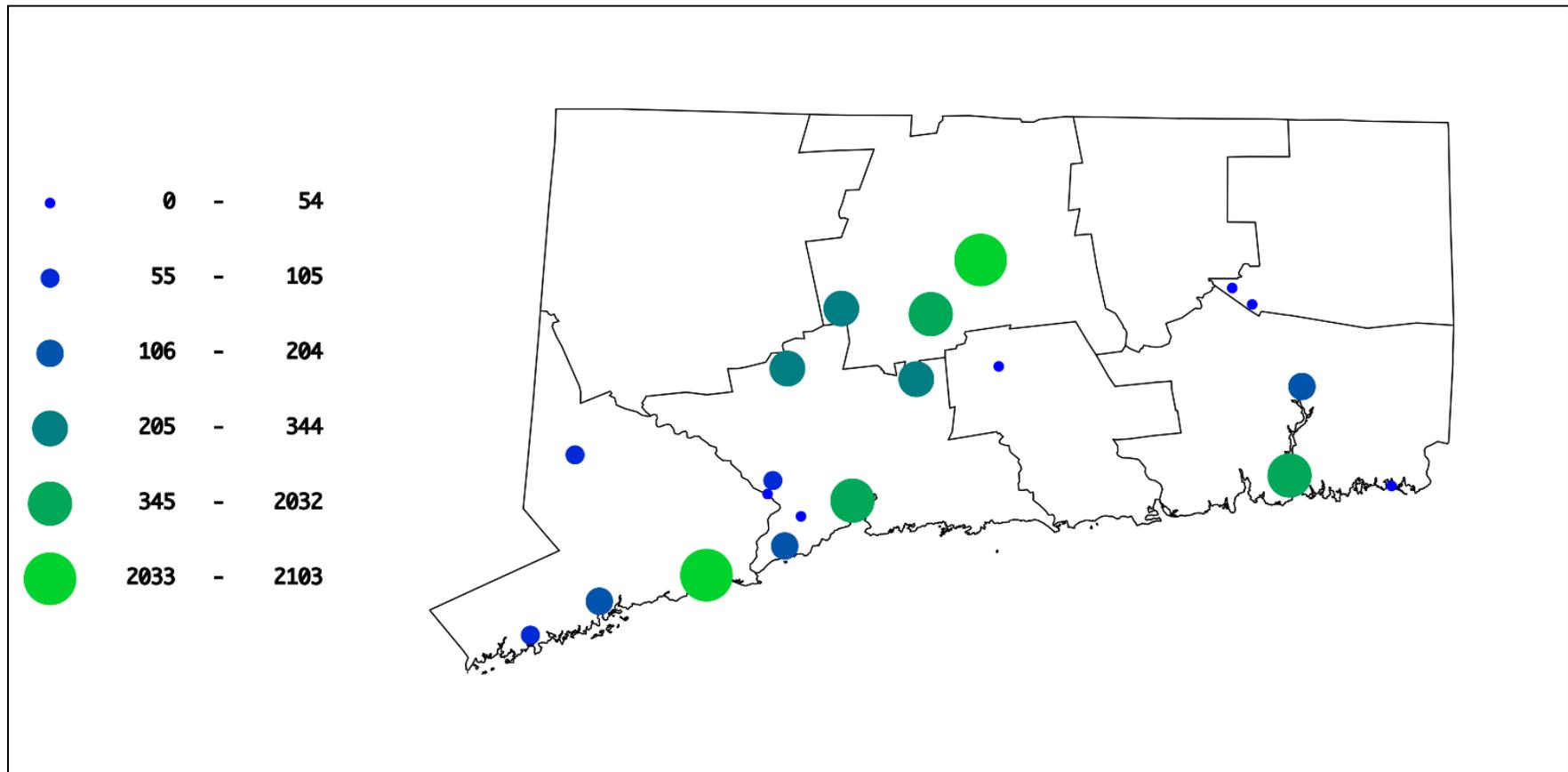
The number of locations reporting diseases also varied over time depending on the type of reports submitted. Figure 1 on the next page provides an overview of the number of unique locations reporting each disease per week between 1888 and 2009. For the remainder of this report, the state of Connecticut was considered a unique reporting location in addition to the individual city reports. Figure 1 indicates that many locations (mostly cities) were included in the reporting system before 1925, after which the number of cities dropped substantially. After 1953, only state reports were included and city reports were discontinued. These patterns were observed nationwide.



**Figure 1,** The number of locations (in color code) reporting on each disease per week between 1888 and 2009

For the majority of the time period between 1888 and 2009, only state level reports were available. It would be of great interest to complete the collection of city data after 1953 and we will seek collaboration with state health offices for this.

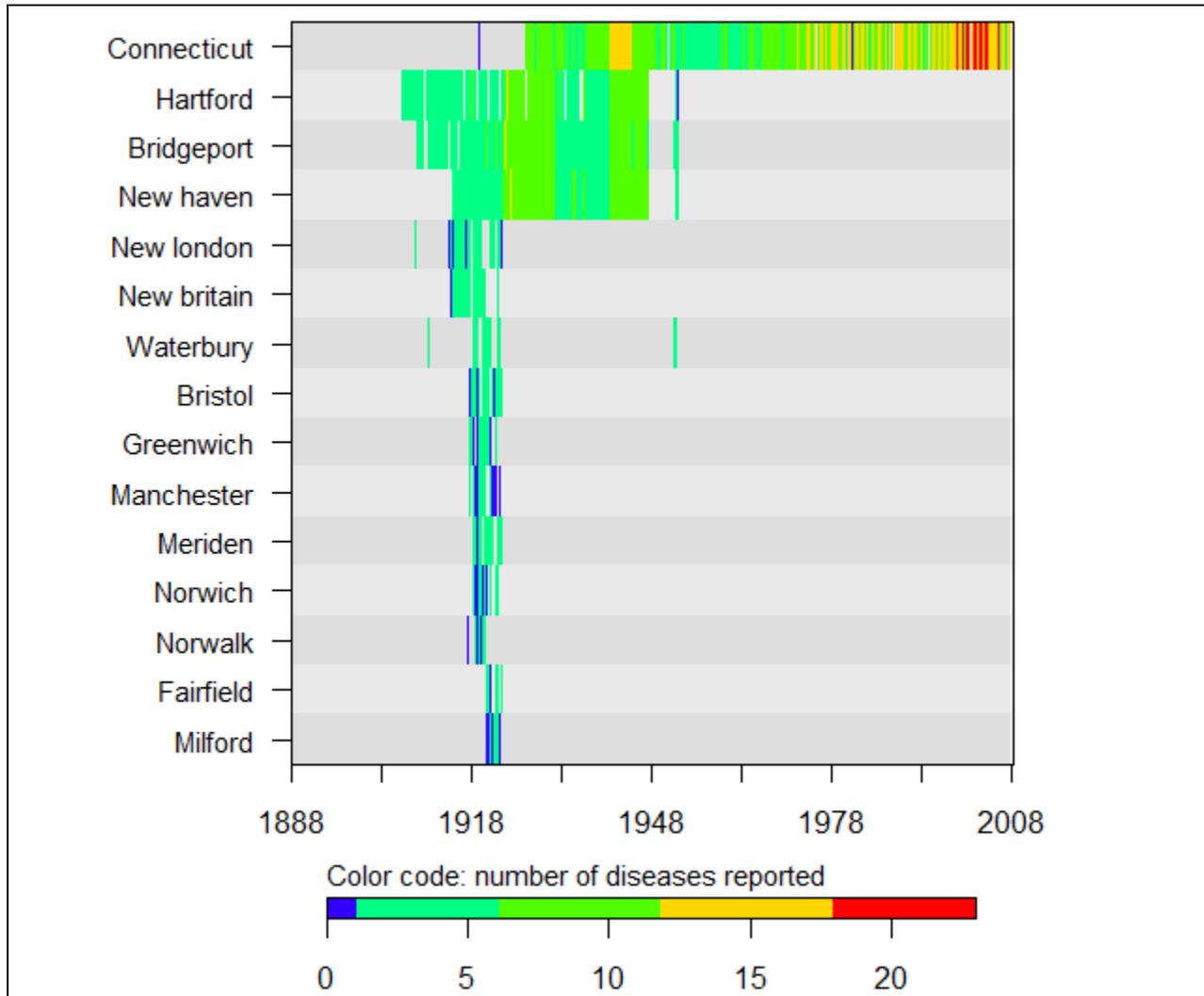
The map in figure 2 (next page) illustrates the geographical distribution of the available city data. It displays the number of weekly reports for any disease available per location that was included (and for which coordinates could be derived)



**Figure 2**, This map displays the number of weekly reports available for any disease per city (note that city reports were only available until 1953)

# Preliminary data, not for publication or official use

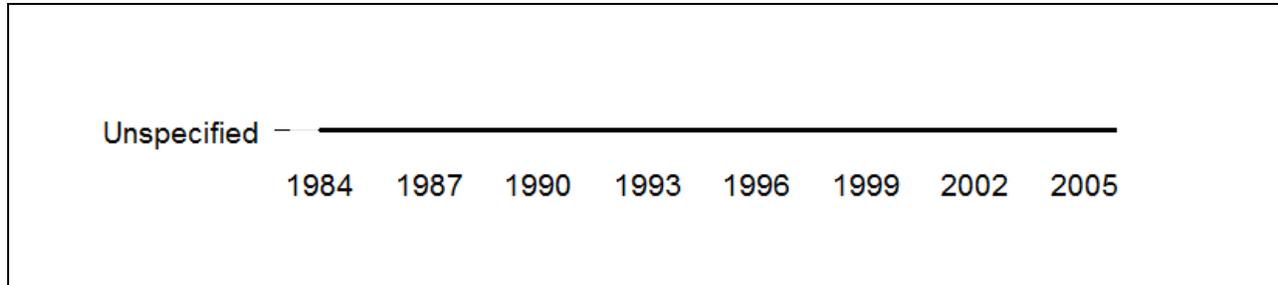
Figure 3 below provides an overview of the number of diseases for which a weekly report was included at the state level and for individual cities. All cities for which at least 100 weekly reports are available have been listed. For most states, it shows a major drop in the number of cities after 1925 and after 1953, city reports were discontinued except for New York City, which became its own reporting jurisdiction.



**Figure 3,** The number of diseases (in color code) that were reported for the state and city level (state on top row) for each week between 1888 and 2009

## Disease specific data for Connecticut

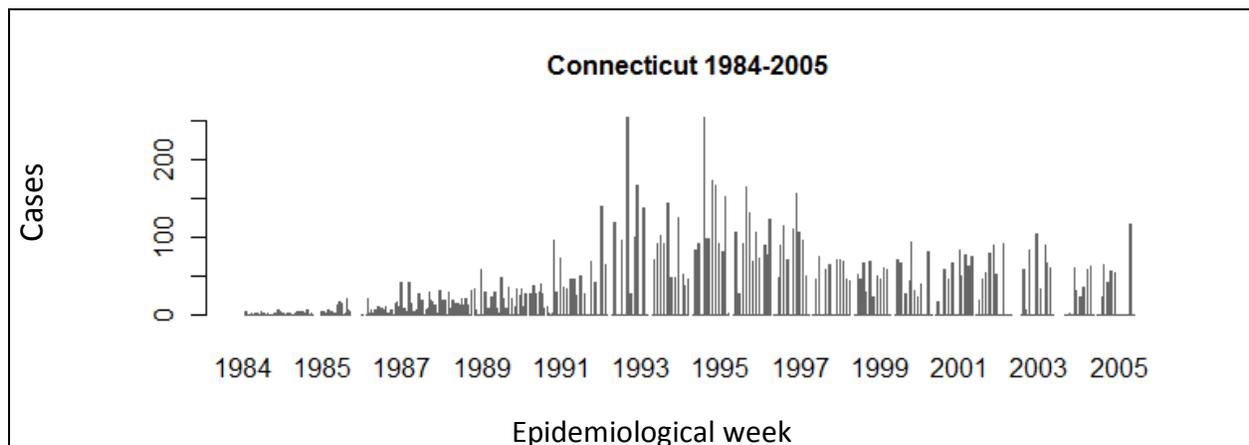
### AIDS



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

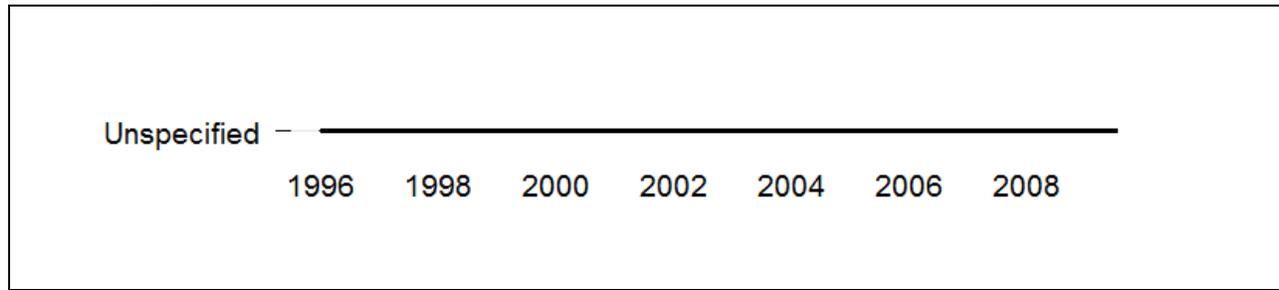
**Table D1**, Summary information for AIDS

Indicator	Connecticut
Report period	1984-2005
Total weeks	948
Total cases	14,398
Max. cases per year	2,498
Year (max)	2004
Max. cases per week	2,224
Week (max)	2004, wk 22
Average cases per year	654
95%CI	(405-903)
Average cases per week	15
95%CI	(10-20)



**Figure D2**, Number of cases reported for AIDS per epidemiological week

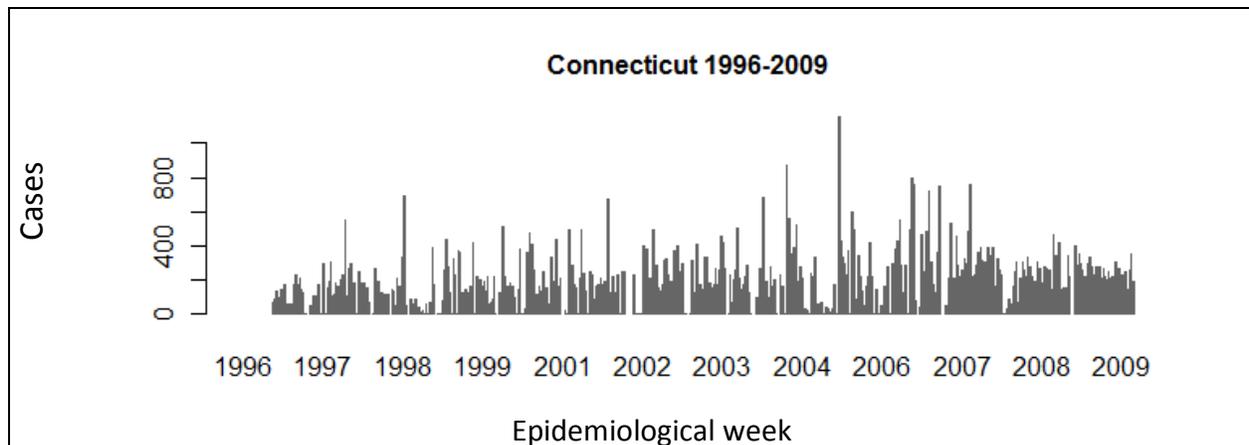
## Chlamydia



**Figure D1,** Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

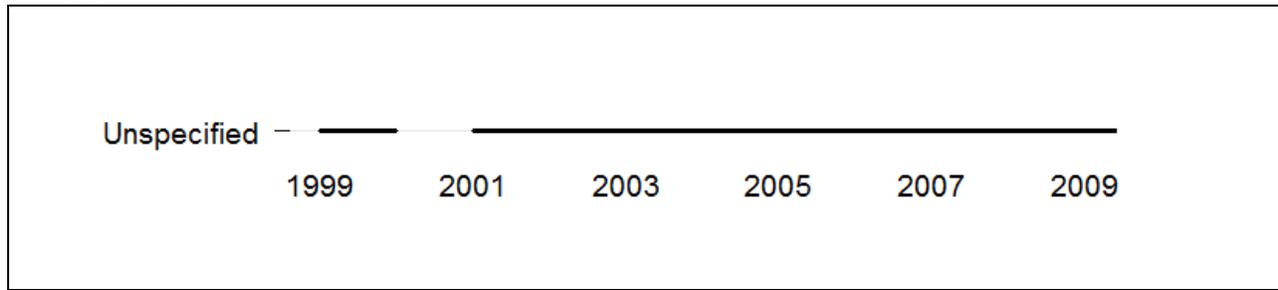
**Table D1,** Summary information for Chlamydia

Indicator	Connecticut
Report period	1996-2009
Total weeks	645
Total cases	103,717
Max. cases per year	10,712
Year (max)	2009
Max. cases per week	1,289
Week (max)	1999, wk 48
Average cases per year	7,408
95%CI	(6,156-8,660)
Average cases per week	161
95%CI	(148-174)



