Project Tycho

Preliminary data for the state of Washington

Tycho database beta test version

June 2011

University of Pittsburgh Graduate School of Public Health

University of Pittsburgh   VMI
                                  Vaccine Modeling Initiative

Table of Content

Disease specific data for Washington, continued

Acknowledgements

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). 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
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

Tycho data report for the state of Washington

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) 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.

Tycho data report for the state of Washington

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.

Tycho data report for the state of Washington

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 Washington. This includes all data that has ever been published at state or city level for Washington 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). 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 [1]. For 1995 to 2009, these reports could be retrieved from the MMWR digital archive on the CDC website [2]. Most weekly reports between 1952 and 1995 could be retrieved from the HathiTrust Digitial Library [3] [4], 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

---

[1] http://www.pubmedcentral.nih.gov/tocrender.fcgi?journal=333&action=archive
[2] http://www.cdc.gov/mmwr
[3] http://catalog.hathitrust.org/Record/003910026
[4] http://catalog.hathitrust.org/Record/003843660

Tycho data report for the state of Washington

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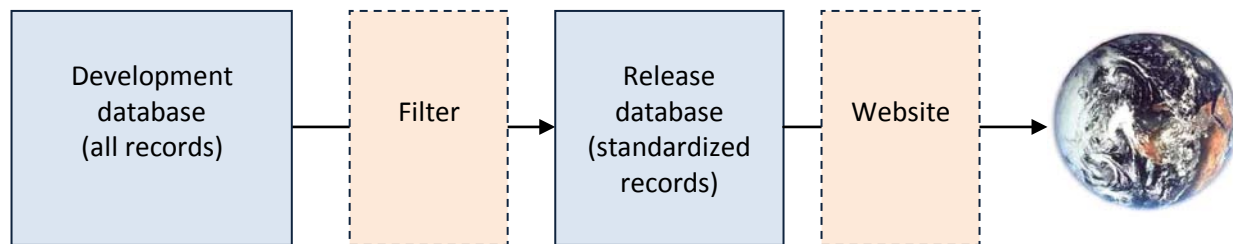
Figure 1, Schematic of the Tycho data architecture. Access to a consistent, standardized subset of the digitized US weekly surveillance reports will be provided for beta testing through a website that will allow querying and downloading of data. Data from the development database will be pushed through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

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 Washington

In this section, an overview of data available for the state of Washington 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 Washington. 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 Washington per each disease and subcategory.

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

Disease	City	State
Aids	-	932
Anthrax	-	143
Botulism	-	1
Brucellosis [undulant fever]	2	213
Chickenpox [varicella]	451	923
Chlamydia	-	675
Cryptosporidiosis	-	238
Dengue	4	4
Diphtheria	2026	2117
Dysentery		
Amebic	-	231
Bacillary	-	207
Unspecified	-	210
Encephalitis		
Lethargic	298	230
Post infectious	-	314
Primary [infectious] including unspecified	342	1091
Escherichia coli		
EHEC 0157	-	178
EHEC non-0157	-	32
O157:H7 NETSS	-	272
O157:H7 PHLIS	-	267
STEC	-	185
Giardiasis	-	378

Tycho data report for the state of Washington

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

Disease	City	State
Gonorrhea		
Civilian	-	748
Unspecified	-	1046
Haemophilus influenzae		
Age <5 non-serotype B	-	73
Age <5 serotype B	-	75
Age <5 unknown serotype	-	111
All ages all serotypes	-	520
Hepatitis		
Acute type A	-	376
Acute type B	-	370
Acute type C	-	105
Acute type NA NB [including C]	-	40
All types, <20 years	-	197
All types, >=20 years	-	100
All types, all ages	45	701
Type A [infectious]	-	1696
Type B [serum]	-	1407
Type NA NB [including C]	-	860
Type unspecified	-	826
Influenza	525	1104
Legionellosis	-	908
Leprosy	1	502
Listeriosis	-	227
Lyme disease	-	511
Malaria		
Military	-	22
Unspecified	-	1776
Measles		
Imported	-	447
Indigenous	-	507
Unspecified	2094	3292
Meningitis		
Aseptic	-	901
Meningococcus	721	1236
Unspecified	168	104

Tycho data report for the state of Washington

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

Disease	City	State
Meningococcal disease		
All serogroups	-	44
Invasive all serogroups	-	193
Invasive serogroup unknown	-	46
Serogroup A C Y and W-135	-	37
Serogroup B	-	45
Serogroup unspecified	35	2317
Mumps	450	2025
Pellagra	299	230
Pneumonia		
Lobar	3	-
Unspecified	57	115
Poliomyelitis		
Non paralytic	-	88
Paralytic	-	381
Total	948	1856
Psittacosis	-	21
Rabies in animals	-	489
Rocky mountain spotted fever	1	483
Rubella	-	1223
Salmonellosis		
NETSS	-	124
PHLIS	-	121
Unspecified	-	379
Scarlet fever		
Including streptococcal sore throat	91	554
Unspecified	2019	1654
Shigellosis		
NETSS	-	123
PHLIS	-	120
Unspecified	-	372
Streptococcal disease, invasive group a	-	117
Streptococcal sore throat	-	99
Syphilis		
Civilian primary and secondary	-	647
Congenital	-	25
Primary and secondary	-	1033
Tetanus	-	243
Toxic shock syndrome	-	252
Trichiniasis	1	6

Tycho data report for the state of Washington

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

Disease	City	State
Tuberculosis [phthisis pulmonalis]		
New active	-	150
Unspecified	543	1620
Tularemia	1	706
Typhoid fever [enteric fever]		
Including paratyphoid fever	337	590
Unspecified	1525	2589
Typhus fever		
Endemic	-	21
Unspecified	21	194
West nile disease		
Neuroinvasive	-	28
Non-neuroinvasive	-	43
Whooping cough [pertussis]	1331	3093

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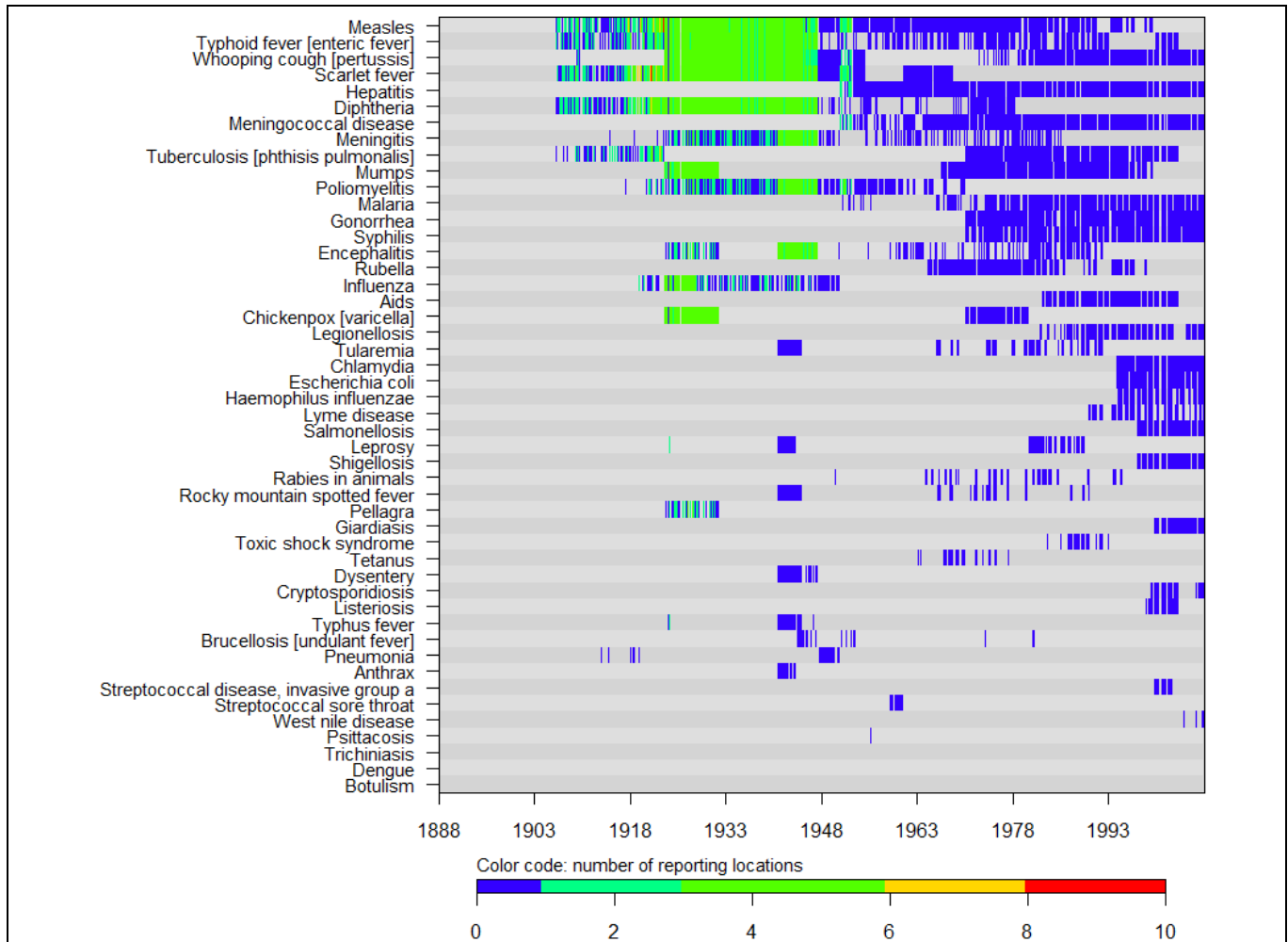
Figure 1, The number of locations (in color code) reporting on each disease per week between 1888 and 2009