**Figure D2,** Number of cases reported for Chlamydia per epidemiological week

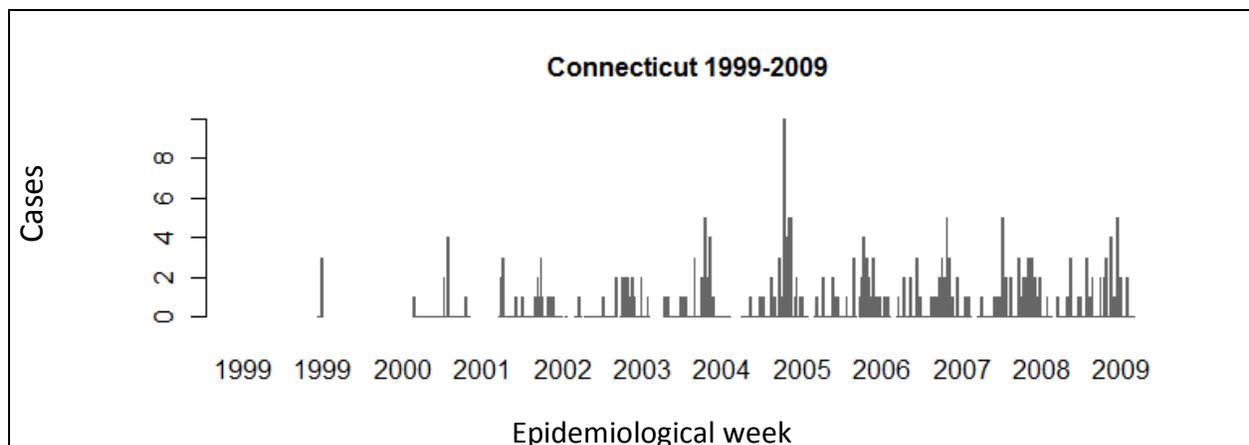
## Cryptosporidiosis



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

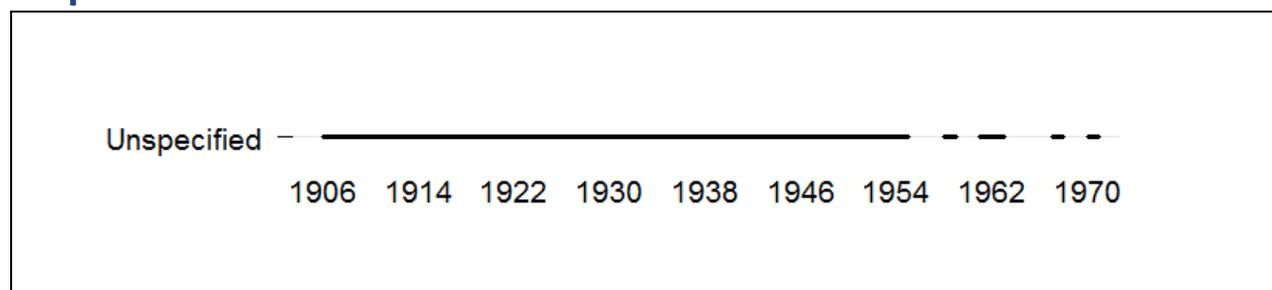
**Table D1**, Summary information for Cryptosporidiosis

Indicator	Connecticut
Report period	1999-2009
Total weeks	399
Total cases	429
Max. cases per year	163
Year (max)	2004
Max. cases per week	140
Week (max)	2004, wk 22
Average cases per year	43
95%CI	(9-77)
Average cases per week	1
95%CI	(0-2)



**Figure D2**, Number of cases reported for Cryptosporidiosis per epidemiological week

## Diphtheria



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

**Table D1**, Summary information for Diphtheria

Indicator	Connecticut	Hartford	Bridgeport	New Haven
Report period	1927-1970	1906-1952	1909-1948	1915-1948
Total weeks	1,163	1,990	1,924	1,666
Total cases	6,337	6,907	5,612	2,713
Max. cases per year	1,215	429	488	478
Year (max)	1928	1922	1922	1920
Max. cases per week	623	37	41	32
Week (max)	1970, wk 44	1920, wk 44	1924, wk 01	1919, wk 51
Average cases per year				
before 1940	363	196	175	104
95%CI	(152-574)	(147-245)	(118-232)	(45-163)
after 1940	66	7	1	1
95%CI	(-4-136)	(3-11)	(0-2)	(0-2)
Average cases per week				
before 1940	8	4	4	2
95%CI	(7-9)	(4-4)	(4-4)	(2-2)
after 1940	2	0	0	0
95%CI	(-1-5)	(0-0)	(0-0)	(0-0)

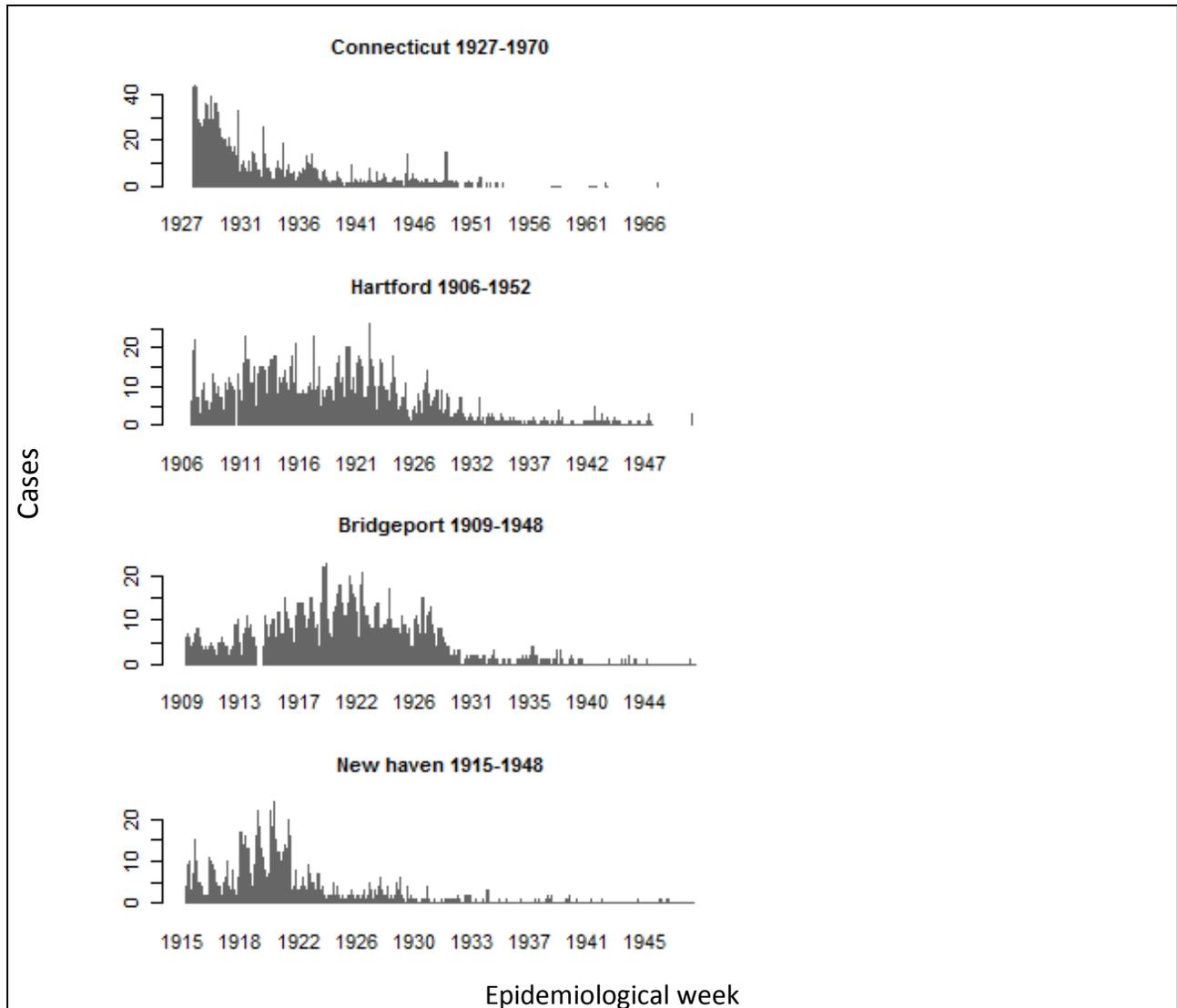


Figure D2, Number of cases reported for Diphtheria per epidemiological week

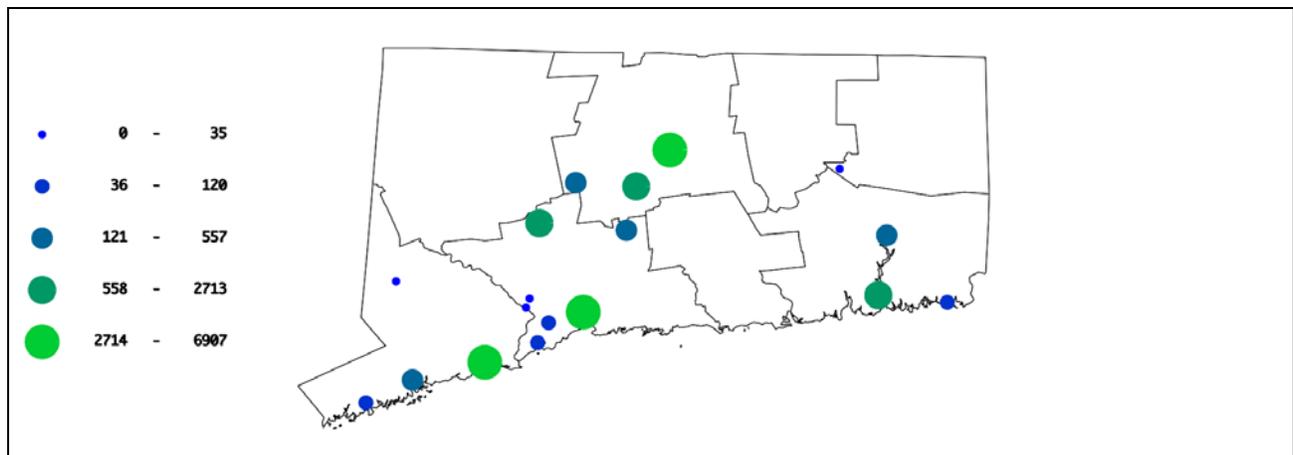
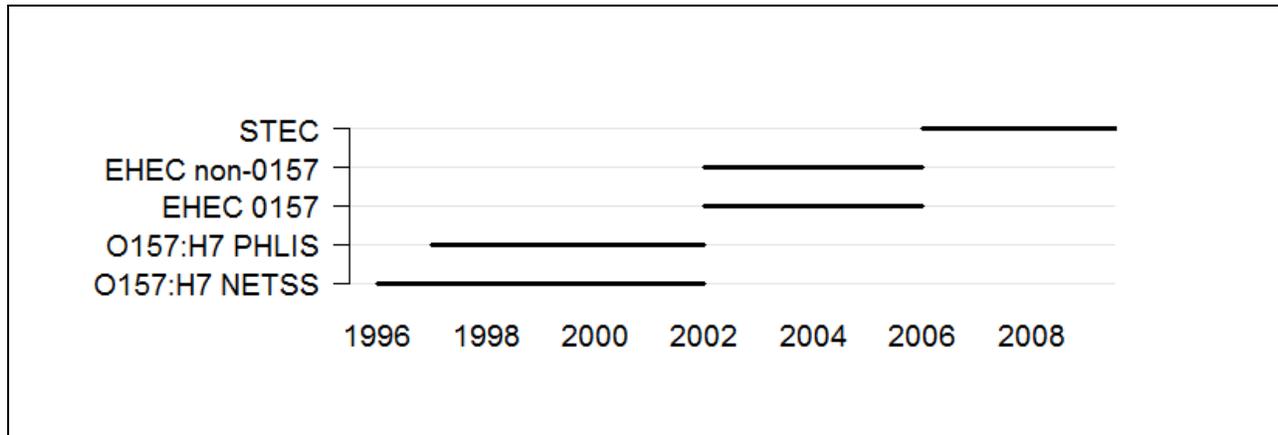


Figure D3, The total number of cases reported for Diphtheria by each city for the entire time period

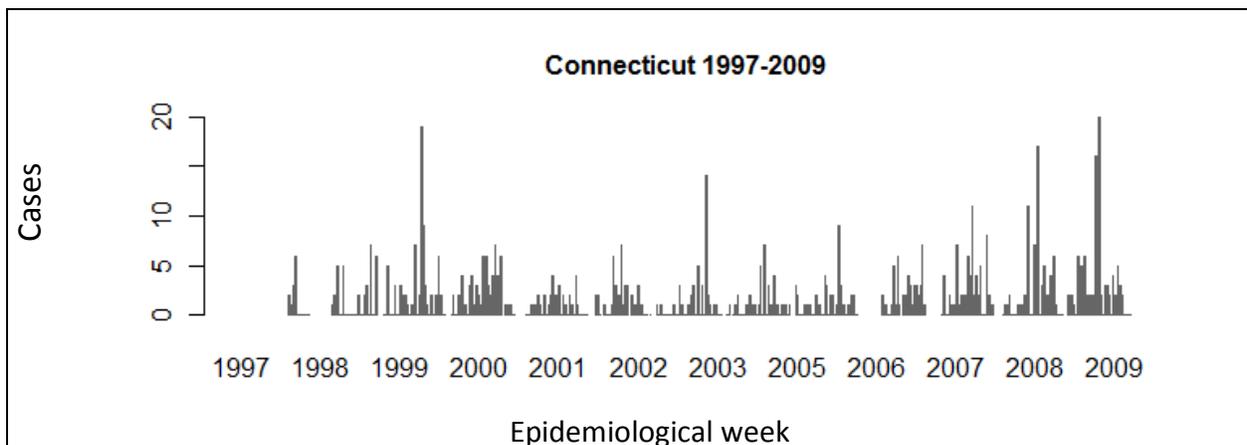
## Escherichia Coli



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

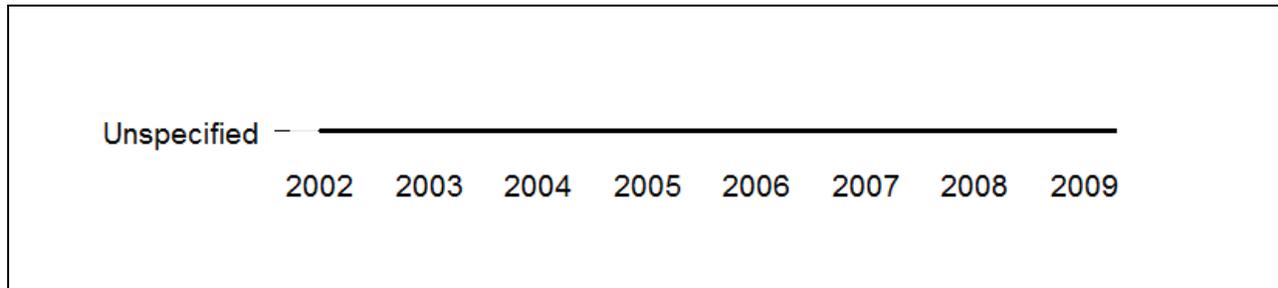
**Table D1**, Summary information for Escherichia Coli (O157:H7 PHLIS, EHEC 0157, and STEC)

Indicator	Connecticut
Report period	1997-2009
Total weeks	512
Total cases	763
Max. cases per year	93
Year (max)	2007
Max. cases per week	28
Week (max)	2006, wk 50
Average cases per year	59
95%CI	(44-74)
Average cases per week	1
95%CI	(1-1)



**Figure D2**, Number of cases reported for Escherichia Coli per epidemiological week

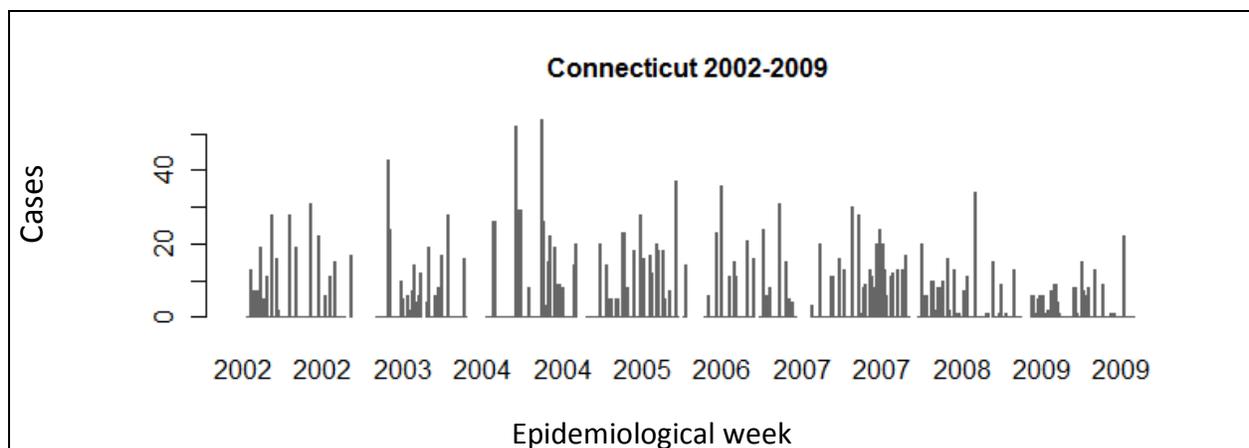
## Giardiasis



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

**Table D1**, Summary information for Giardiasis

Indicator	Connecticut
Report period	2002-2009
Total weeks	358
Total cases	7,710
Max. cases per year	5,777
Year (max)	2006
Max. cases per week	5,542
Week (max)	2006, wk 35
Average cases per year	964
95%CI	(-664-2,592)
Average cases per week	22
95%CI	(-8-52)



**Figure D2**, Number of cases reported for Giardiasis per epidemiological week

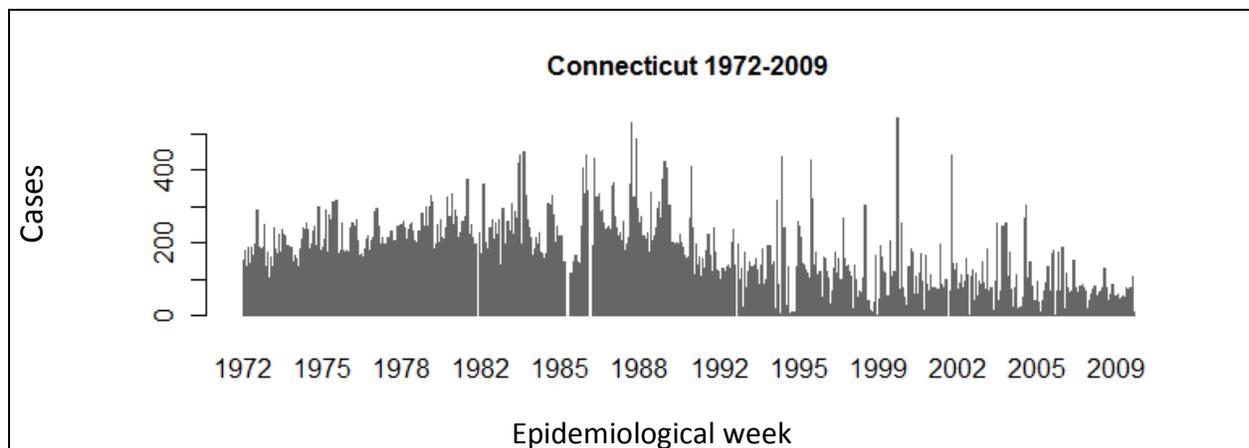
## Gonorrhea



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

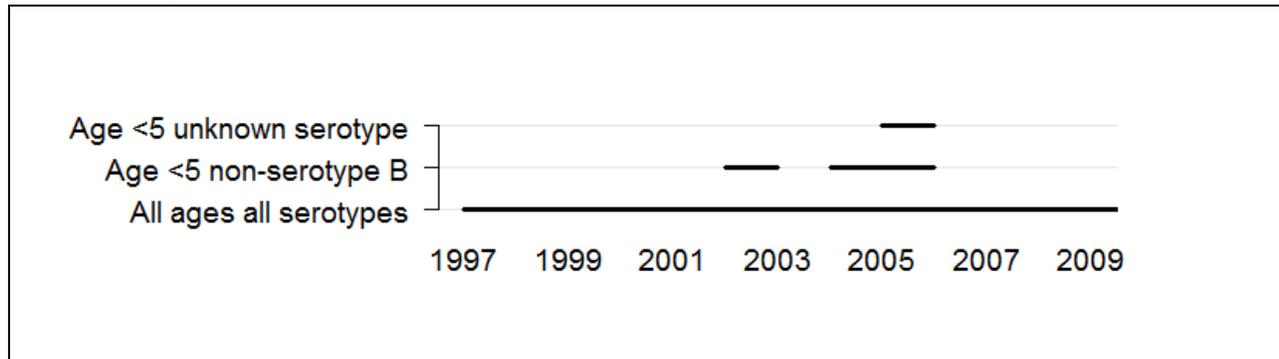
**Table D1**, Summary information for Gonorrhea (Unspecified and Civilian)

Indicator	Connecticut
Report period	1972-2009
Total weeks	1,803
Total cases	222,743
Max. cases per year	11,303
Year (max)	1987
Max. cases per week	2,020
Week (max)	2001, wk 44
Average cases per year	5,862
95%CI	(4,858-6,866)
Average cases per week	124
95%CI	(119-129)



**Figure D2**, Number of cases reported for Gonorrhea per epidemiological week

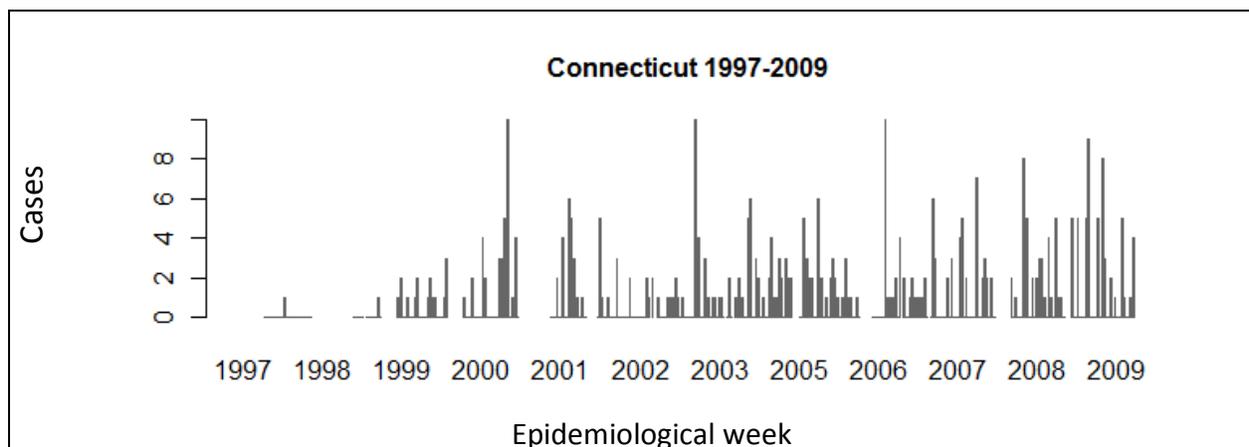
## Haemophilus Influenzae



**Figure D1,** Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

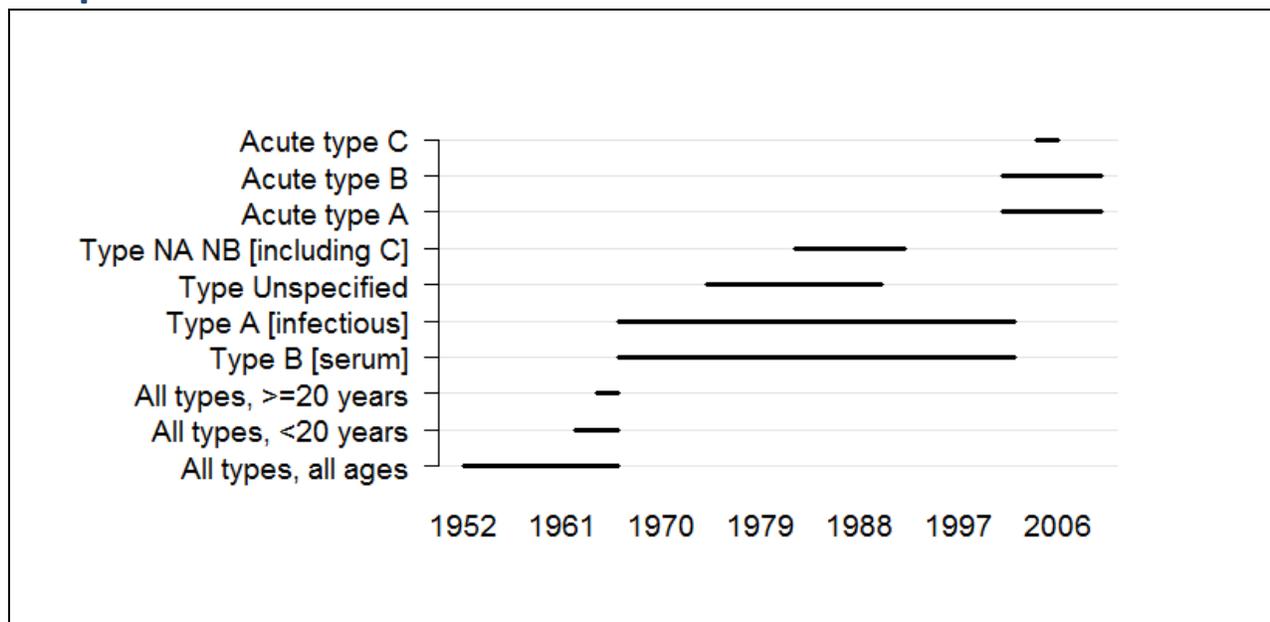
**Table D1,** Summary information for Haemophilus Influenzae (All ages all serotypes)

Indicator	Connecticut
Report period	1997-2009
Total weeks	514
Total cases	450
Max. cases per year	60
Year (max)	2005
Max. cases per week	24
Week (max)	2002, wk 46
Average cases per year	35
95%CI	(24-46)
Average cases per week	1
95%CI	(1-1)



**Figure D2,** Number of cases reported for Haemophilus Influenzae per epidemiological week

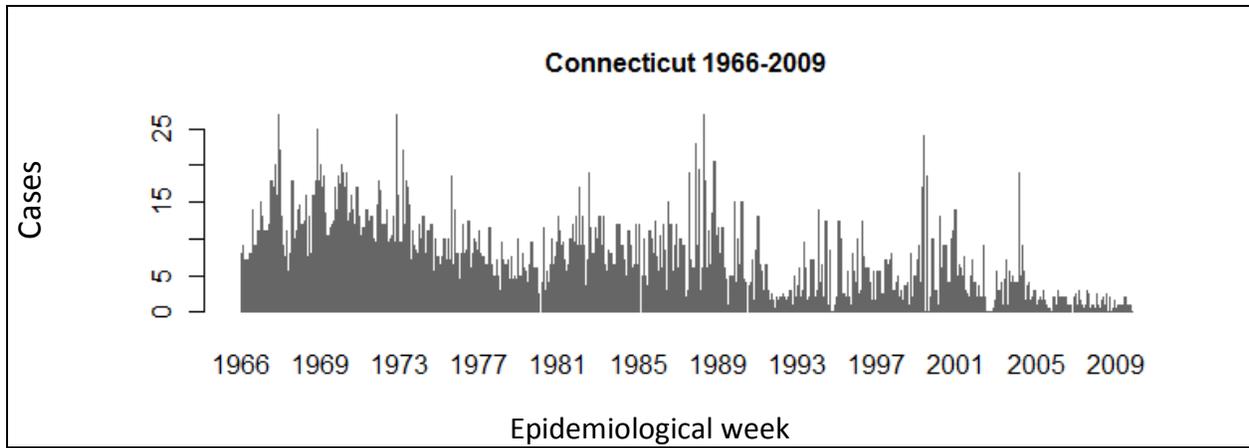
## Hepatitis



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

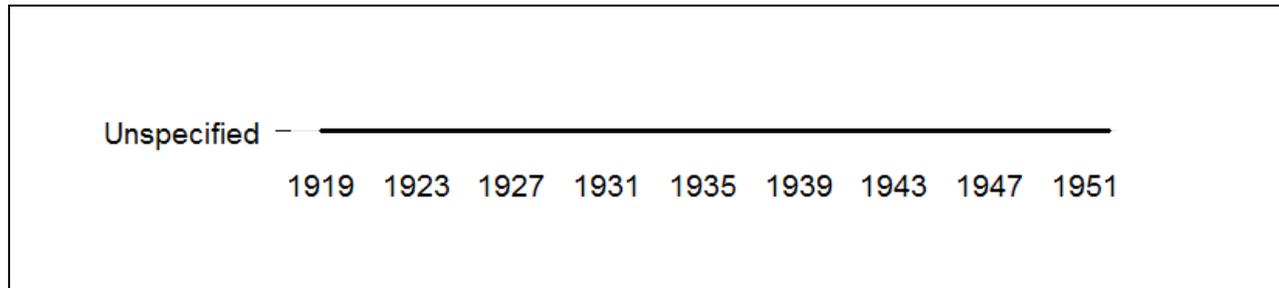
**Table D1**, Summary information for Hepatitis (Type A [Infectious], Acute type A, Type B [Serum], and Acute type B)

Indicator	Connecticut
Report period	1966-2009
Total weeks	2,076
Total cases	11,088
Max. cases per year	2,165
Year (max)	1970
Max. cases per week	1,634
Week (max)	1970, wk 44
Average cases per year	
before 1990	387
95%CI	(228-546)
after 1990	74
95%CI	(54-94)
Average cases per week	
before 1990	8
95%CI	(5-11)
after 1990	2
95%CI	(2-2)



**Figure D2,** Number of cases reported for Hepatitis per epidemiological week

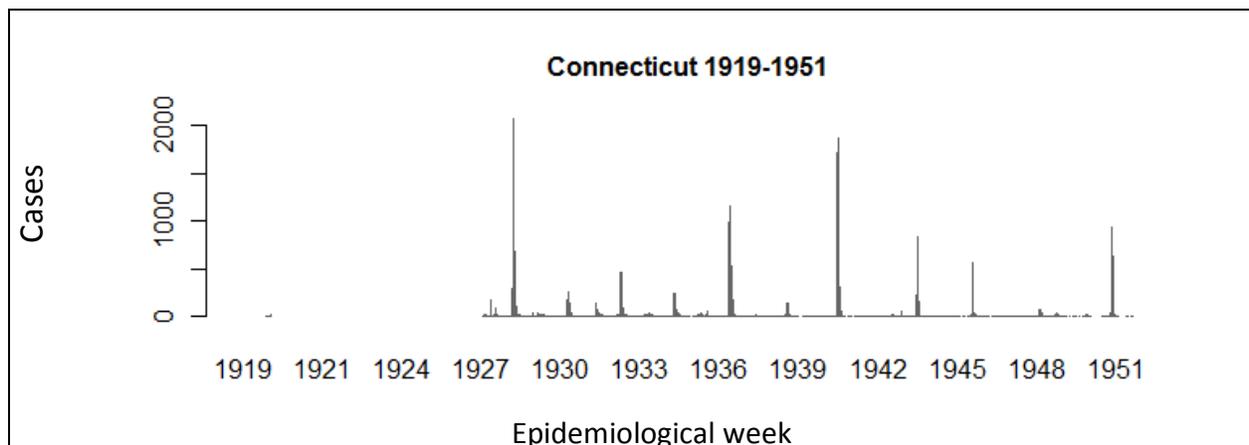
## Influenza



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

**Table D1**, Summary information for Influenza

Indicator	Connecticut
Report period	1919-1951
Total weeks	925
Total cases	42,297
Max. cases per year	16,795
Year (max)	1929
Max. cases per week	5,053
Week (max)	1929, wk 09
Average cases per year	1,567
95%CI	(246-2,888)
Average cases per week	46
95%CI	(28-64)



**Figure D2**, Number of cases reported for Influenza per epidemiological week

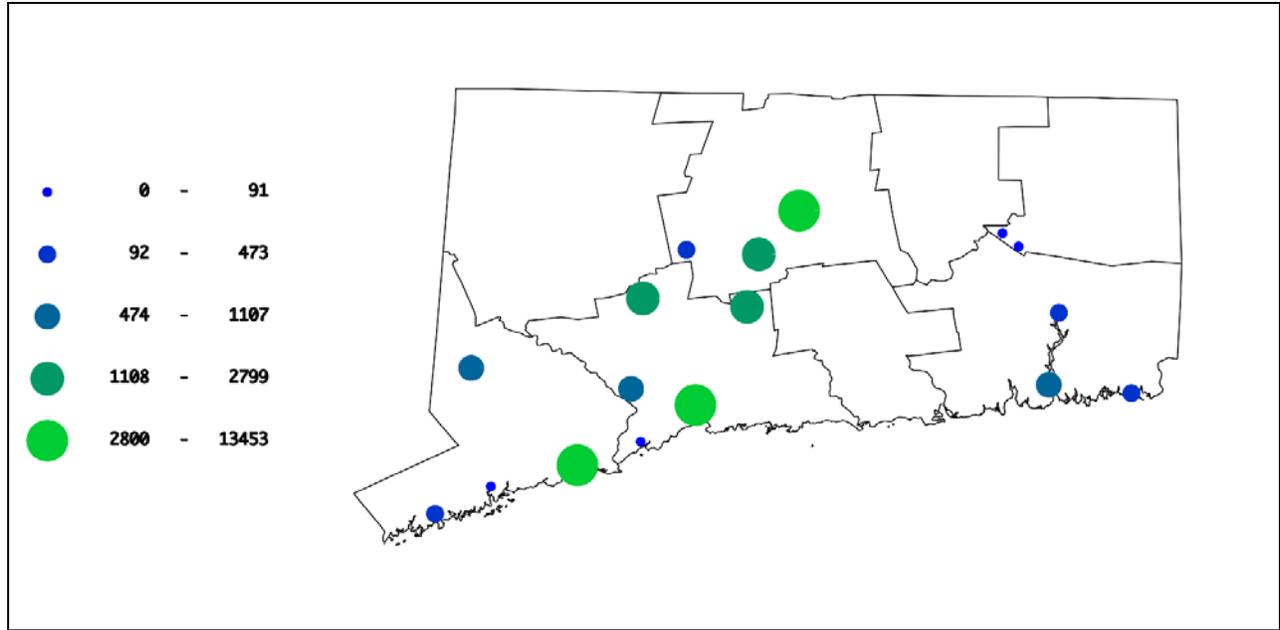
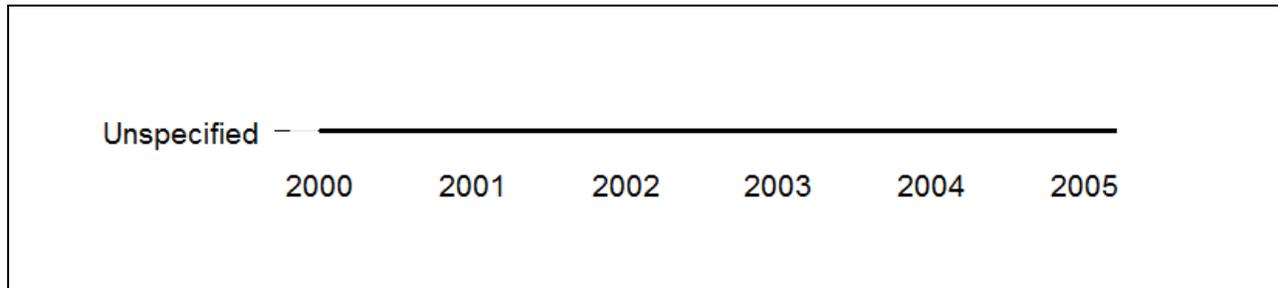