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)

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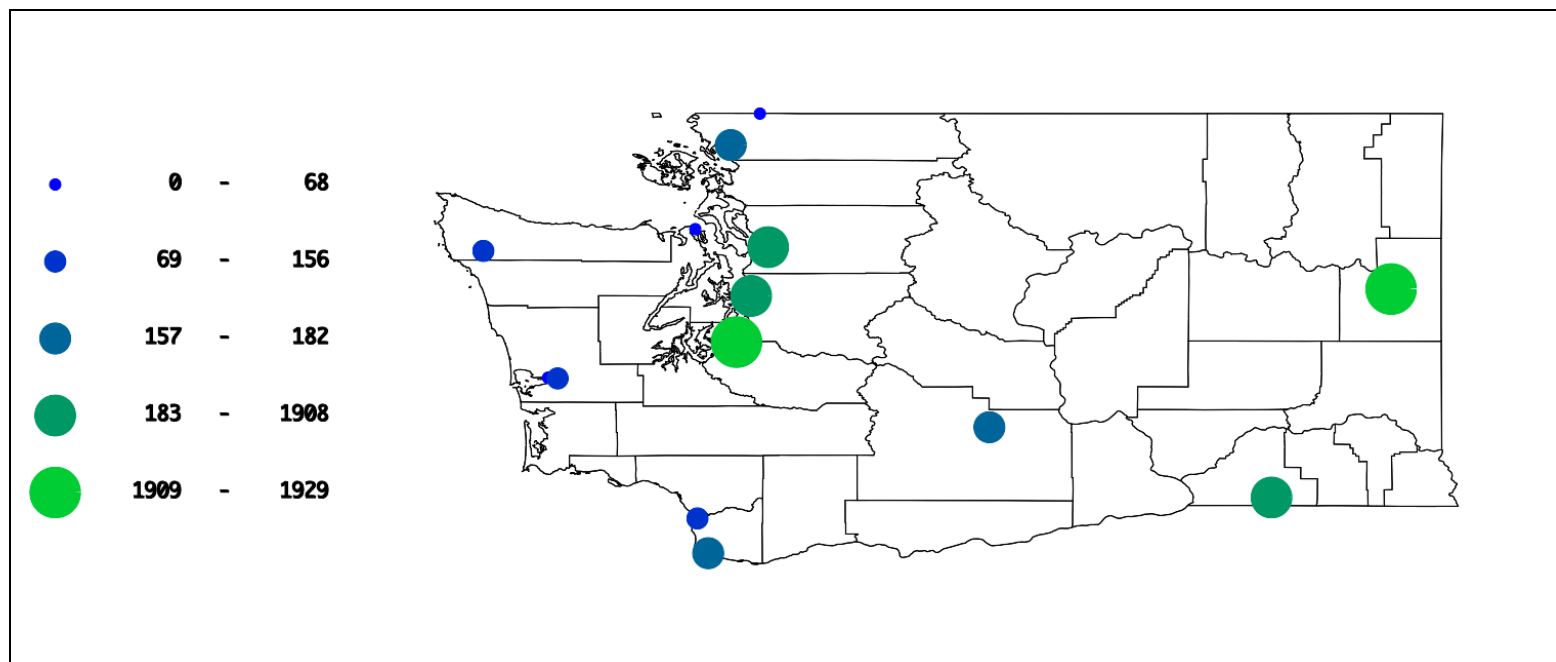
Figure 2, This map displays the number of weekly reports available for any disease per city (note that city reports were only available until 1953)

Figure 3 below provides an overview of the number of diseases for which a weekly report was included at the state state level and for individual cities. All cities for which at least 100 weekly reports were 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.
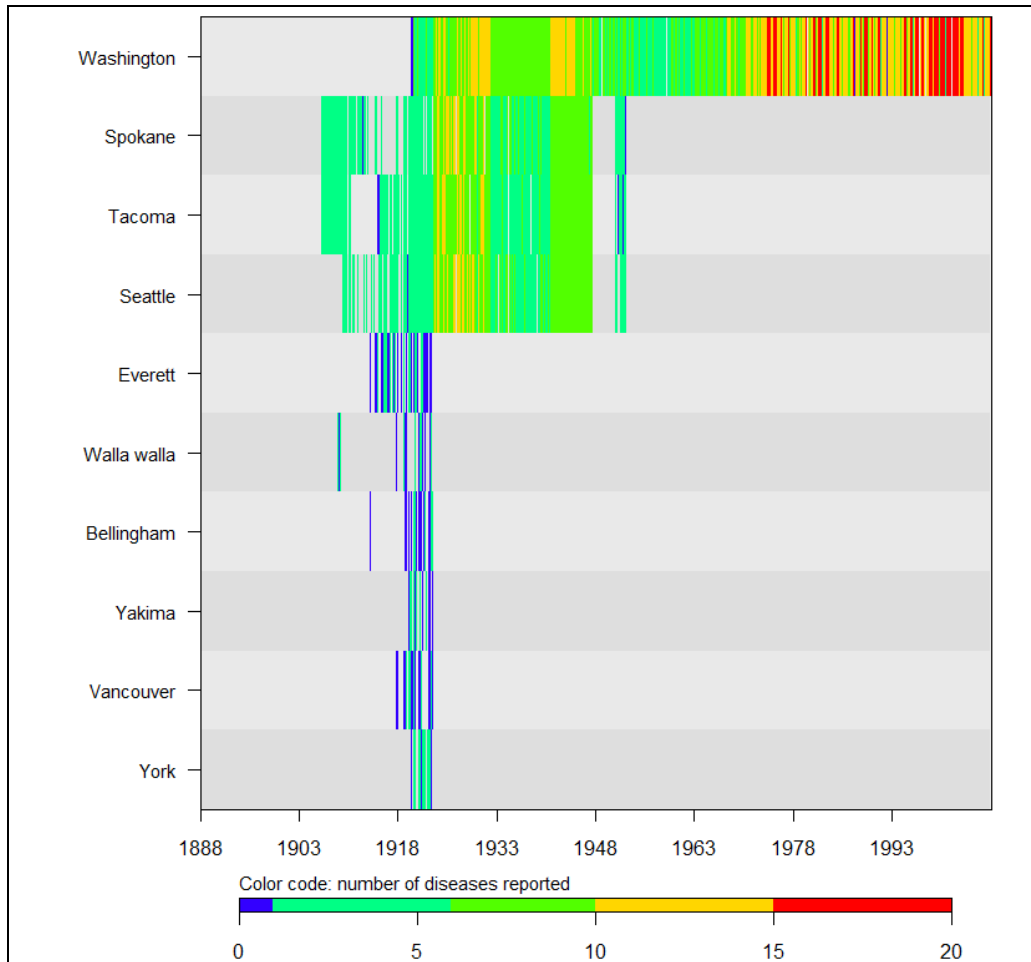
Figure 3, The number of diseases (in color code) that were reported for the state and city level (state on top row) for each week between 1888 and 2009
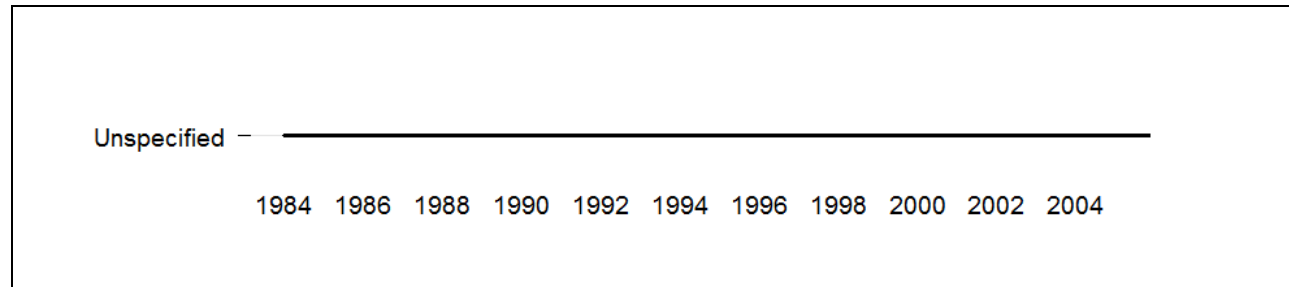
Tycho data report for the state of Washington