Figure D3, Total number of cases reported for Influenza by each city for the entire time period

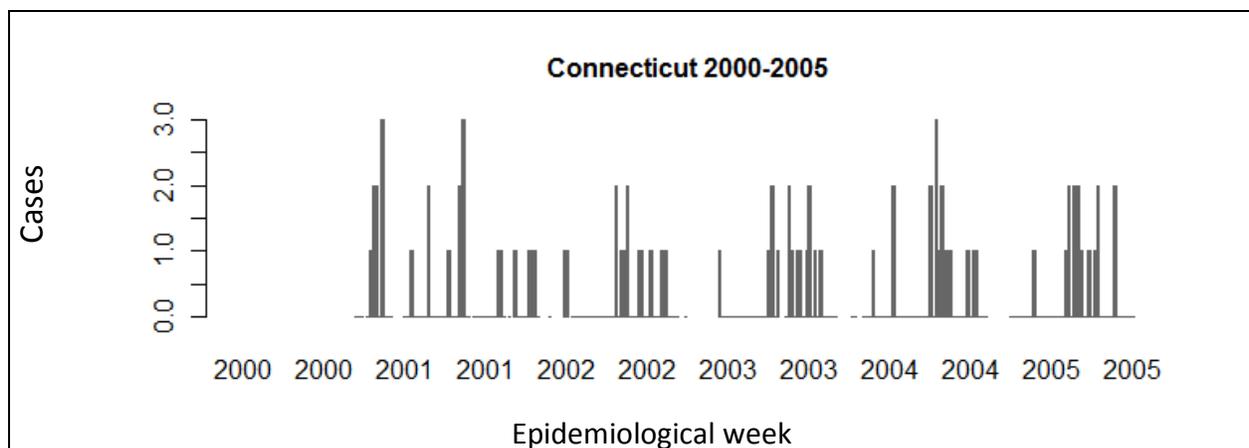
## Listeriosis



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

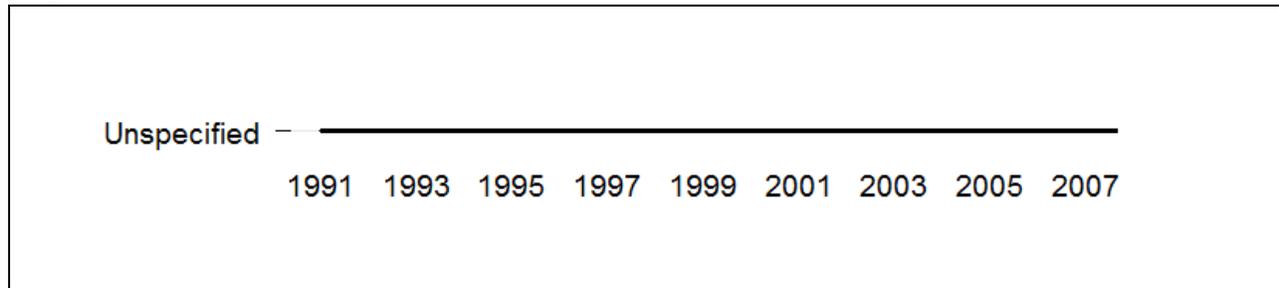
**Table D1**, Summary information for Listeriosis

Indicator	Connecticut
Report period	2000-2005
Total weeks	227
Total cases	102
Max. cases per year	26
Year (max)	2001
Max. cases per week	11
Week (max)	2001, wk 28
Average cases per year	17
95%CI	(11-23)
Average cases per week	0
95%CI	(0-0)



**Figure D2**, Number of cases reported for Listeriosis per epidemiological week

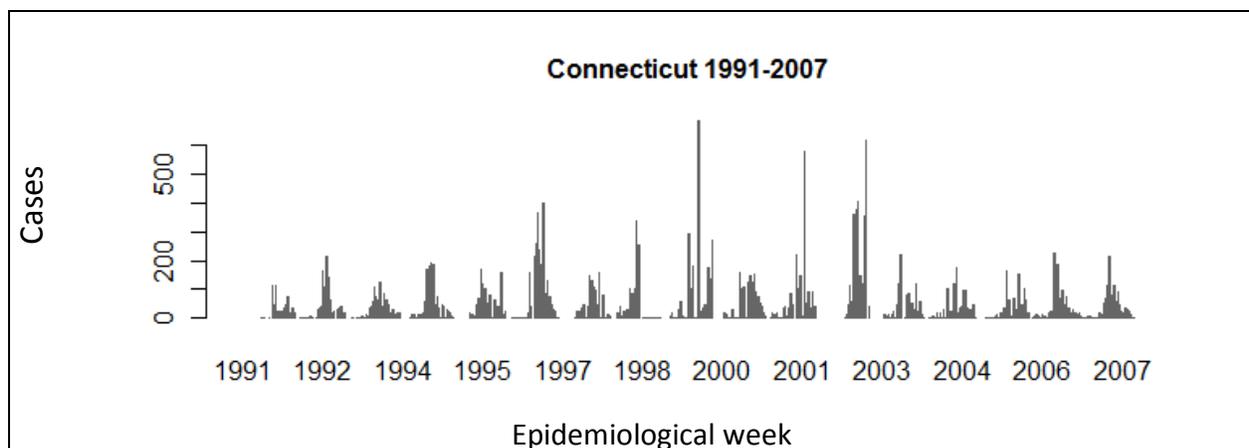
## Lyme Disease



**Figure D1,** Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

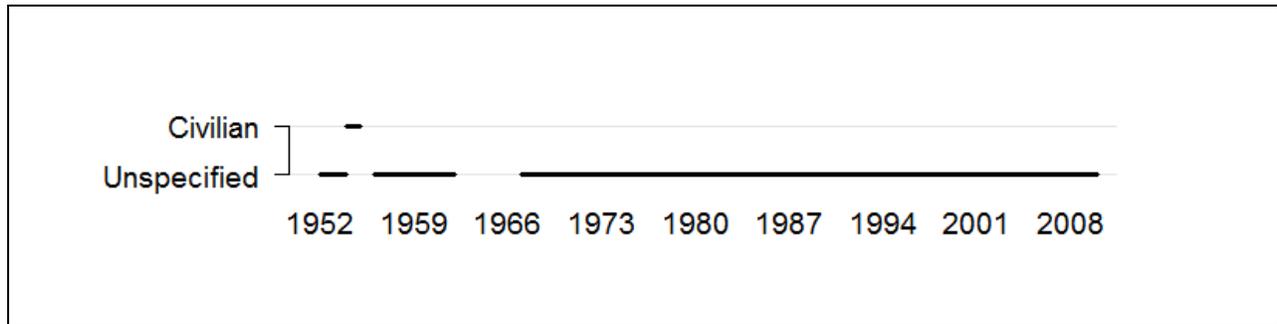
**Table D1,** Summary information for Lyme Disease

Indicator	Connecticut
Report period	1991-2007
Total weeks	679
Total cases	31,176
Max. cases per year	3,567
Year (max)	2002
Max. cases per week	1,804
Week (max)	1998, wk 52
Average cases per year	1,834
95%CI	(1,437-2,231)
Average cases per week	46
95%CI	(38-54)



**Figure D2,** Number of cases reported for Lyme Disease per epidemiological week

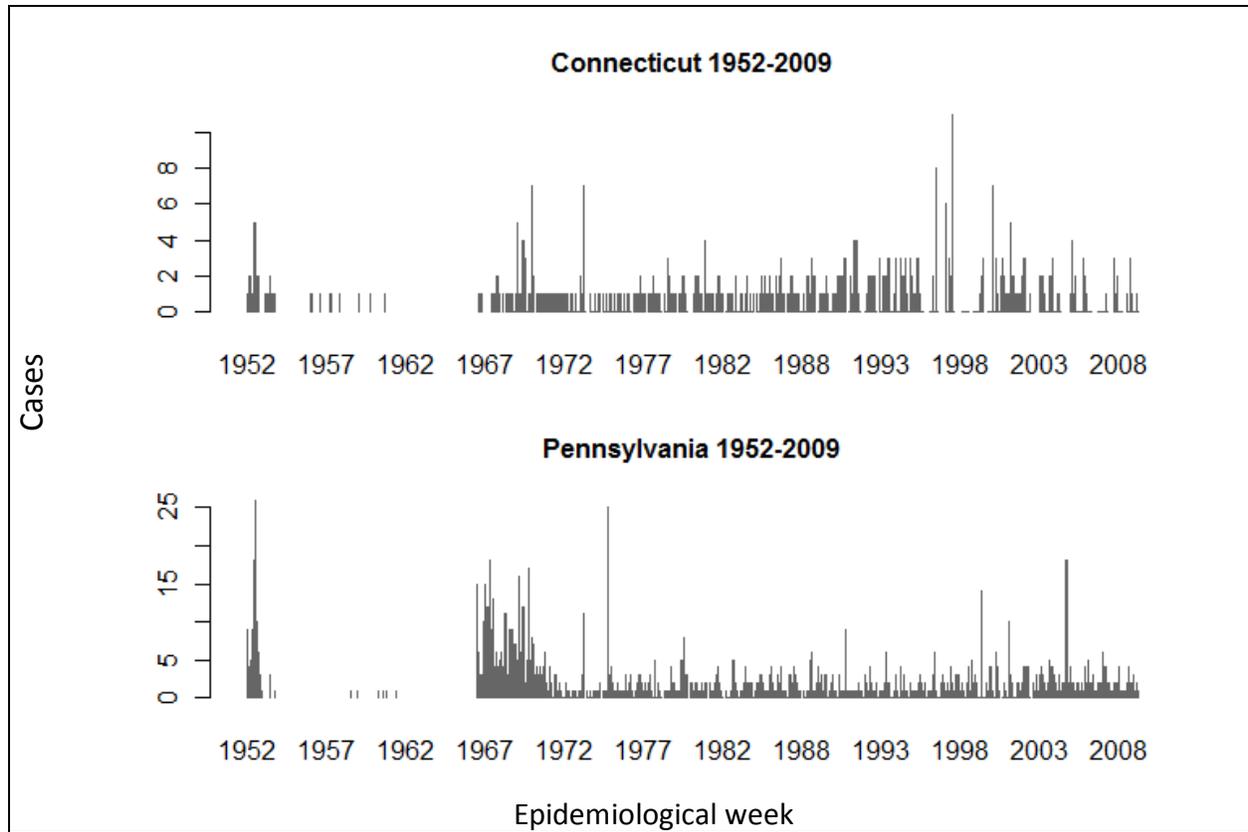
## Malaria



**Figure D1,** Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

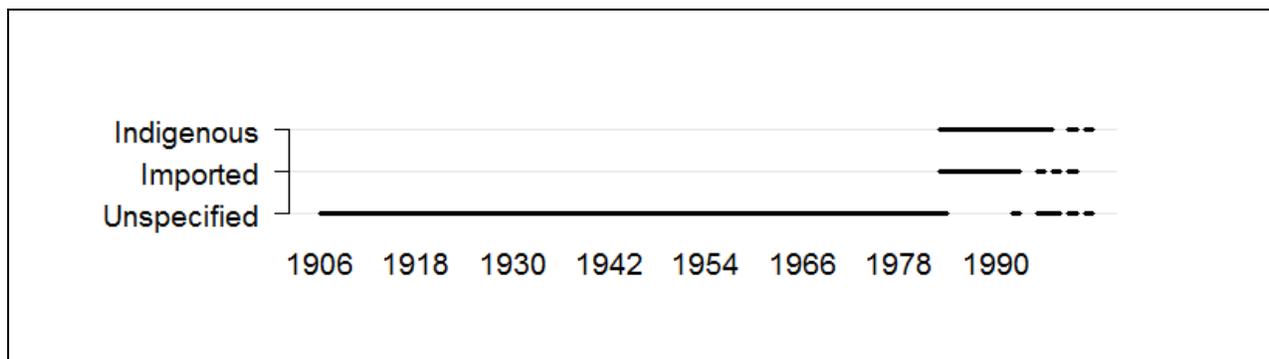
**Table D1,** Summary information for Malaria (Unspecified and Civilian)

Indicator	Connecticut
Report period	1952-2009
Total weeks	1,626
Total cases	656
Max. cases per year	40
Year (max)	1970
Max. cases per week	17
Week (max)	1970, wk 33
Average cases per year	13
95%CI	(11-15)
Average cases per week	0
95%CI	(0-0)



**Figure D2**, Number of cases reported for Malaria per epidemiological week

## Measles



**Figure D1,** Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

**Table D1,** Summary information for Measles (Unspecified)

Indicator	Connecticut	Hartford	Bridgeport	New Haven
Report period	1927-2001	1906-1953	1909-1953	1915-1953
Total weeks	2,875	1,775	1,755	1,626
Total cases	373,820	23,096	11,754	38,431
Max. cases per year	29,227	2,675	853	5,969
Year (max)	1952	1939	1923	1935
Max. cases per week	2,238	278	138	641
Week (max)	1955, wk 06	1939, wk 05	1926, wk 01	1935, wk 13
Average cases per year				
before 1970	8,349	513	280	1,068
95%CI	(6,133-10,565)	(345-681)	(209-351)	(592-1,544)
after 1970	340	-	-	-
95%CI	(65-615)	-	-	-
Average cases per week				
before 1970	169	13	7	24
95%CI	(157-181)	(12-14)	(6-8)	(21-27)
after 1970	9	-	-	-
95%CI	(7-11)	-	-	-

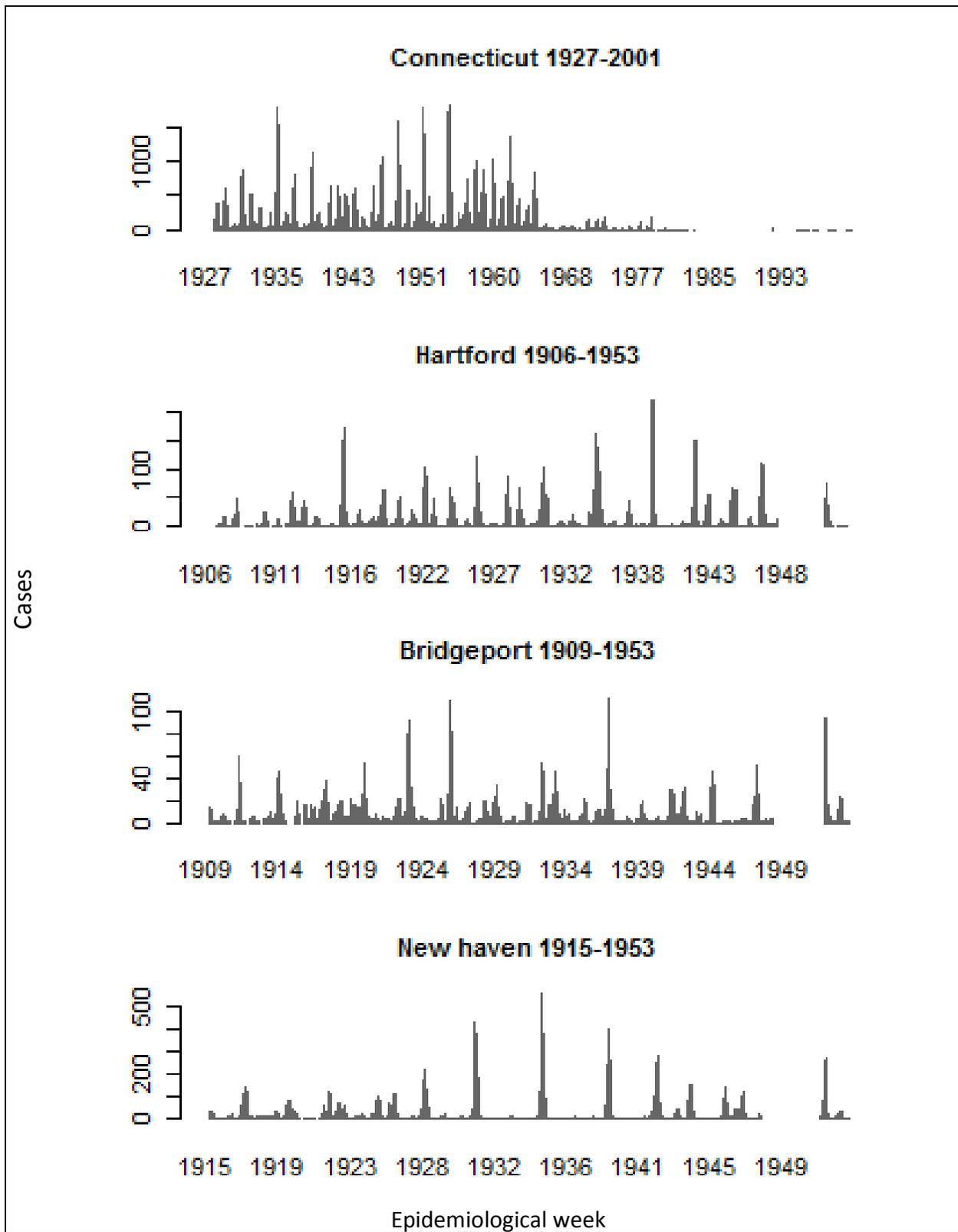


Figure D2, Number of cases reported for Measles per epidemiological week

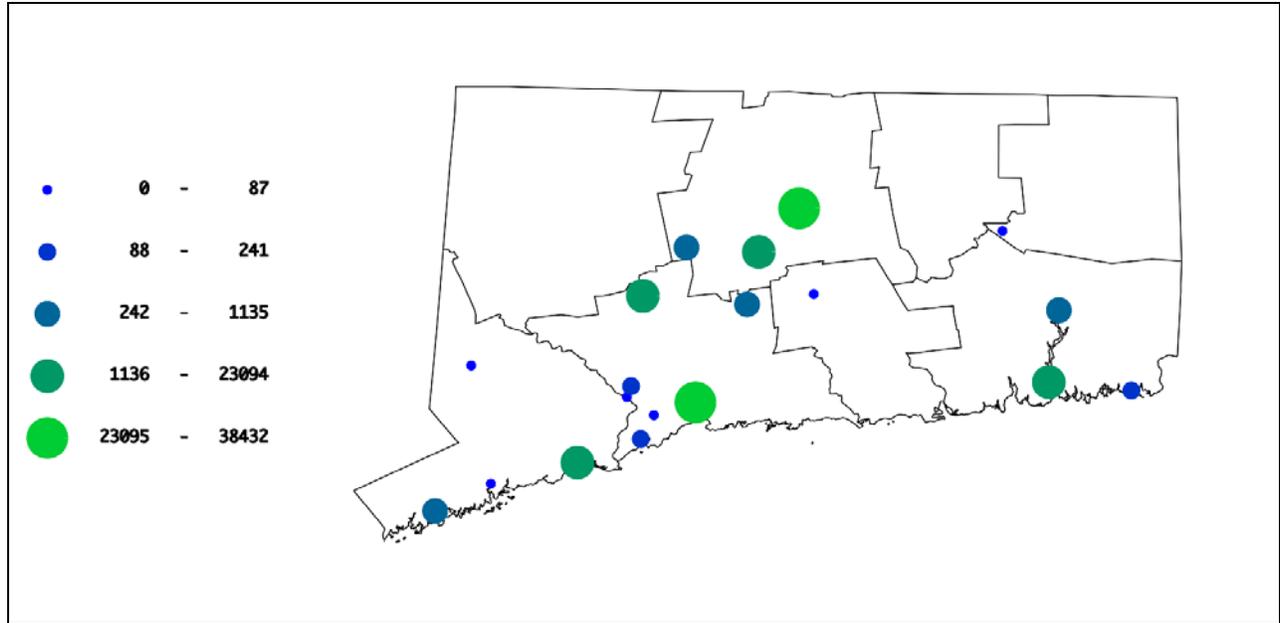


Figure D3, Total number of cases reported by each city for Measles for the entire time period

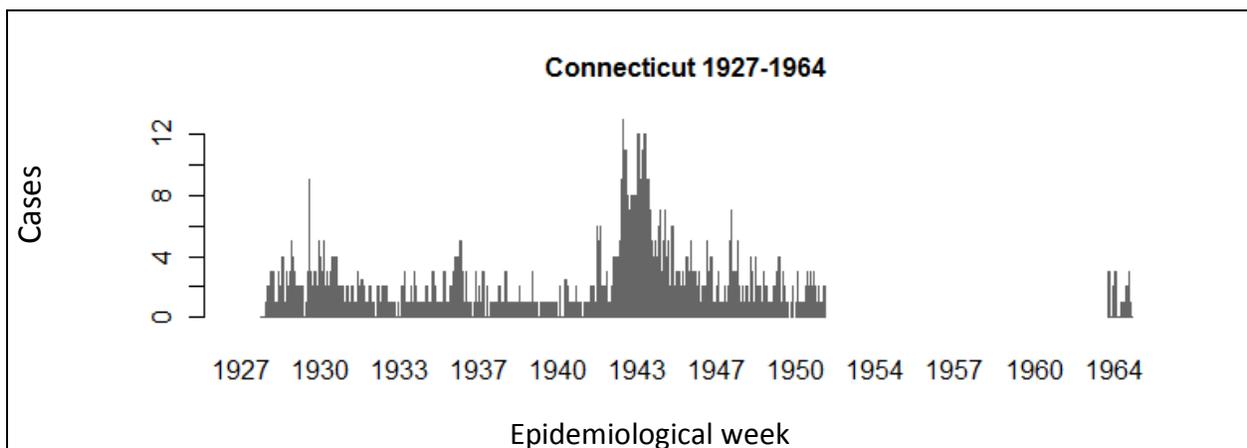
## Meningitis



**Figure D1,** Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

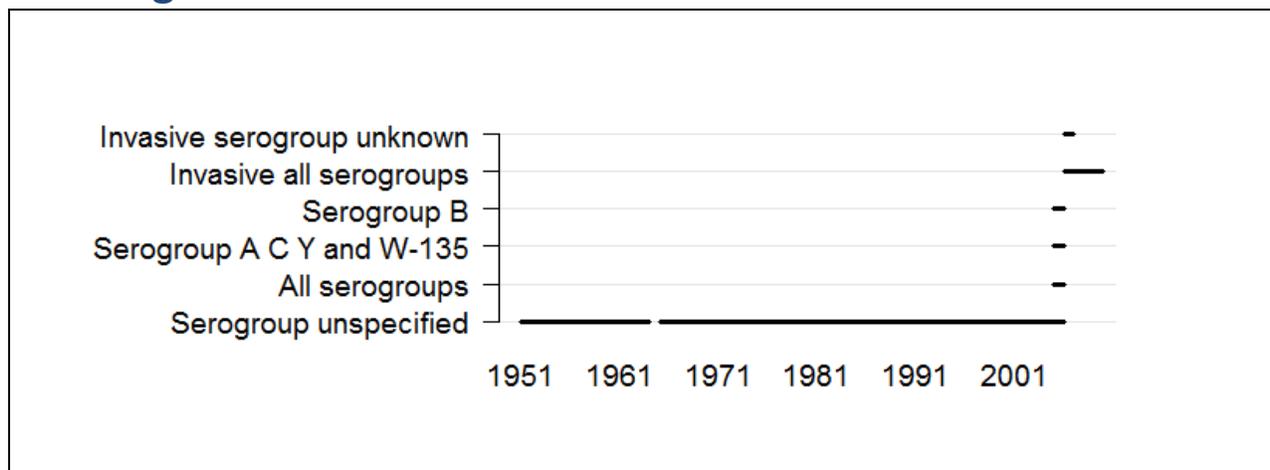
**Table D1,** Summary information for Meningitis (Meningococcus)

Indicator	Connecticut
Report period	1927-1964
Total weeks	1,212
Total cases	1,650
Max. cases per year	297
Year (max)	1943
Max. cases per week	17
Week (max)	1943, wk 15
Average cases per year	63
95%CI	(35-91)
Average cases per week	1
95%CI	(1-1)



**Figure D2,** Number of cases reported for Meningitis per epidemiological week

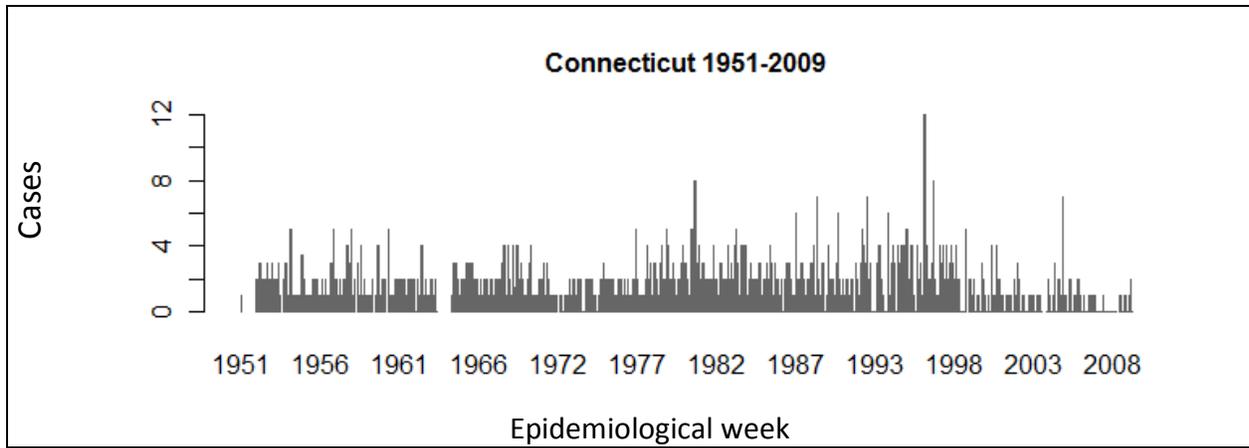
## Meningococcal Disease



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

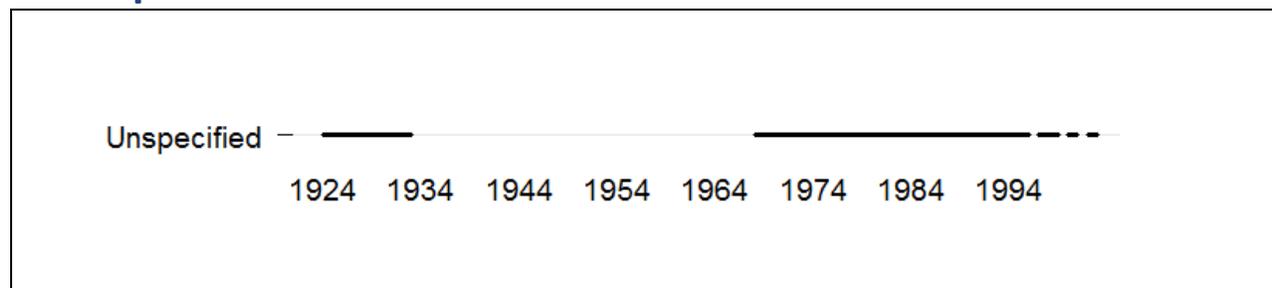
**Table D1**, Summary information for Meningococcal Disease (Serogroup unspecified and Invasive all serogroups)

Indicator	Connecticut
Report period	1951-2009
Total weeks	2,467
Total cases	2,351
Max. cases per year	457
Year (max)	1952
Max. cases per week	414
Week (max)	1952, wk 53
Average cases per year	
before 1980	50
95%CI	(20-80)
after 1980	31
95%CI	(24-38)
Average cases per week	
before 1980	1
95%CI	(0-2)
after 1980	1
95%CI	(1-1)



**Figure D2,** Number of cases reported for Meningococcal Disease per epidemiological week

## Mumps



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

**Table D1**, Summary information for Mumps

<b>Indicator</b>	<b>Connecticut</b>	<b>Hartford</b>	<b>Bridgeport</b>	<b>New Haven</b>
Report period	1968-2002	1924-1932	1924-1932	1924-1932
Total weeks	1,393	399	447	446
Total cases	19,968	1,169	294	3,625
Max. cases per year	3,991	332	86	1,105
Year (max)	1969	1932	1931	1928
Max. cases per week	230	42	8	128
Week (max)	1968, wk 12	1932, wk 02	1927, wk 10	1928, wk 10
Average cases per year				
before 1980	1,519	130	33	403
95%CI	(659-2,379)	(55-205)	(12-54)	(94-712)
after 1980	12	-	-	-
95%CI	(6-18)	-	-	-
Average cases per week				
before 1980	30	3	1	8
95%CI	(27-33)	(3-3)	(1-1)	(6-10)
after 1980	0	-	-	-
95%CI	(0-0)	-	-	-

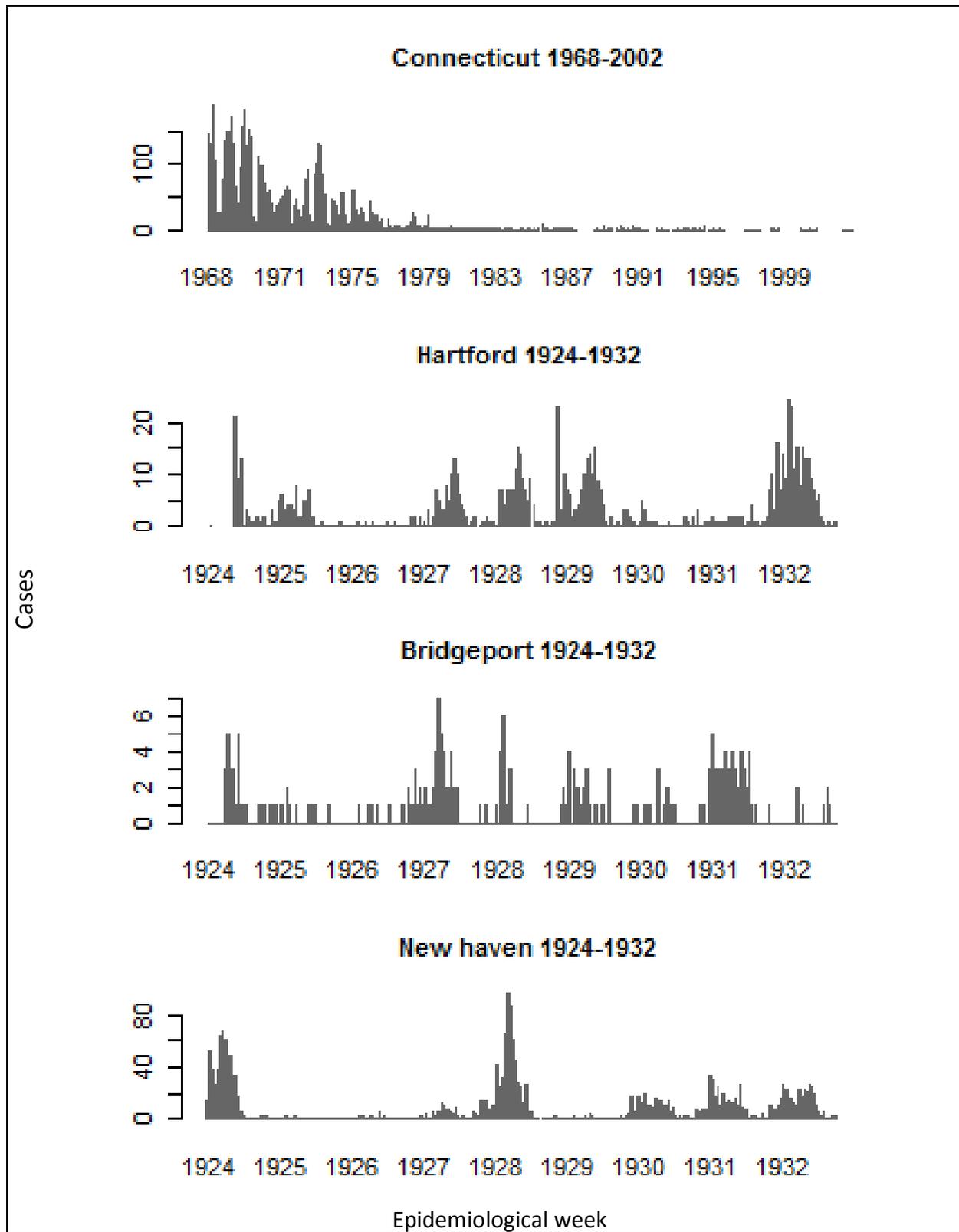


Figure D2, Number of cases reported for Mumps per epidemiological week

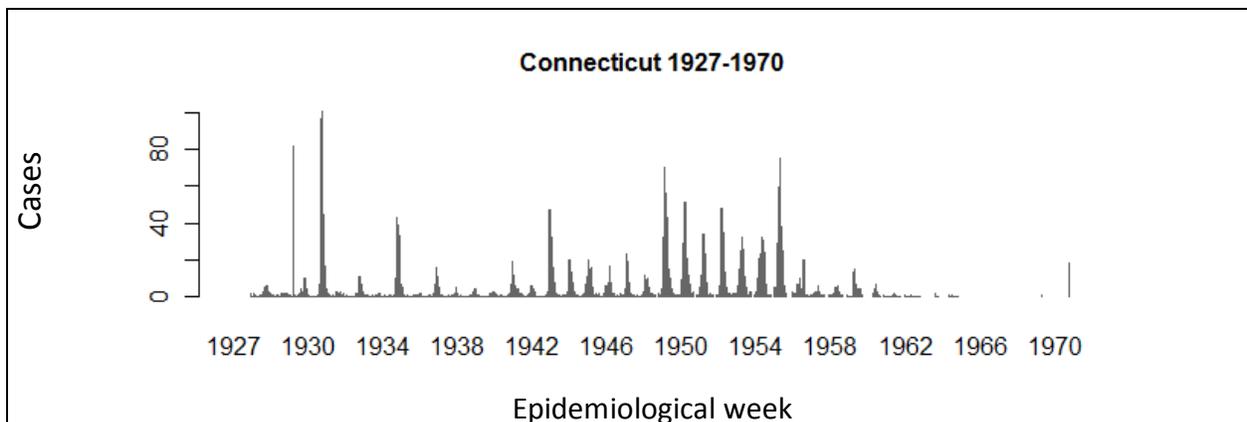
## Poliomyelitis



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

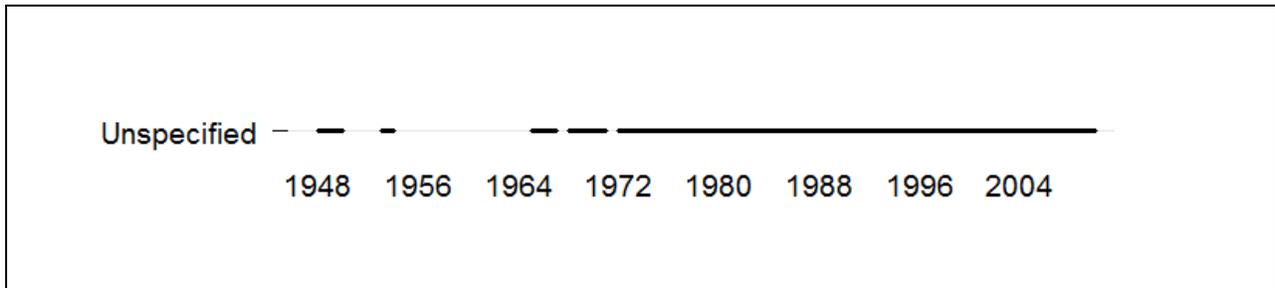
**Table D1**, Summary information for Poliomyelitis (Total)