Disease specific data for Washington
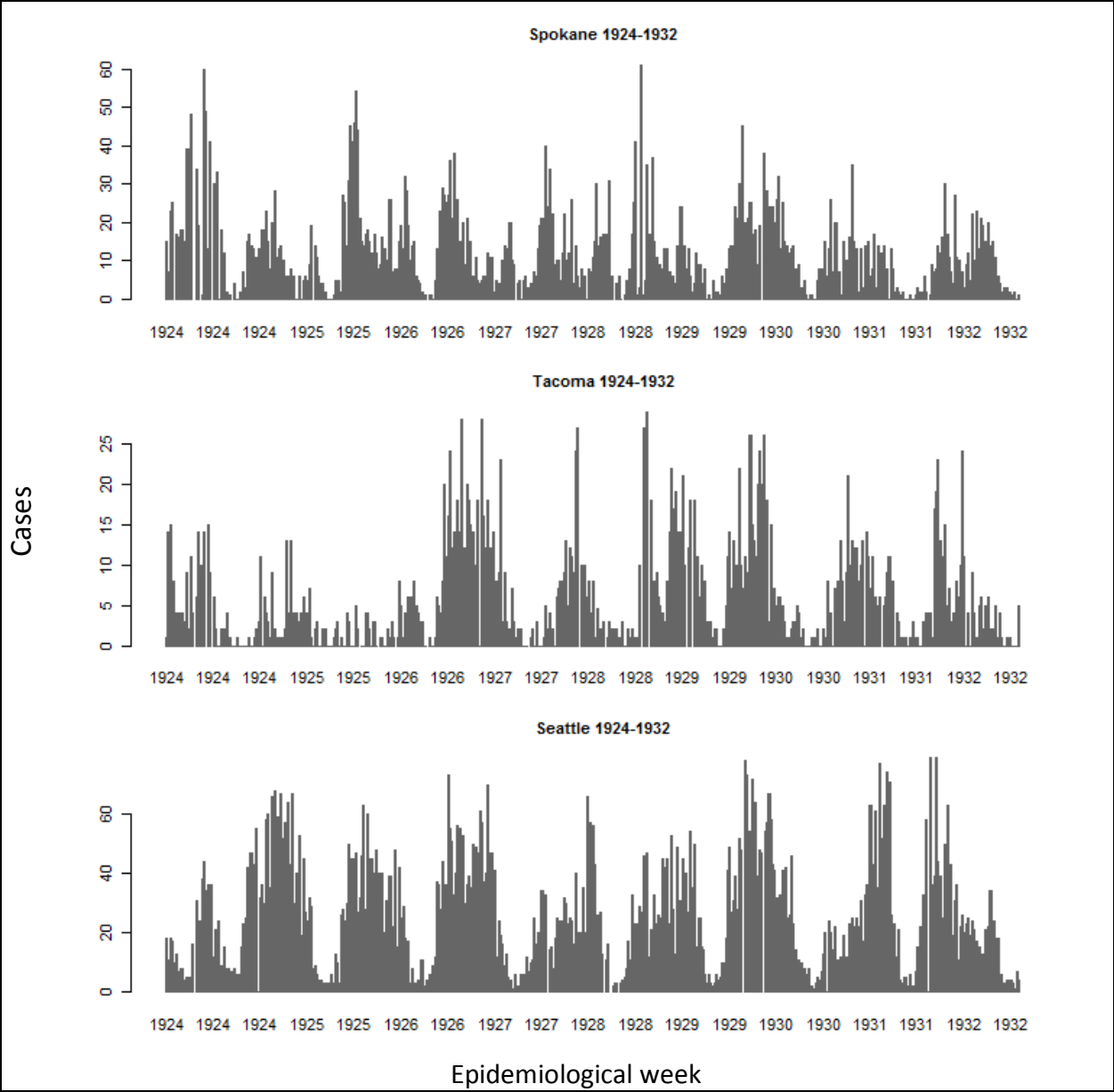
AIDS

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.
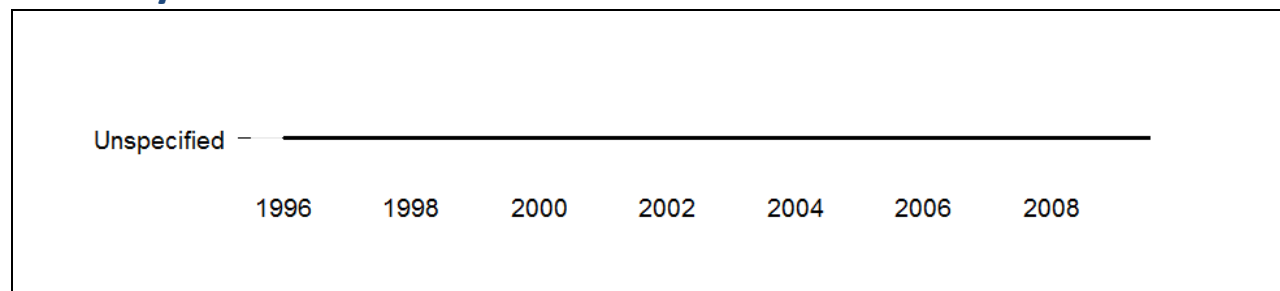
Table D1, Summary information for AIDS

Indicator	Washington
Report period	1984-2005
Total weeks	932
Total cases	10,188
Max. cases per year	1,477
Year (max)	1993
Max. cases per week	550
Week (max)	1993, wk 22
Average cases per year	463
95%CI	(326-600)
Average cases per week	11
95%CI	(9-13)
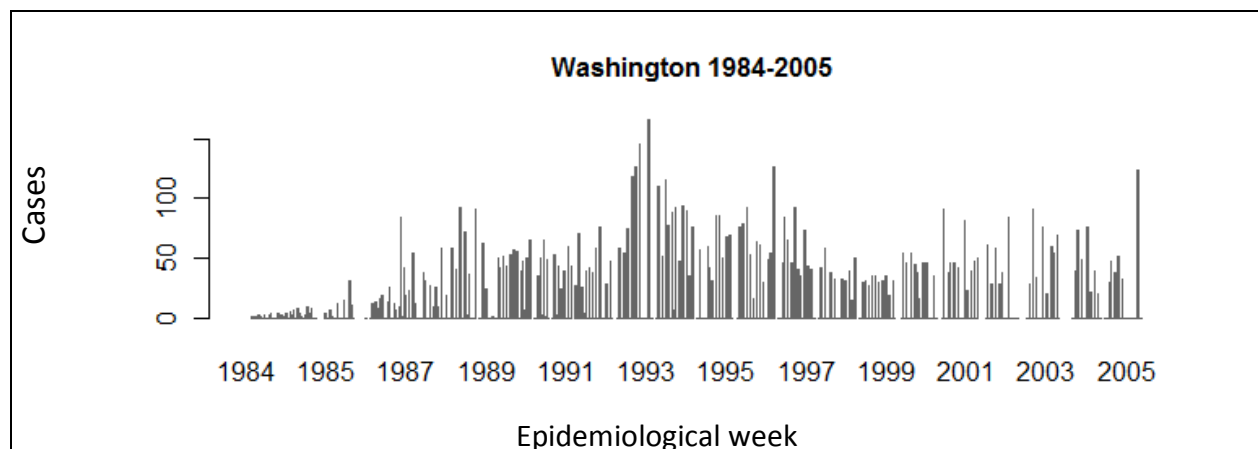
Figure D2, Number of cases reported for AIDS per epidemiological week

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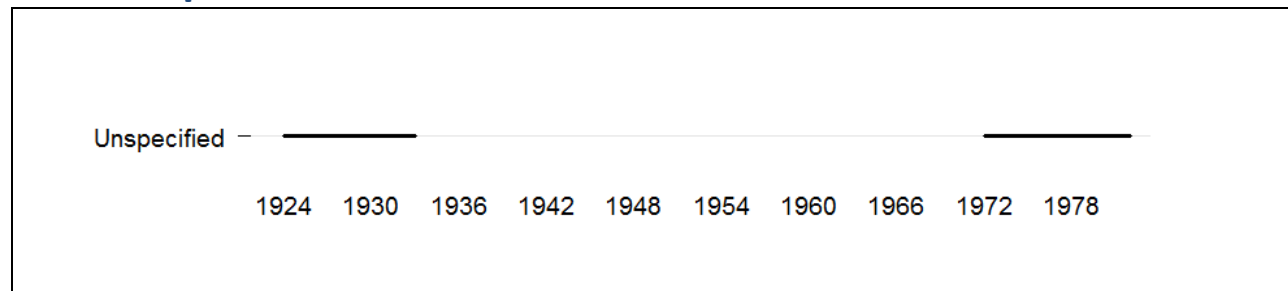
Figure D1, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

Table D1, Summary information for Chickenpox

Indicator	Spokane	Tacoma	Seattle
Report period	1924-1932	1924-1932	1924-1932
Total weeks	430	431	444
Total cases	5,354	2,610	11,730
Max. cases per year	789	487	1,733
Year (max)	1928	1929	1925
Max. cases per week	84	34	91
Week (max)	1928, wk 49	1929, wk 45	1925, wk 05
Average cases per year			
before 1995	595	290	1,303
95%CI	(470-720)	(201-379)	(1,056-1,550)
after 1995	-	-	-
95%CI	-	-	-
Average cases per week			
before 1995	12	6	26
95%CI	(11-13)	(5-7)	(24-28)
after 1995	-	-	-
95%CI	-	-	-