Indicator	Connecticut
Report period	1927-1970
Total weeks	1,627
Total cases	6,998
Max. cases per year	1,075
Year (max)	1931
Max. cases per week	162
Week (max)	1931, wk 35
Average cases per year	
before 1960	205
95%CI	(122-288)
after 1960	5
95%CI	(-2-12)
Average cases per week	
before 1960	5
95%CI	(4-6)
after 1960	0
95%CI	(0-0)



**Figure D2**, Number of cases reported for Poliomyelitis per epidemiological week

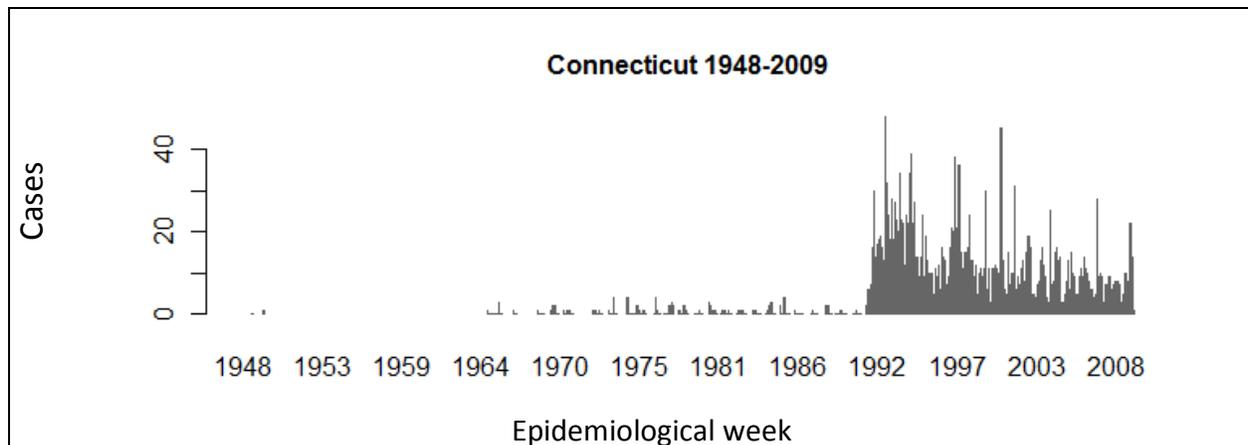
## Rabies in Animals



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

**Table D1**, Summary information for Rabies in Animals

Indicator	Connecticut
Report period	1948-2009
Total weeks	1,615
Total cases	7,491
Max. cases per year	1,188
Year (max)	1991
Max. cases per week	1,030
Week (max)	1991, wk 29
Average cases per year	166
95%CI	(88-244)
Average cases per week	5
95%CI	(4-6)



**Figure D2**, Number of cases reported for Rabies in Animals per epidemiological week

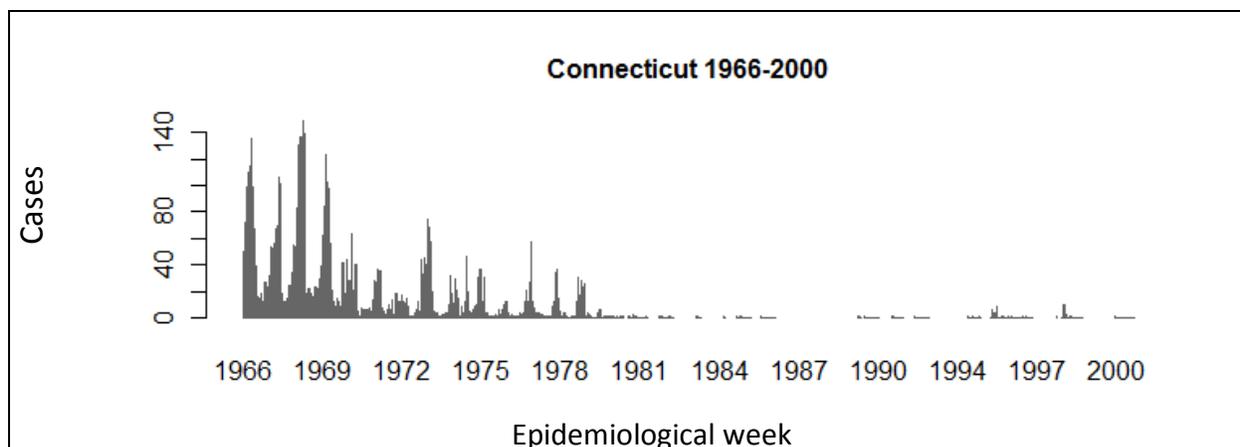
## Rubella



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

**Table D1**, Summary information for Rubella

Indicator	Connecticut
Report period	1966-2000
Total weeks	1,167
Total cases	12,768
Max. cases per year	3,039
Year (max)	1968
Max. cases per week	213
Week (max)	1968, wk 24
Average cases per year	
before 1970	1,868
95%CI	(774-2,962)
after 1970	137
95%CI	(46-228)
Average cases per week	
before 1970	37
95%CI	(32-42)
after 1970	4
95%CI	(3-5)



**Figure D2**, Number of cases reported for Rubella per epidemiological week

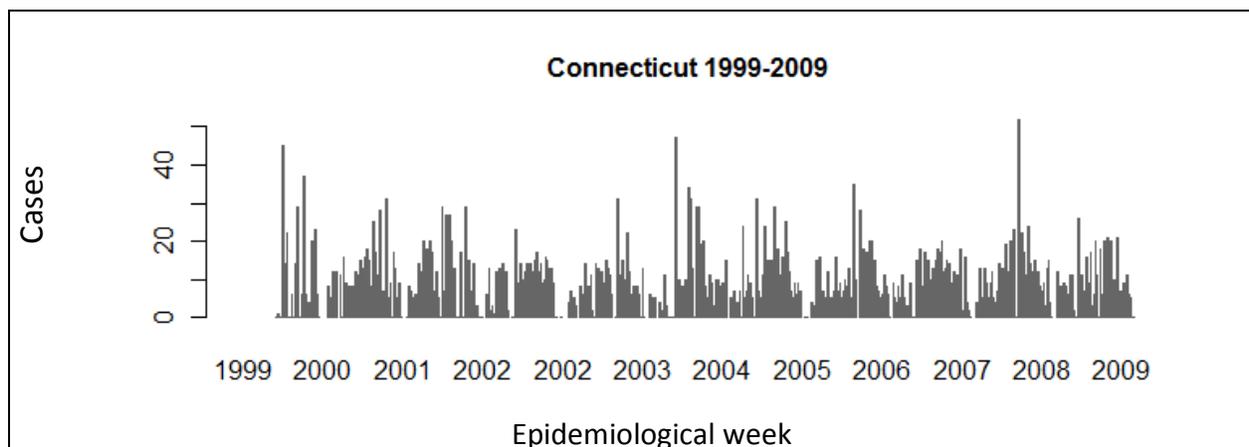
## Salmonellosis



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

**Table D1**, Summary information for Salmonellosis (PHLIS and Unspecified)

Indicator	Connecticut
Report period	1999-2009
Total weeks	505
Total cases	4,640
Max. cases per year	467
Year (max)	2008
Max. cases per week	121
Week (max)	1999, wk 24
Average cases per year	422
95%CI	(399-445)
Average cases per week	9
95%CI	(8-10)



**Figure D2**, Number of cases reported for Salmonellosis per epidemiological week

## Scarlet Fever



**Figure D1,** Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

**Table D1,** Summary information for Scarlet Fever (Including streptococcal sore throat and Unspecified)

Indicator	Connecticut	Hartford	Bridgeport	New Haven
Report period	1927-1969	1906-1953	1909-1953	1915-1953
Total weeks	1,845	1,978	1,924	1,678
Total cases	311,578	10,837	10,732	7,417
Max. cases per year	44,872	1,038	1,172	651
Year (max)	1967	1924	1937	1921
Max. cases per week	1,817	68	80	59
Week (max)	1967, wk 05	1924, wk 15	1937, wk 18	1921, wk 07
Average cases per year	8,421	241	256	206
95%CI	(4,015-12,827)	(190-292)	(191-321)	(149-263)
Average cases per week	169	5	6	4
95%CI	(156-182)	(5-5)	(6-6)	(4-4)

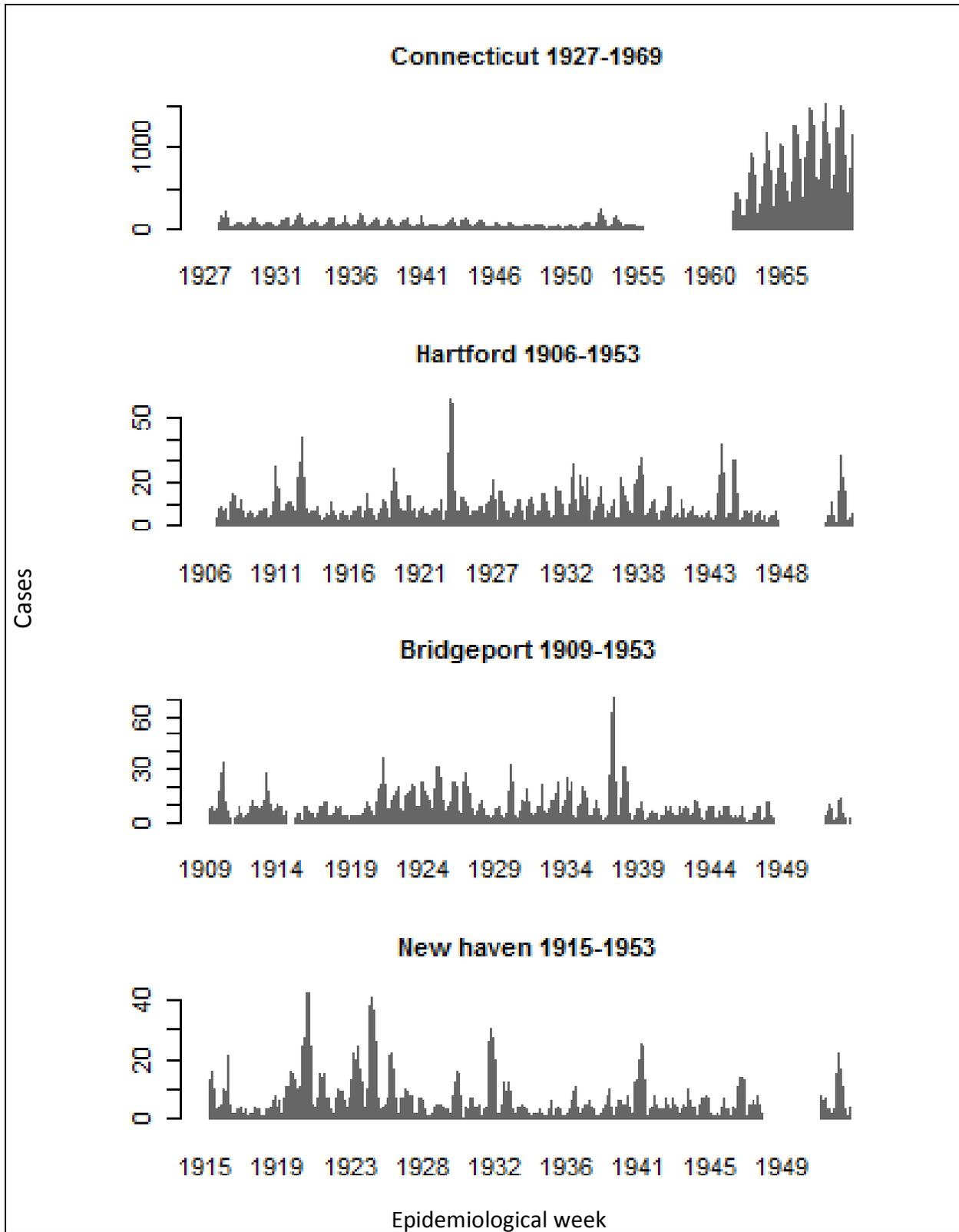


Figure D2, Number of cases reported for Scarlet Fever per epidemiological week

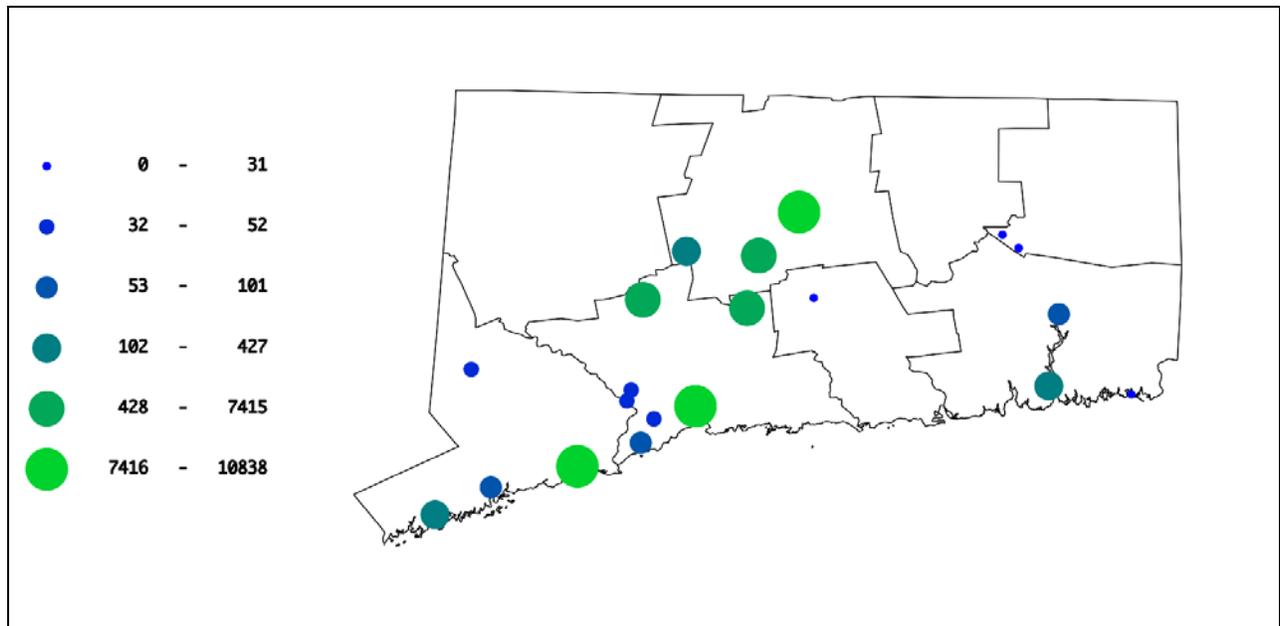
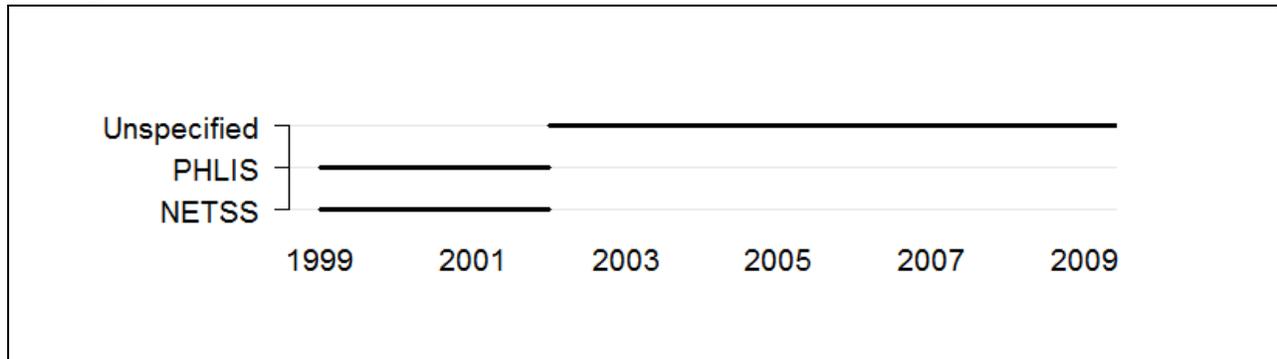


Figure D3, Total number of all cases reported for Scarlet Fever by each city for the entire time period

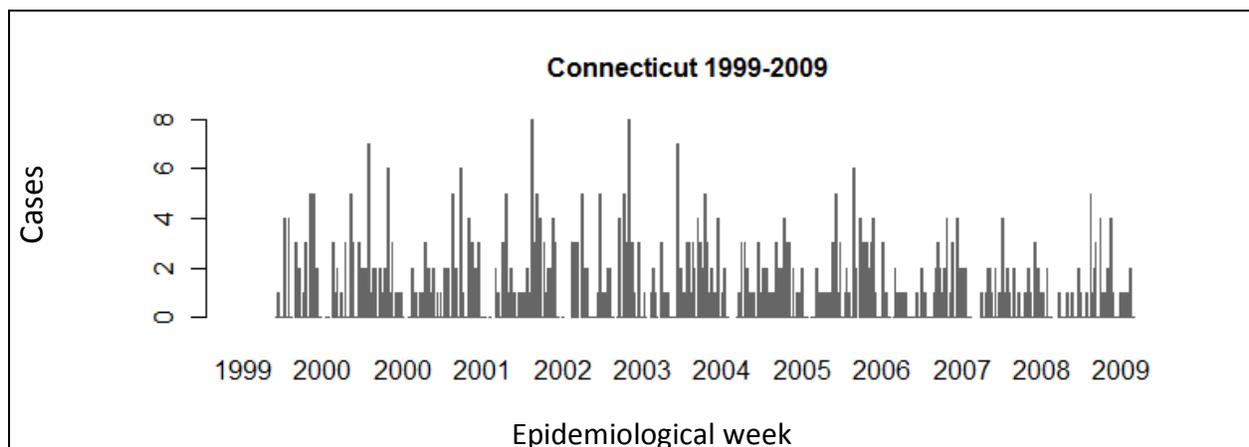
## Shigellosis



**Figure D1,** Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

**Table D1,** Summary information for Shigellosis (Unspecified and PHLIS)

Indicator	Connecticut
Report period	1999-2009
Total weeks	500
Total cases	879
Max. cases per year	278
Year (max)	1999
Max. cases per week	248
Week (max)	1999, wk 33
Average cases per year	80
95%CI	(34-126)
Average cases per week	2
95%CI	(1-3)