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

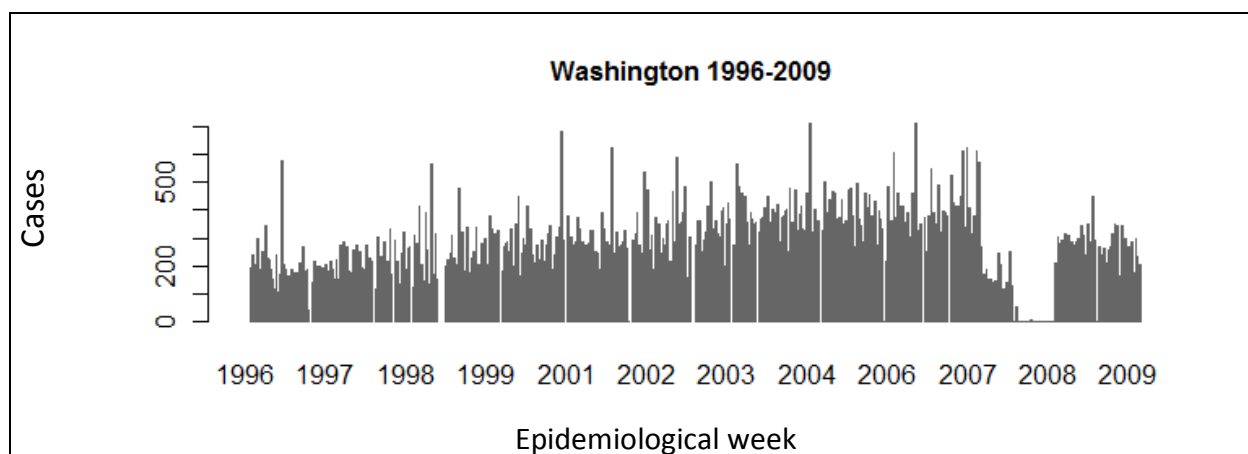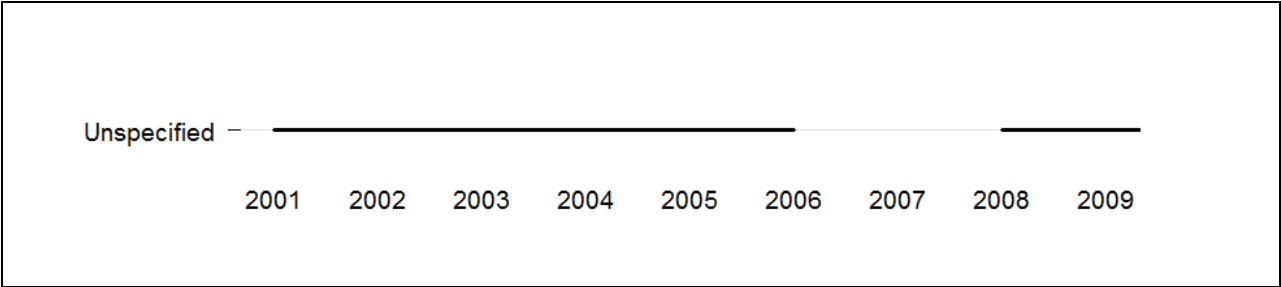*Tycho data report for the state of Washington*

Chlamydia
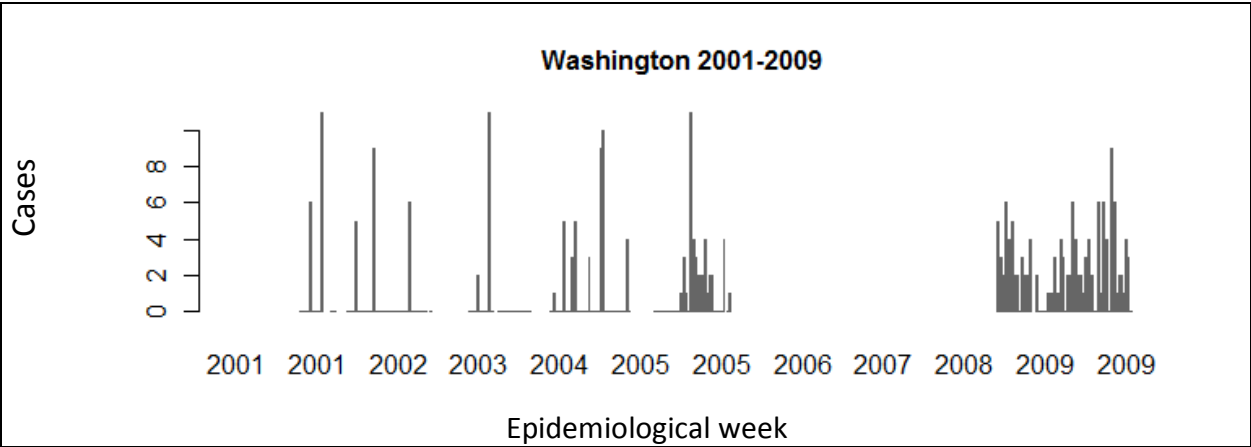

Figure D1, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

Table D1, Summary information for Chlamydia

Indicator	Washington
Report period	1996-2009
Total weeks	675
Total cases	178,395
Max. cases per year	21,472
Year (max)	2004
Max. cases per week	5,612
Week (max)	2004, wk 23
Average cases per year	12,742
95%CI	(10,222-15,262)
Average cases per week	264
95%CI	(243-285)

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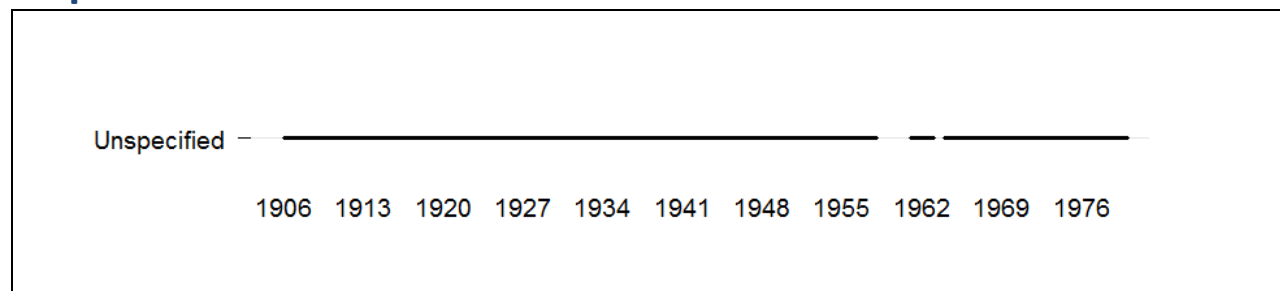
Cryptosporidiosis
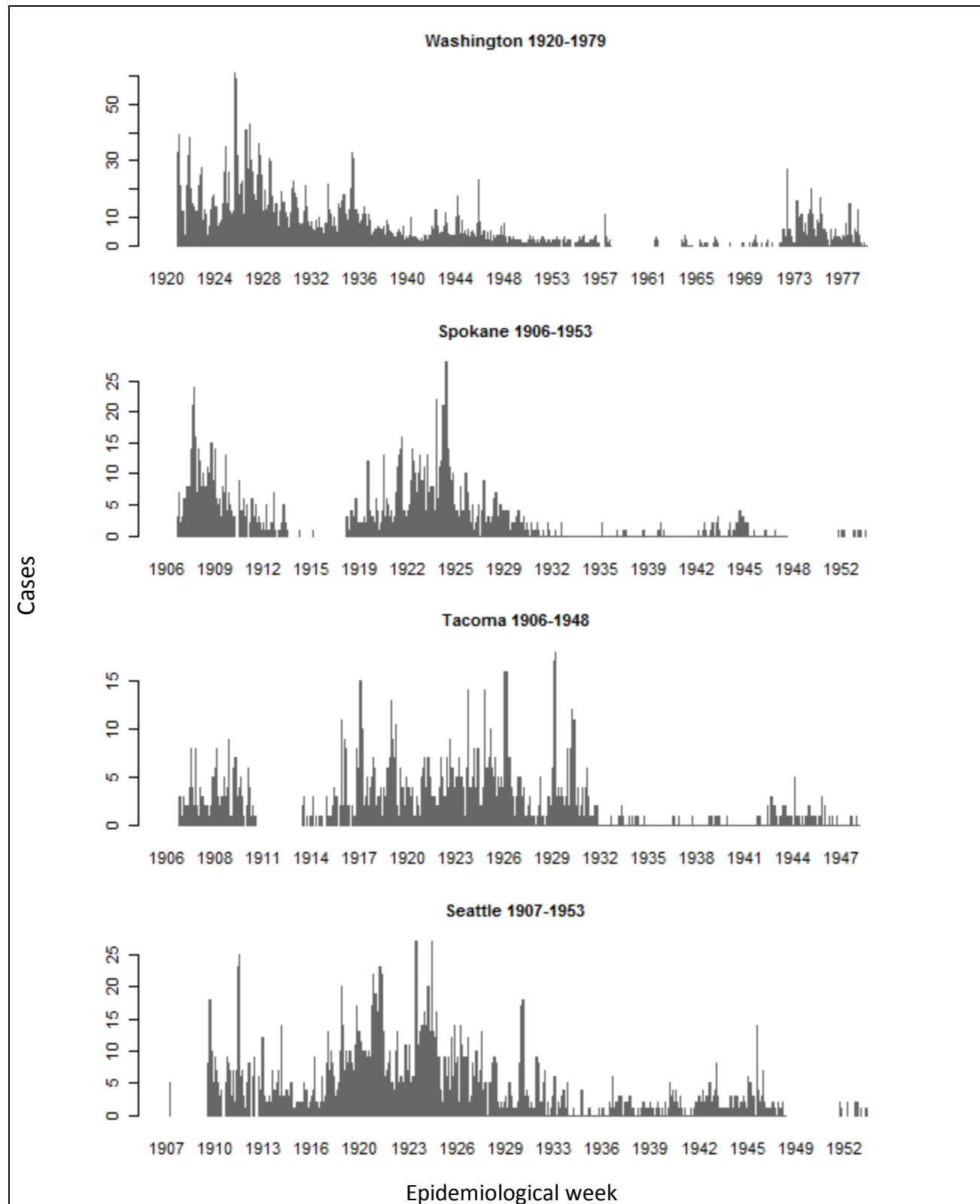

Figure D1, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.

Table D1, Summary information for Cryptosporidiosis

Indicator	Washington
Report period	2001-2009
Total weeks	238
Total cases	319
Max. cases per year	93
Year (max)	2009
Max. cases per week	18
Week (max)	2003, wk 40
Average cases per year	46
95%CI	(24-68)
Average cases per week	1
95%CI	(1-1)

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

Tycho data report for the state of Washington

Dengue

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.
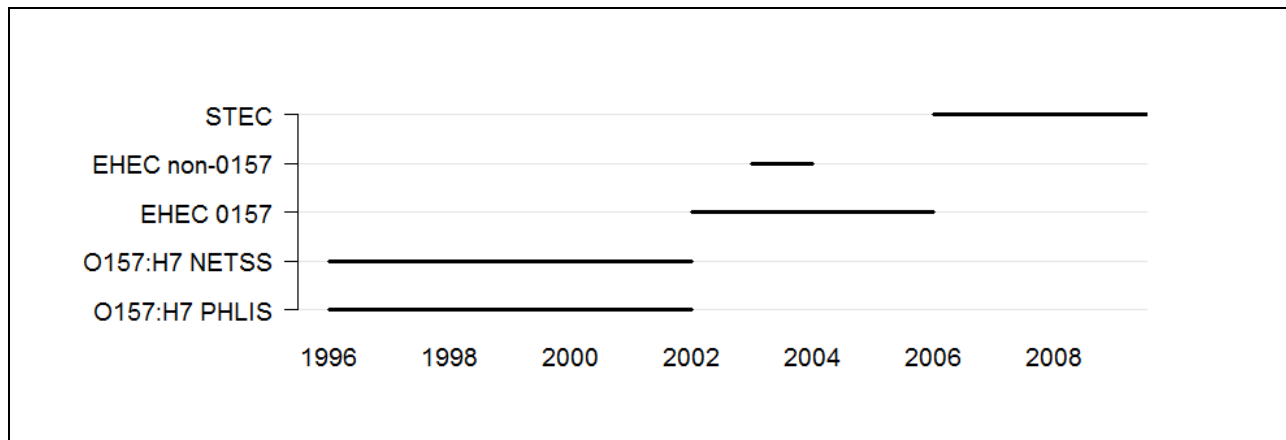
Table D1, Summary information for Dengue

Indicator	Seattle
Report period	1924-1924
Total weeks	3
Total cases	0

Tycho data report for the state of Washington

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	Washington	Spokane	Tacoma	Seattle
Report period	1920-1979	1906-1953	1906-1948	1907-1953
Total weeks	2,117	1,713	1,654	1,713
Total cases	11,801	3,084	2,418	4,513
Max. cases per year	921	361	245	433
Year (max)	1927	1908	1926	1924
Max. cases per week	325	58	32	32
Week (max)	1962, wk 43	1925, wk 13	1944, wk 09	1920, wk 48
Average cases per year				
before 1940	441	88	70	126
95%CI	(328-554)	(49-127)	(48-92)	(85-167)
after 1940	73	9	14	35
95%CI	(47-99)	(1-17)	(0-28)	(13-57)
Average cases per week				
before 1940	9	2	2	3
95%CI	(9-9)	(2-2)	(2-2)	(3-3)
after 1940	2	0	0	1
95%CI	(1-3)	(0-0)	(0-0)	(1-1)

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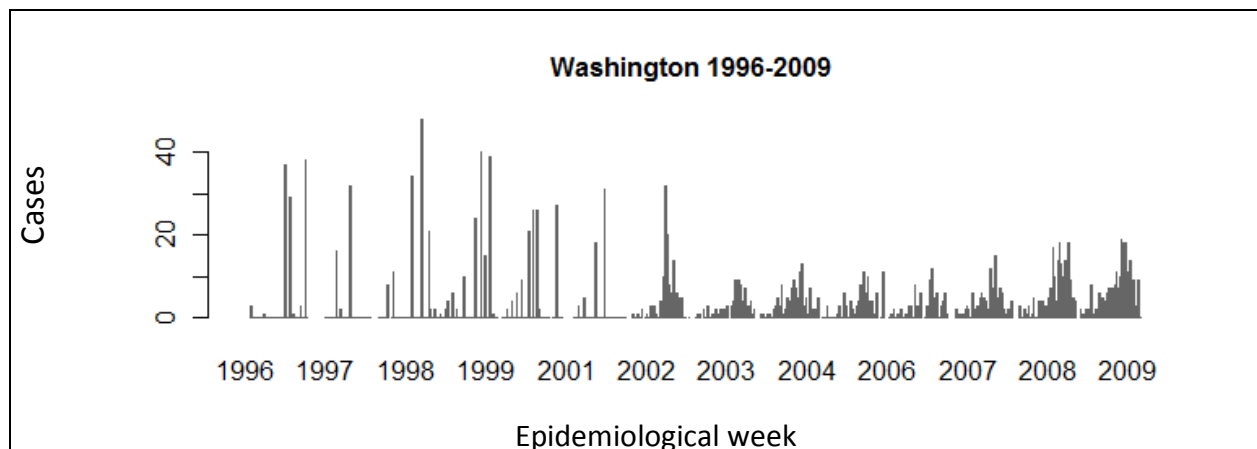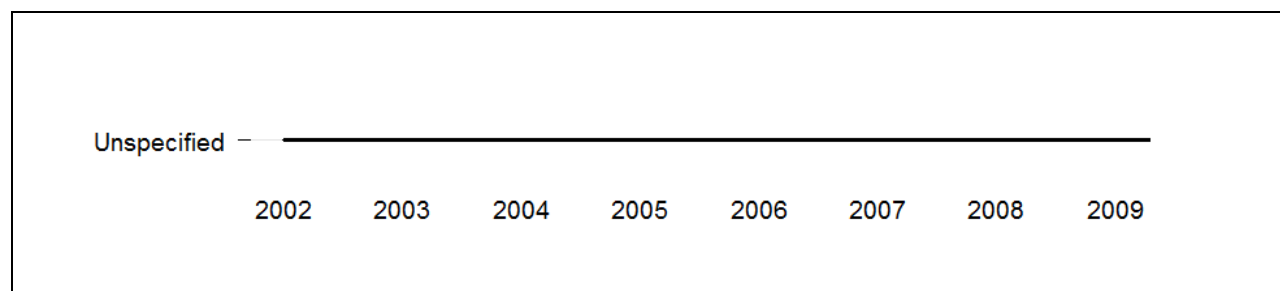
Escherichia Coli
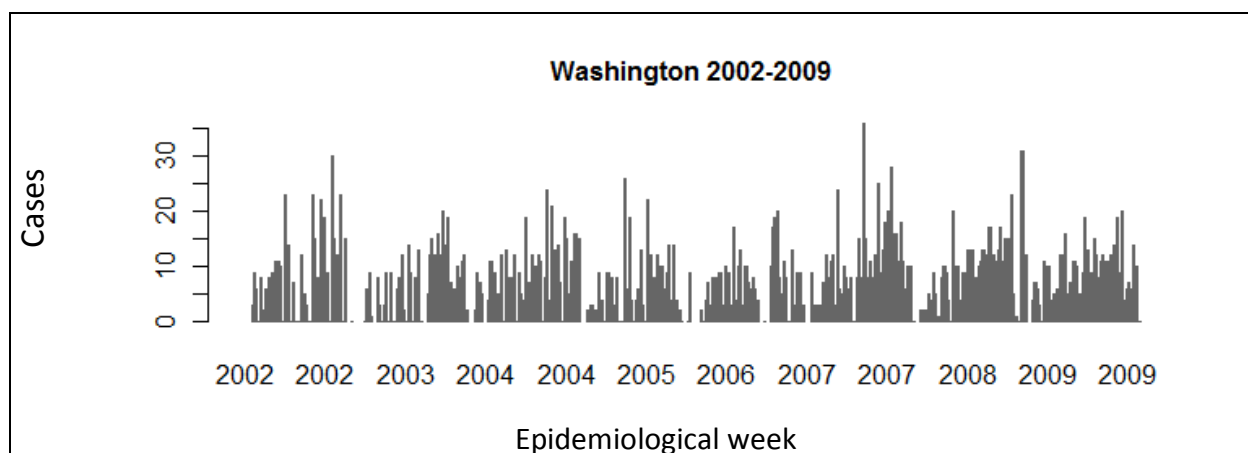

Figure D1, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.
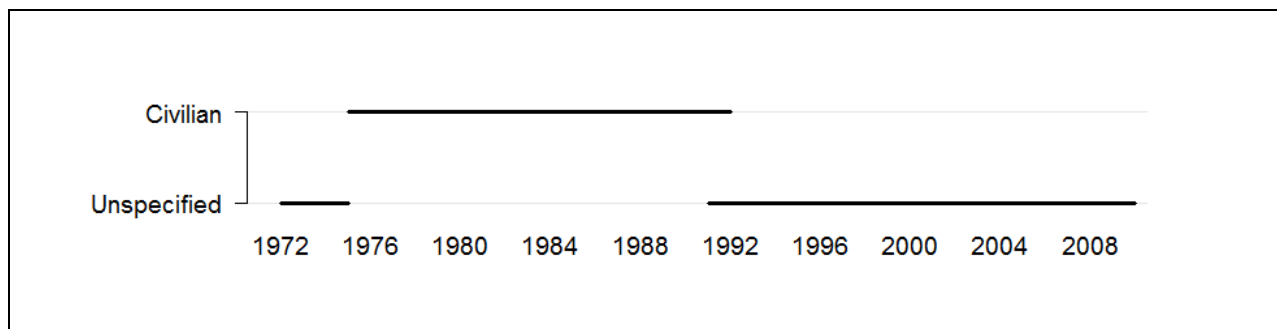
Table D1, Summary information for Escherichia Coli (0157:H7 PHLIS, EHEC 0157, and STEC)