**Figure D2,** Number of cases reported for Shigellosis per epidemiological week

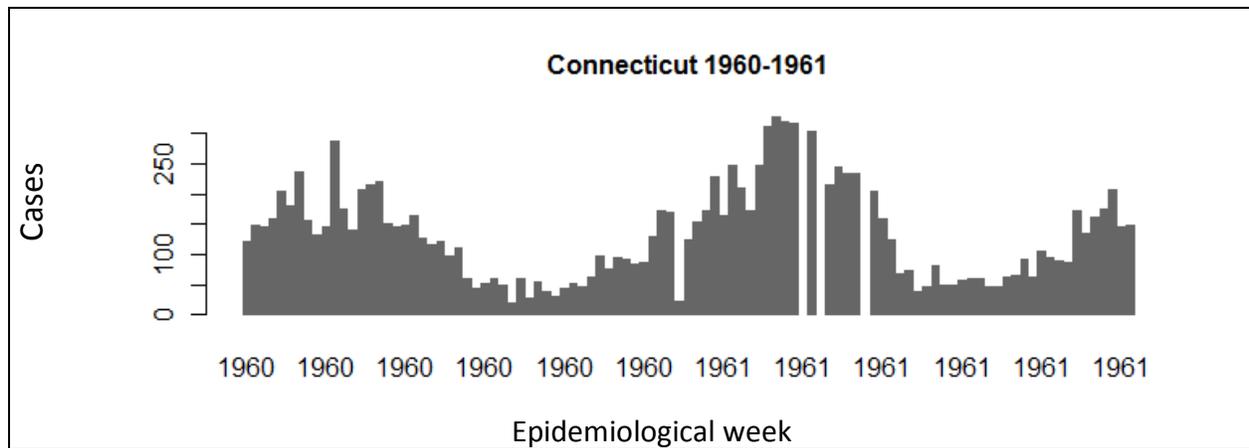
## Streptococcal Sore Throat



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

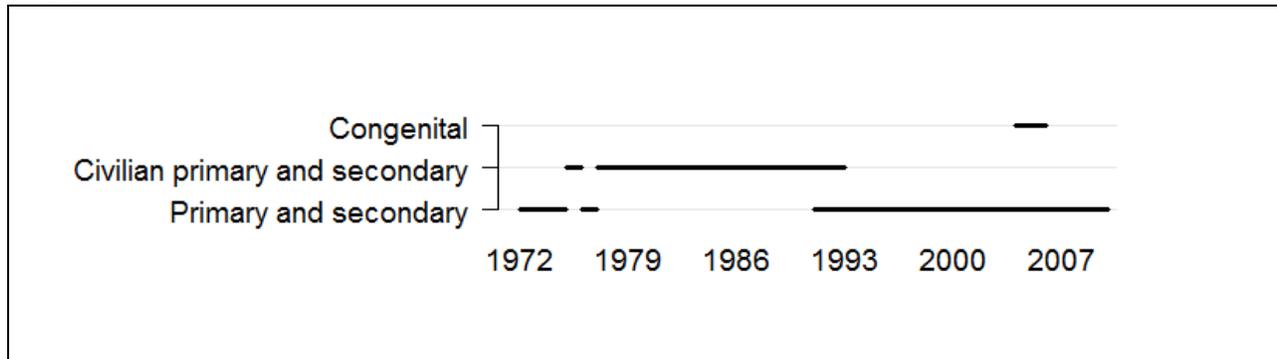
**Table D1**, Summary information for Streptococcal Sore Throat

Indicator	Connecticut
Report period	1960-1961
Total weeks	101
Total cases	14,065
Max. cases per year	8,023
Year (max)	1961
Max. cases per week	425
Week (max)	1961, wk 08
Average cases per year	7,032
95%CI	(-5,553-19,617)
Average cases per week	139
95%CI	(122-156)



**Figure D2**, Number of cases reported for Streptococcal Sore Throat per epidemiological week

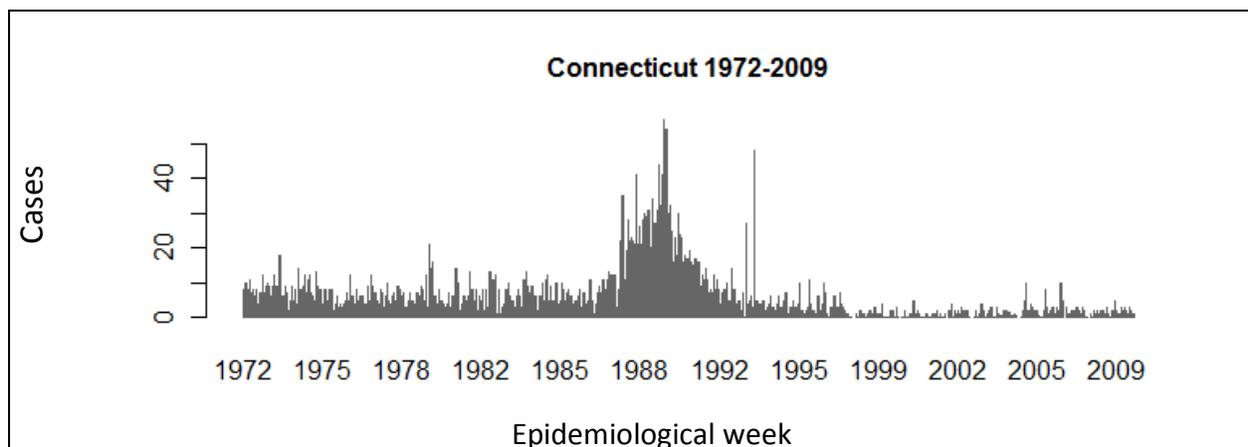
## Syphilis



**Figure D1,** Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

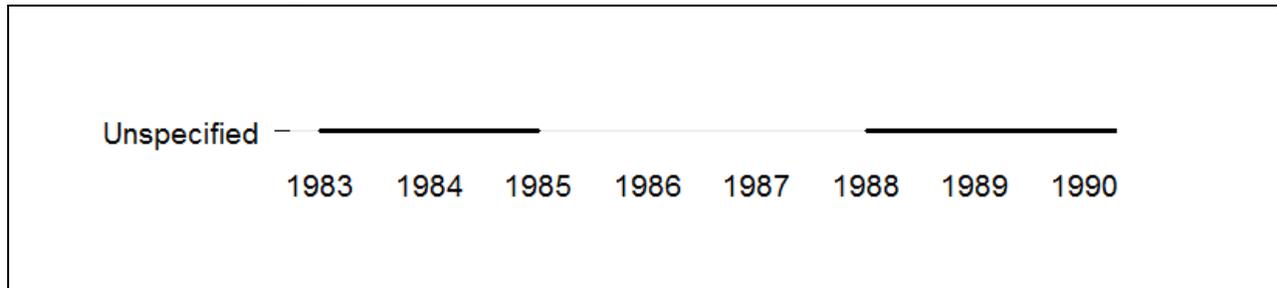
**Table D1,** Summary information for Syphilis (Primary and secondary and Civilian primary and secondary)

Indicator	Connecticut
Report period	1972-2009
Total weeks	1,790
Total cases	7,242
Max. cases per year	1,099
Year (max)	1989
Max. cases per week	90
Week (max)	1995, wk 47
Average cases per year	191
95%CI	(117-265)
Average cases per week	4
95%CI	(4-4)



**Figure D2,** Number of cases reported for Syphilis per epidemiological week

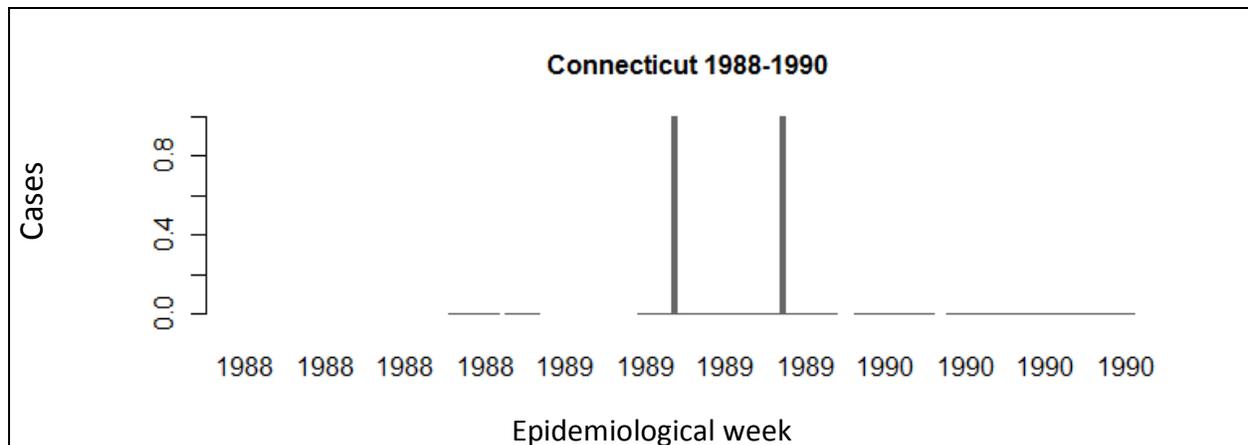
## Toxic Shock Syndrome



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

**Table D1**, Summary information for Toxic shock syndrome

Indicator	Connecticut
Report period	1988-1990
Total weeks	97
Total cases	2
Max. cases per year	2
Year (max)	1989
Max. cases per week	1
Week (max)	1989, wk 24
Average cases per year	1
95%CI	(-2-4)
Average cases per week	0
95%CI	(0-0)



**Figure D2**, Number of cases reported for Toxic Shock Syndrome per epidemiological week

## Tuberculosis



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

**Table D1**, Summary information for Tuberculosis (Unspecified)

Indicator	Connecticut	Hartford	Bridgeport	New Haven
Report period	1975-2005	1906-1923	1909-1923	1915-1923
Total weeks	1,463	720	645	432
Total cases	7,013	3,105	2,789	3,205
Max. cases per year	2,840	300	261	461
Year (max)	1991	1915	1918	1920
Max. cases per week	2,702	25	27	38
Week (max)	1991, wk 29	1915, wk 07	1910, wk 51	1921, wk 13
Average cases per year	226	172	186	356
95%CI	(47-405)	(139-205)	(160-212)	(305-407)
Average cases per week	5	4	4	7
95%CI	(1-9)	(4-4)	(4-4)	(7-7)

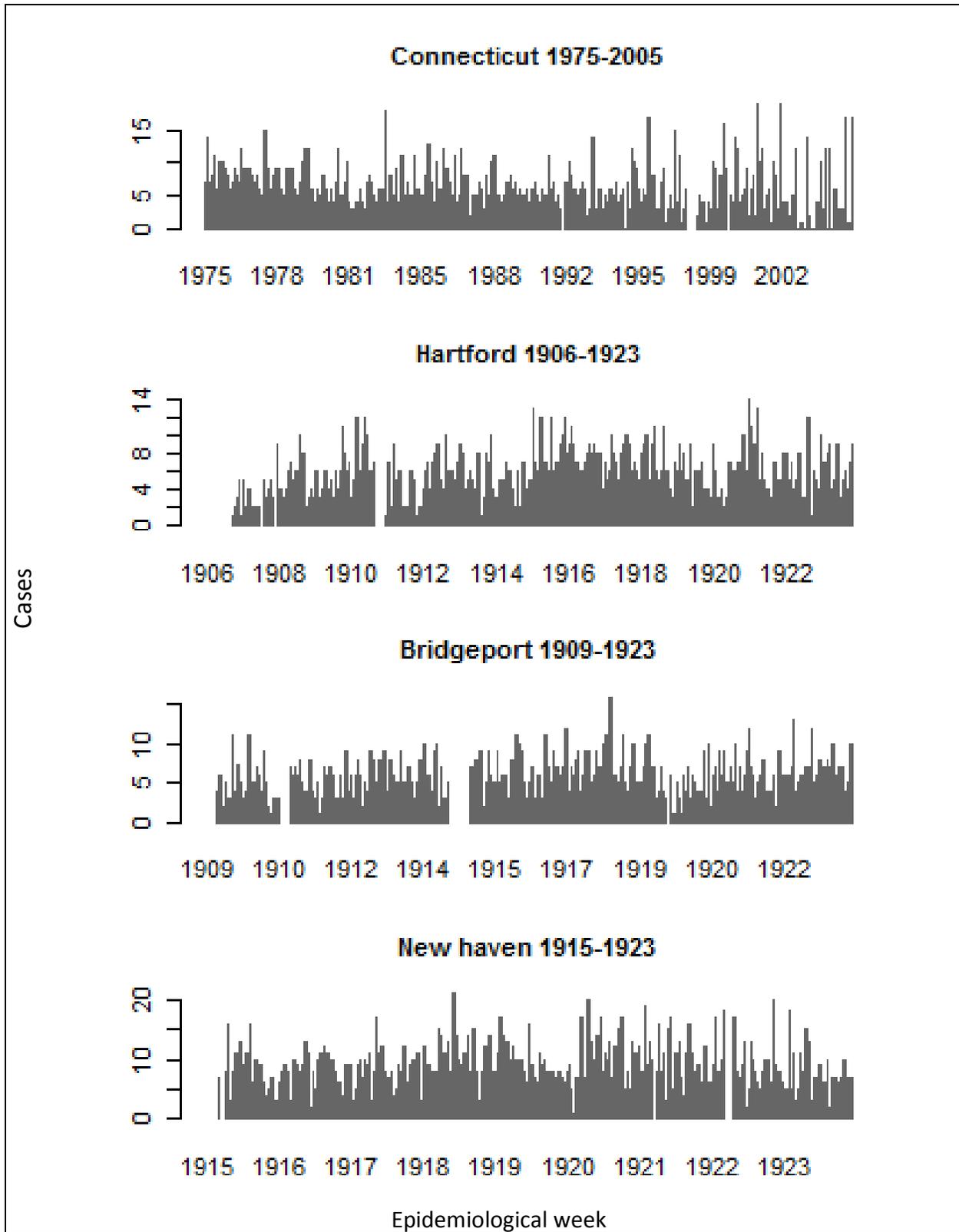


Figure D2, Number of cases reported for Tuberculosis per epidemiological week

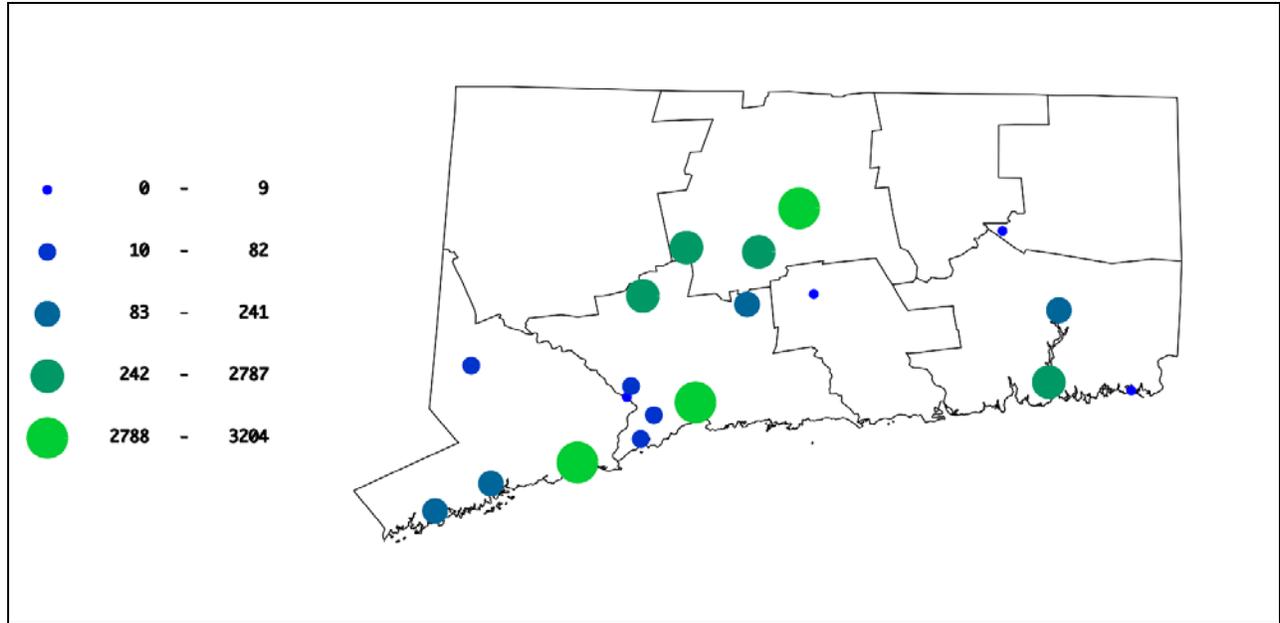
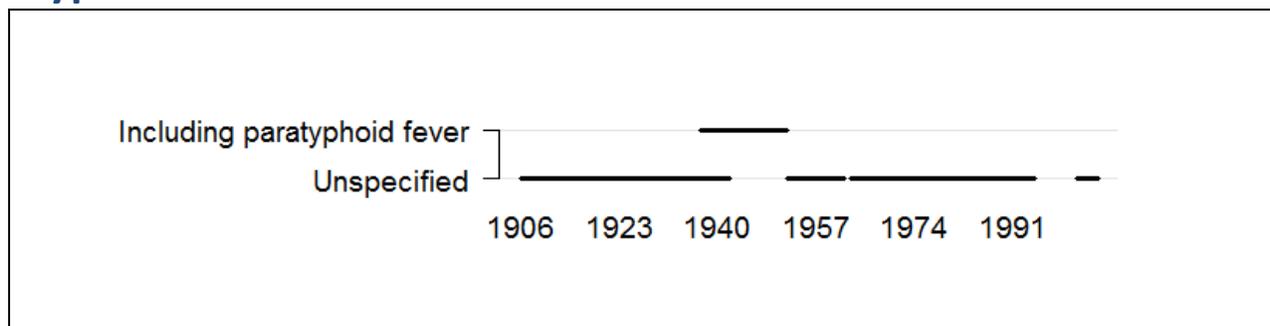