Indicator	Washington
Report period	1996-2009
Total weeks	630
Total cases	2,076
Max. cases per year	262
Year (max)	2009
Max. cases per week	77
Week (max)	1997, wk 47
Average cases per year	148
95%CI	(116-180)
Average cases per week	3
95%CI	(2-4)

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

Giardiasis
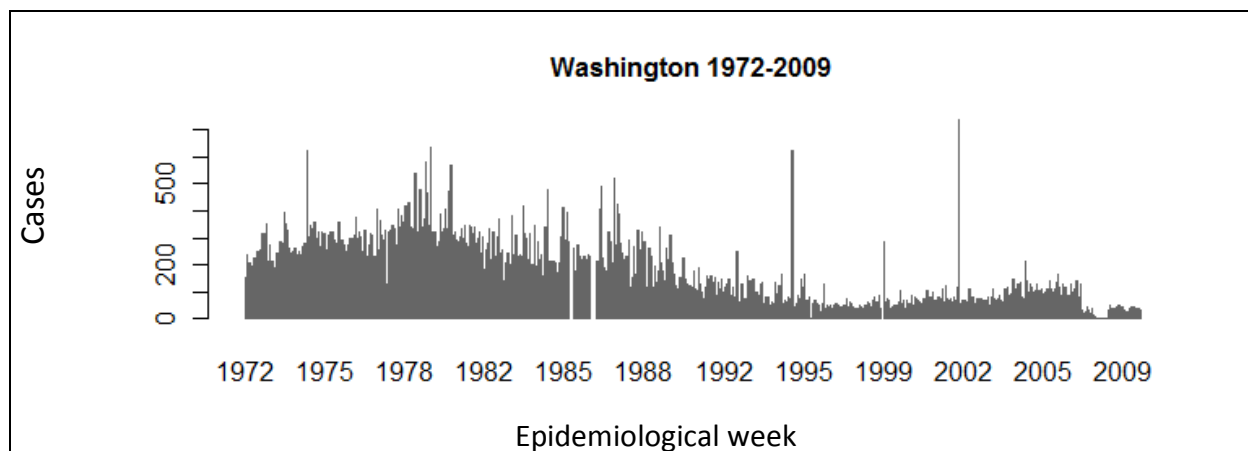

Figure D1, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.
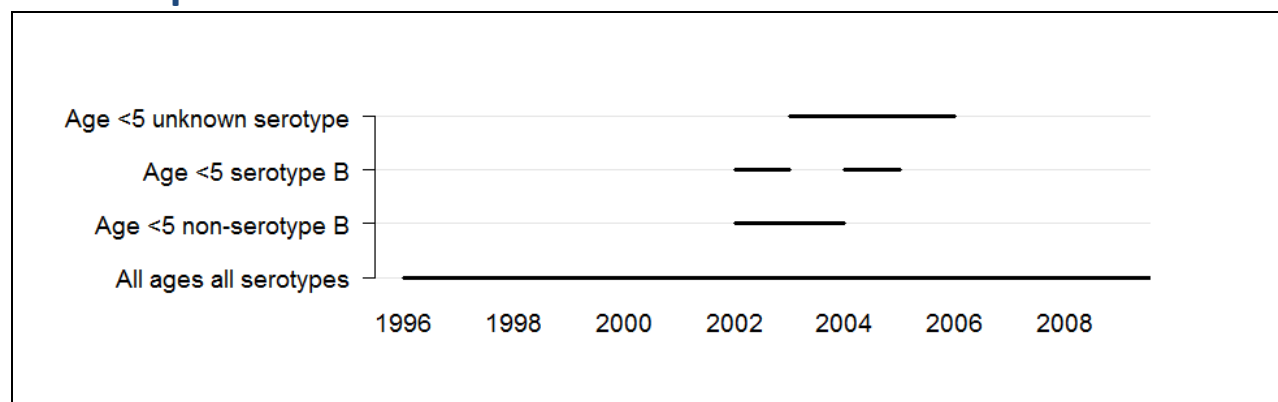
Table D1, Summary information for Giardiasis

Indicator	Washington
Report period	2002-2009
Total weeks	377
Total cases	3,134
Max. cases per year	534
Year (max)	2007
Max. cases per week	41
Week (max)	2007, wk 52
Average cases per year	392
95%CI	(332-452)
Average cases per week	8
95%CI	(7-9)

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

Tycho data report for the state of Washington

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	Washington
Report period	1972-2009
Total weeks	1,794
Total cases	238,734
Max. cases per year	14,223
Year (max)	1979
Max. cases per week	2,565
Week (max)	2001, wk 44
Average cases per year	6,282
95%CI	(4,902-7,662)
Average cases per week	133
95%CI	(127-139)

Figure D2, Number of cases reported for Gonorrhea per epidemiological week
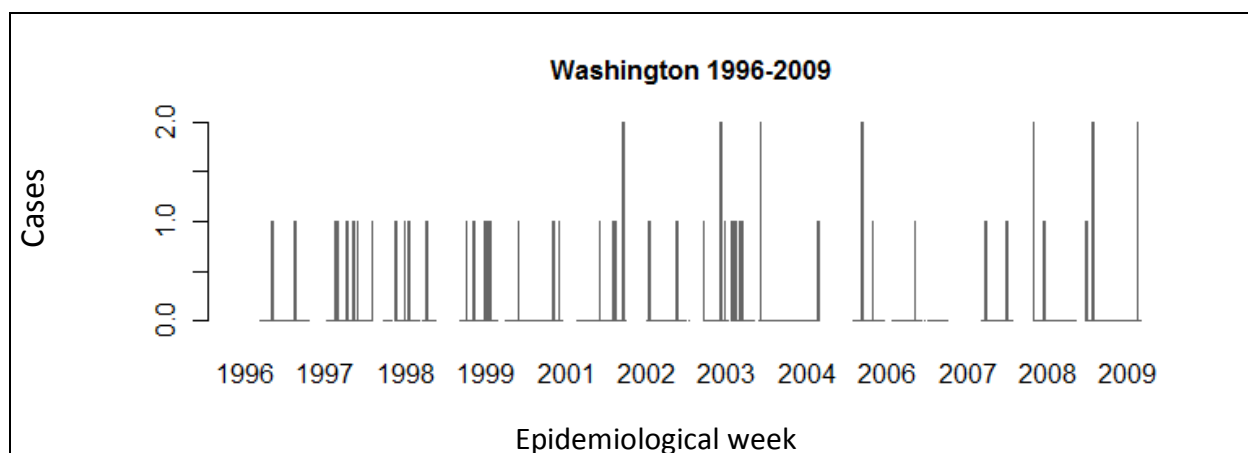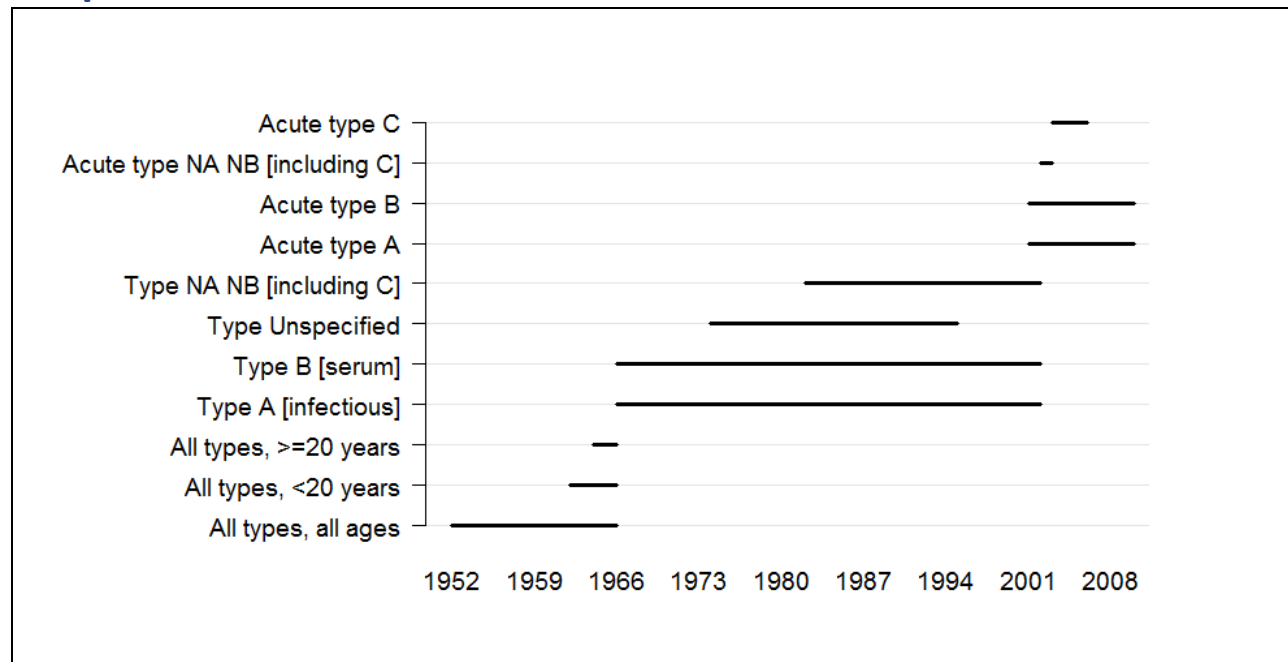
Tycho data report for the state of Washington