Figure D3, Total number of cases reported for Tuberculosis by each city for the entire time period

## Typhoid Fever



**Figure D1**, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

**Table D1**, Summary information for Typhoid Fever (Unspecified and Including paratyphoid fever)

Indicator	Connecticut	Hartford	Bridgeport	New Haven
Report period	1927-2005	1906-1953	1909-1952	1915-1952
Total weeks	2,727	1,568	1,446	1,509
Total cases	1,991	1,391	461	978
Max. cases per year	222	133	32	136
Year (max)	1953	1914	1912	1915
Max. cases per week	117	30	8	15
Week (max)	1953, wk 02	1910, wk 36	1917, wk 34	1915, wk 35
Average cases per year				
before 1950	60	32	12	29
95%CI	(45-75)	(21-43)	(9-15)	(17-41)
after 1950	12	1	1	4
95%CI	(2-22)	-	-	-
Average cases per week				
before 1950	1	1	0	1
95%CI	(1-1)	(1-1)	(0-0)	(1-1)
after 1950	0	1	1	1
95%CI	(0-0)	-	-	(0-2)

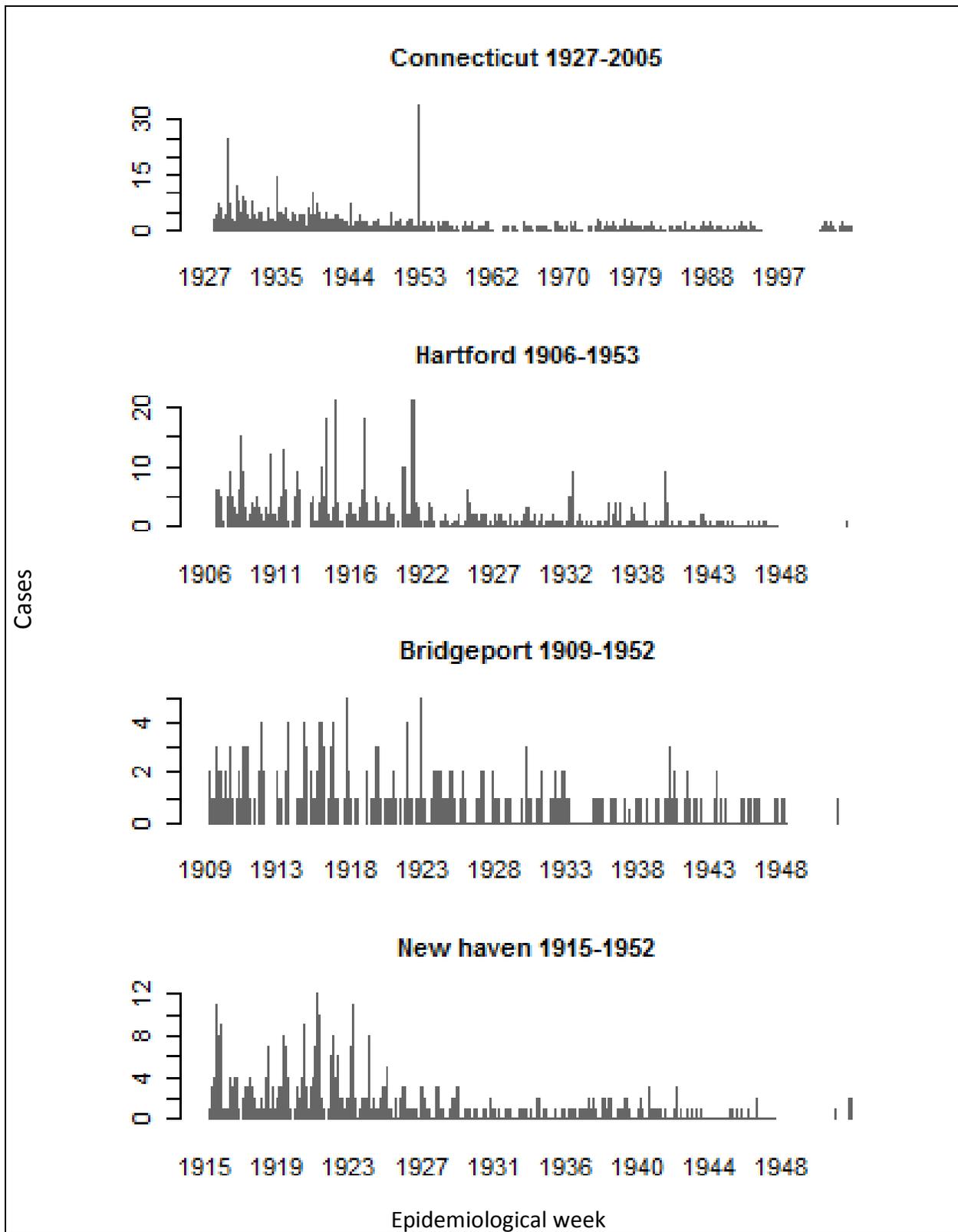


Figure D2, Number of cases reported for Typhoid Fever per epidemiological week

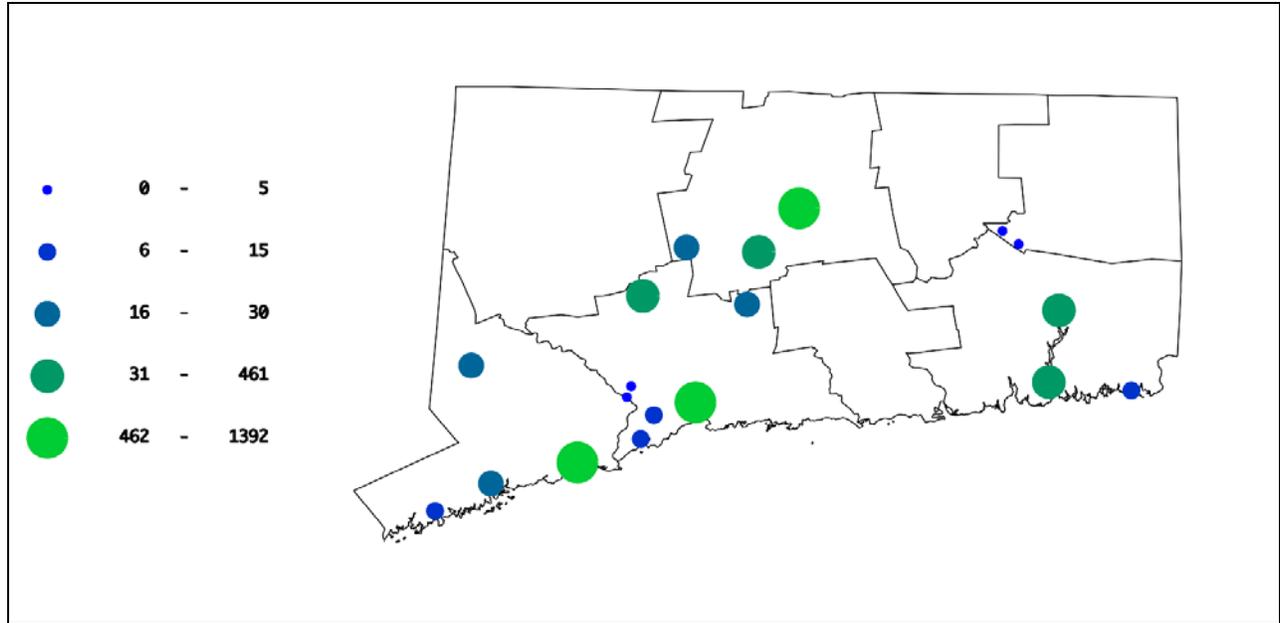
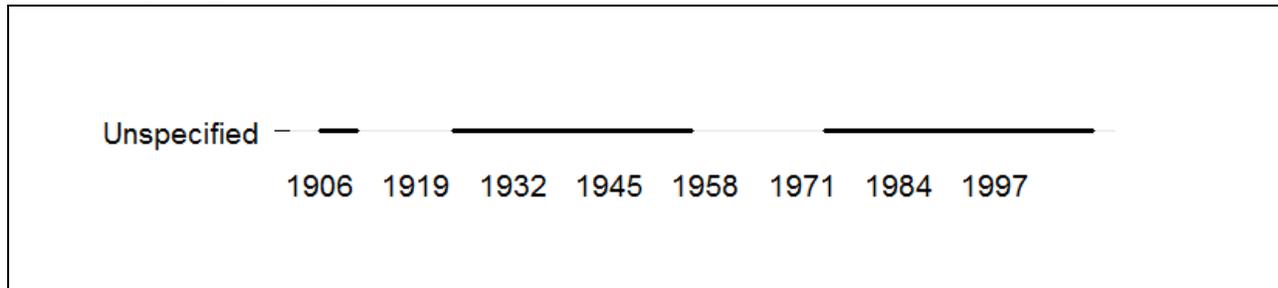


Figure D3, Total number of all cases reported for Typhoid Fever by each city for the entire time period

## Whooping Cough



**Figure D1,** Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

**Table D1,** Summary information for Whooping Cough

Indicator	Connecticut	Bridgeport	New Haven
Report period	1937-2009	1909-1953	1924-1953
Total weeks	2,007	1,225	1,241
Total cases	39,637	1,799	9,451
Max. cases per year	4,059	185	1,080
Year (max)	1942	1932	1936
Max. cases per week	163	15	73
Week (max)	1950, wk 10	1932, wk 28	1936, wk 35
Average cases per year			
before 1960	2,038	62	350
95%CI	(1,455-2,621)	(44-80)	(250-450)
after 1960	25	-	-
95%CI	(18-32)	-	-
Average cases per week			
before 1960	42	1	8
95%CI	(40-44)	(1-1)	(8-8)
after 1960	1	-	-
95%CI	(1-1)	-	-

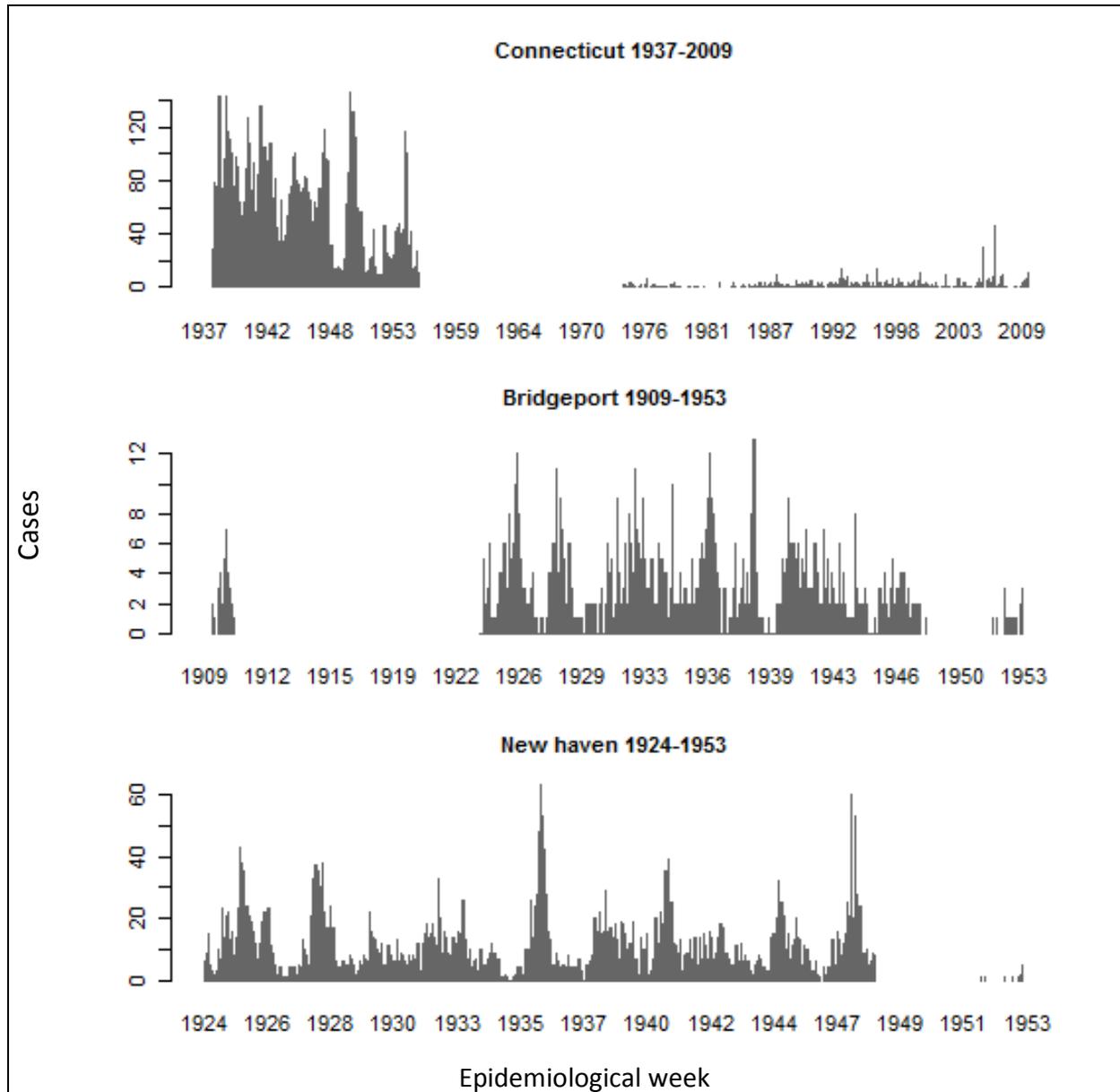


Figure D2, Number of cases reported for Whooping Cough per epidemiological week

## Diseases with fragmented data over time

For a number of diseases, only fragmented data was available for Connecticut. This was due to the inclusion of these diseases in the weekly surveillance system for short periods or widely dispersed periods of time. Due to these wide gaps or limited data, data for these diseases were not presented in separate sections but in the table below. Work will continue to include more data in the Tycho database and complete data sets for each disease where possible. Collaborations with local, state and federal public health agencies will be made to explore data availability and opportunities to include these in the Tycho database.

**Table 2,** Summary information on the occurrence of diseases with fragmented data in Connecticut

<b>Disease</b>	<b>Report type</b>	<b>Report period</b>	<b>Number of reports</b>	<b>Total cases</b>
Anthrax	State	1942-1945	143	0
Brucellosis [undulant fever]	City	1953-1953	1	1
Brucellosis [undulant fever]	State	1943-1982	293	632
Chickenpox [varicella]	City	1924-1981	1295	11792
Chickenpox [varicella]	State	1972-2007	514	58037
Dengue	City	1924-1925	16	0
Dysentery	State	1942-1948	249	260
Encephalitis	City	1941-1981	1001	26
Encephalitis	State	1942-1991	891	505
Legionellosis	State	1982-2009	614	433
Leprosy	State	1942-1990	305	2
Pellagra	City	1924-1932	292	7
Pneumonia	City	1917-1923	854	2797
Pneumonia	State	1920-1951	171	6042
Psittacosis	State	1956-1961	24	25
Rocky mountain spotted fever	State	1942-2004	722	150
Streptococcal disease, invasive group a	State	2003-2009	145	270
Streptococcus pneumoniae invasive disease	State	2007-2009	18	11
Tetanus	State	1962-1981	332	18
Toxic shock syndrome	State	1983-1990	100	5
Trichiniasis	City	1952-1953	7	7
Trichiniasis	State	1952-1955	28	33
Tularemia	State	1942-1988	304	38
Typhus fever	City	1924-1925	38	1
Typhus fever	State	1942-1958	201	10
West Nile disease	State	2003-2008	80	11

## Project Tycho

This report provides preliminary data for the state of Connecticut available in the Tycho database. This database is currently being beta tested and these data cannot be used for publication or other official use at this time. An open access release to the general public is planned for later in 2011.

Please visit the Tycho website for more information and to query the database at: [www.tycho.pitt.edu](http://www.tycho.pitt.edu). For further information regarding the Tycho project, contact Dr. Wilbert van Panhuis at the University of Pittsburgh Graduate School of Public Health.

Dr. Wilbert van Panhuis, MD PdD  
Graduate School of Public Health  
University of Pittsburgh  
130 DeSoto Street  
704 Parran Hall  
Pittsburgh, 15261 PA  
Tel: 412-624-7693  
Email: [wav10@pitt.edu](mailto:wav10@pitt.edu)