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	Washington
Report period	1996-2009
Total weeks	520
Total cases	63
Max. cases per year	9
Year (max)	2003
Max. cases per week	2
Week (max)	1998, wk 29
Average cases per year	4
95%CI	(3-5)
Average cases per week	0
95%CI	(0-0)

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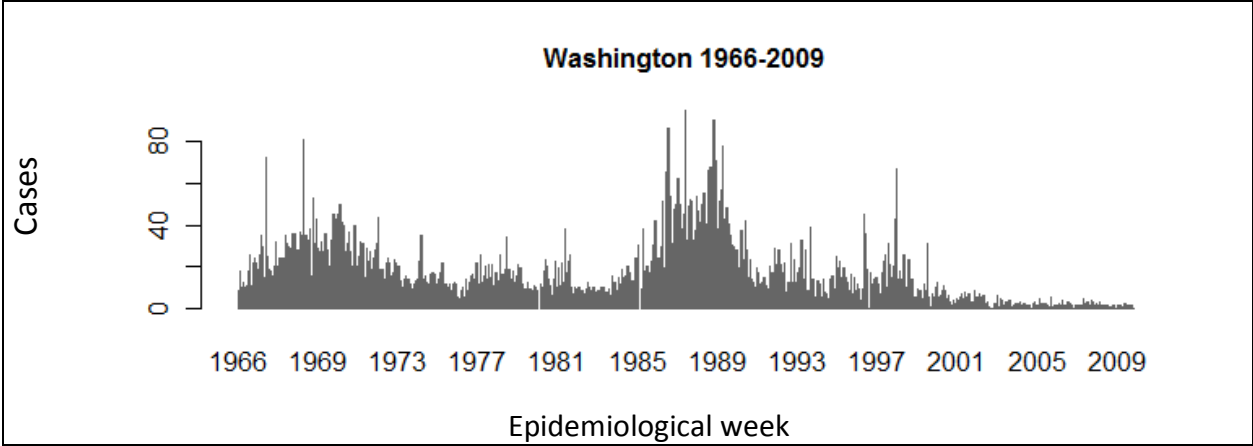
Hepatitis

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	Washington
Report period	1966-2009
Total weeks	2,082
Total cases	26,084
Max. cases per year	3,659
Year (max)	1970
Max. cases per week	2,440
Week (max)	1970, wk 44
Average cases per year	
before 1990	863
95%CI	(566-1,160)
after 1990	238
95%CI	(147-329)
Average cases per week	
before 1990	18
95%CI	(14-22)
after 1990	5
95%CI	(5-5)

Tycho data report for the state of Washington

Figure D2, Number of cases reported for Hepatitis per epidemiological week
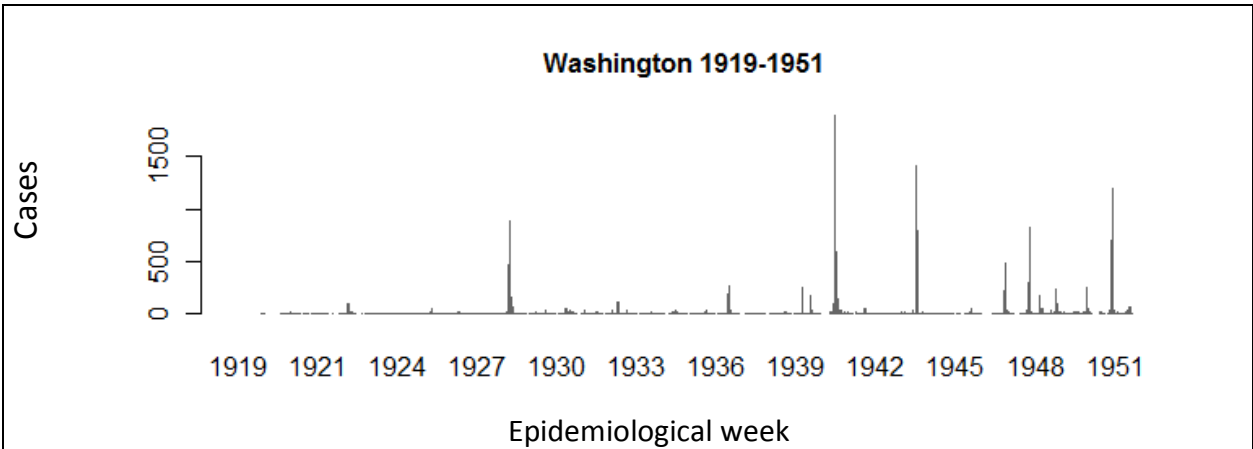
Influenza
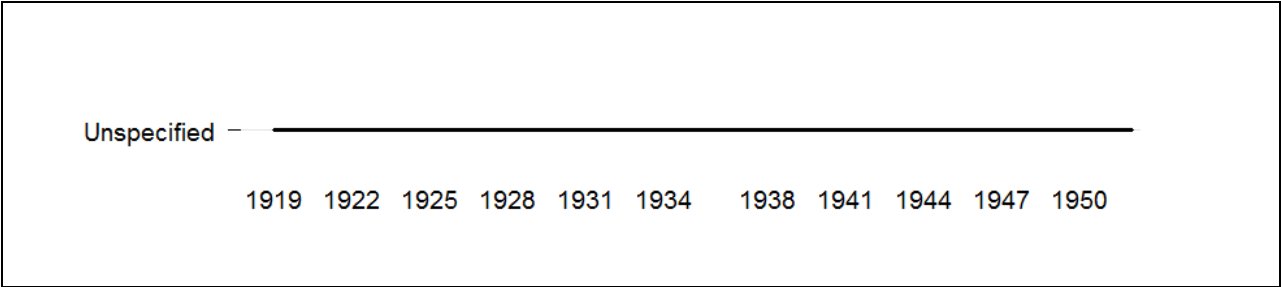
Figure D1, Weeks between 1888 and 2009 for which data on the disease and subcategories (if applicable) are available in the Tycho database. If no subcategory was reported, a subcategory of "Unspecified" was assigned.
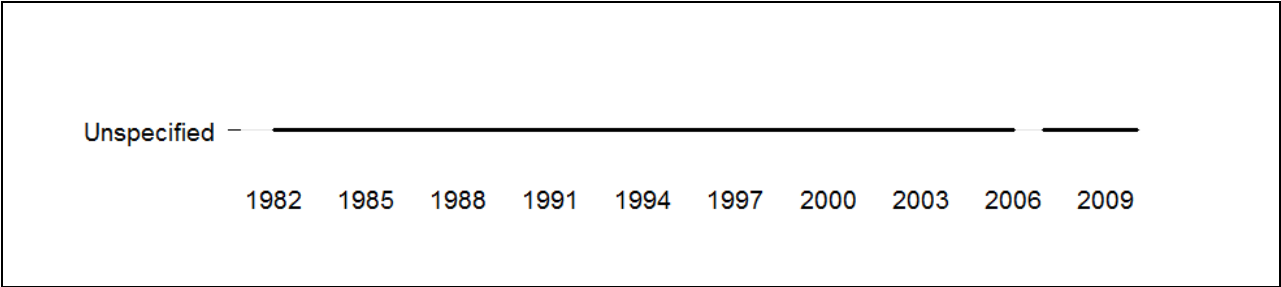
Table D1, Summary information for Influenza

Indicator	Washington
Report period	1919-1951
Total weeks	1,104
Total cases	32,768
Max. cases per year	11,019
Year (max)	1951
Max. cases per week	5,847
Week (max)	1951, wk 11
Average cases per year	993
95%CI	(268-1,718)
Average cases per week	30
95%CI	(17-43)

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

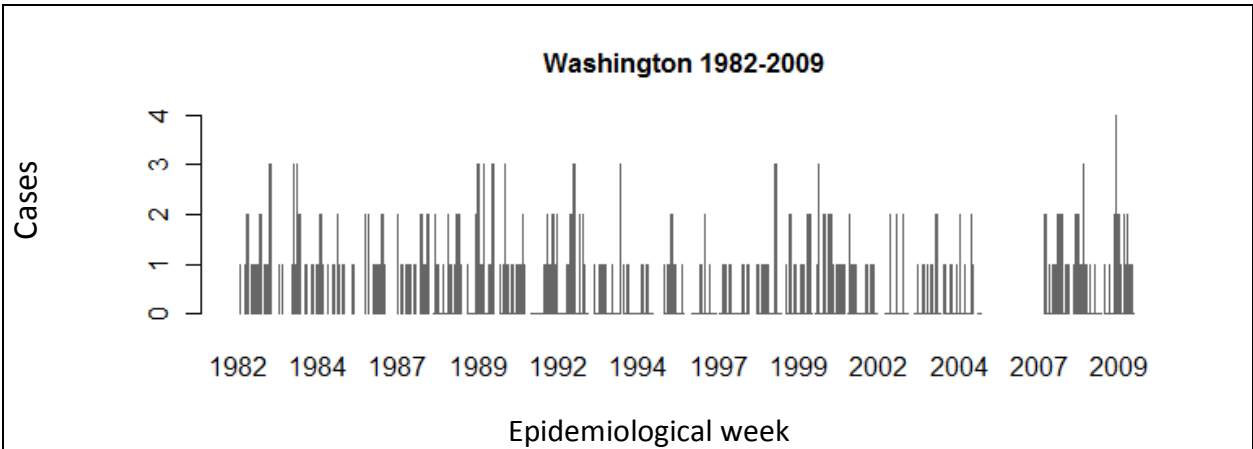*Tycho data report for the state of Washington*

Legionellosis

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.
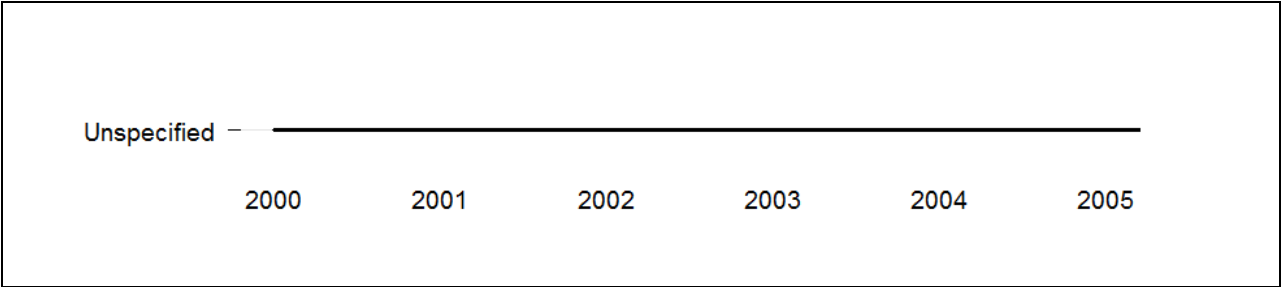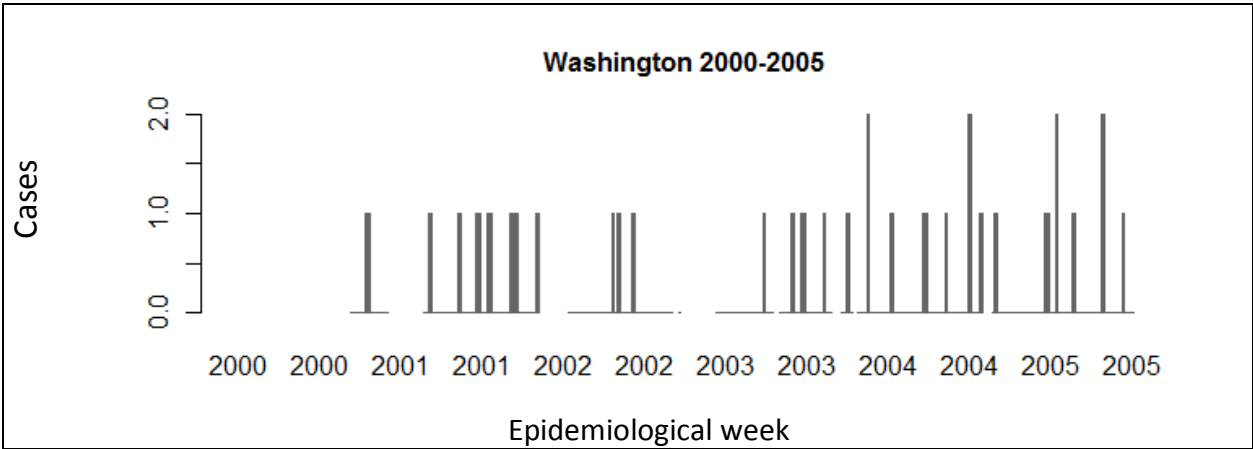
Table D1, Summary information for Legionellosis

Indicator	Washington
Report period	1982-2009
Total weeks	908
Total cases	363
Max. cases per year	27
Year (max)	2009
Max. cases per week	4
Week (max)	1987, wk 30
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 Legionellosis per epidemiological week

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.
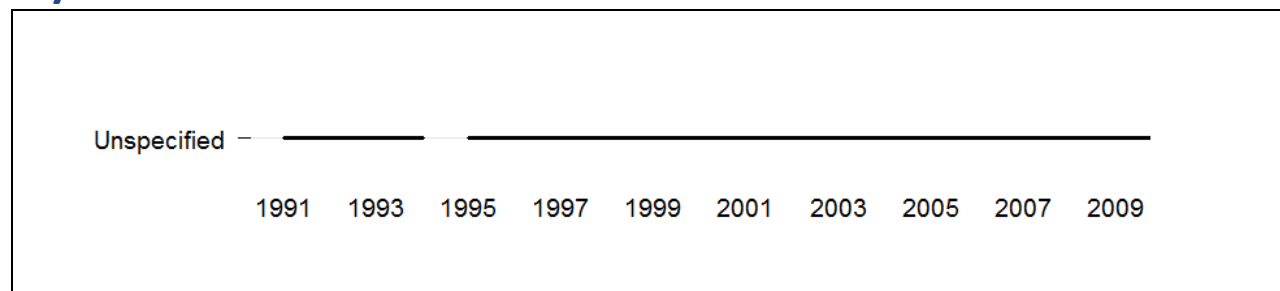
Table D1, Summary information for Listeriosis

Indicator	Washington
Report period	2000-2005
Total weeks	227
Total cases	45
Max. cases per year	10
Year (max)	2001
Max. cases per week	2
Week (max)	2002, wk 18
Average cases per year	8
95%CI	(5-11)
Average cases per week	0
95%CI	(0-0)

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

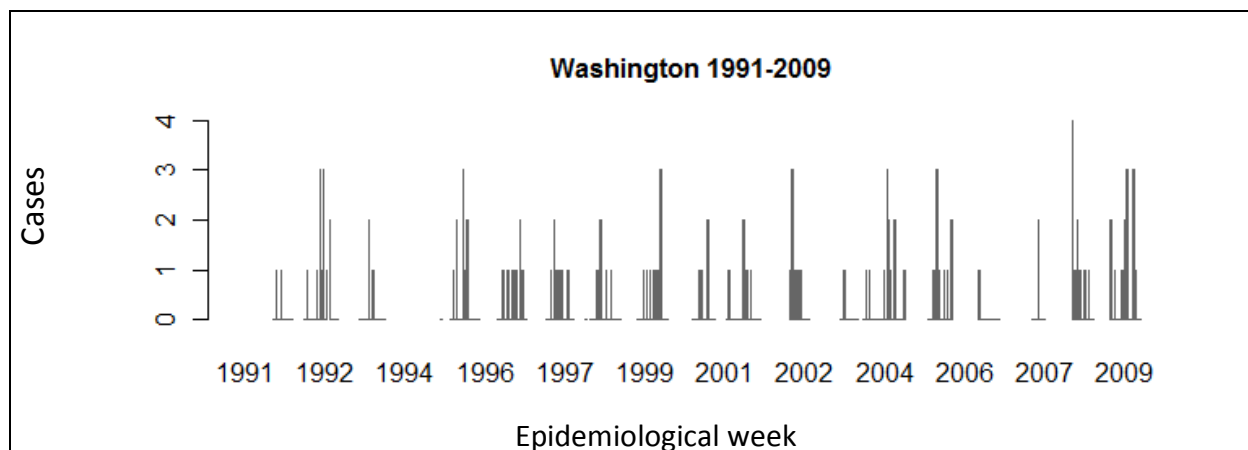*Tycho data report for the state of Washington*

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.
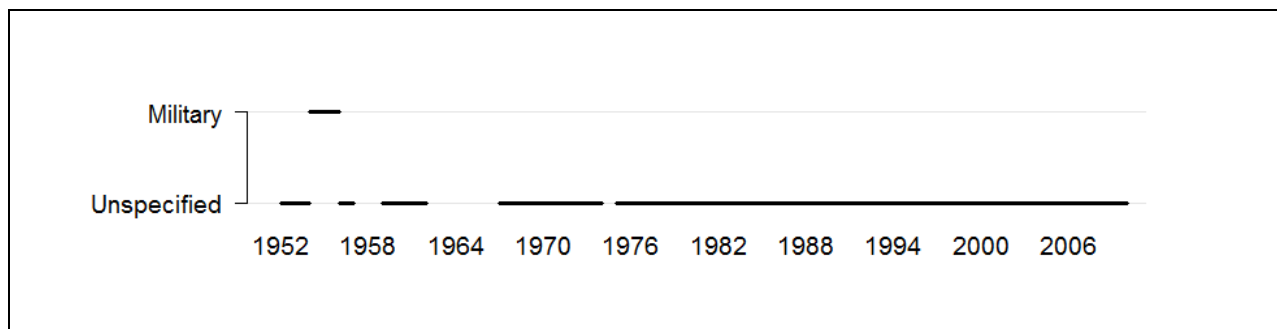
Table D1, Summary information for Lyme Disease

Indicator	Washington
Report period	1991-2009
Total weeks	511
Total cases	153
Max. cases per year	17
Year (max)	2009
Max. cases per week	7
Week (max)	2007, wk 50
Average cases per year	8
95%CI	(6-10)
Average cases per week	0
95%CI	(0-0)

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

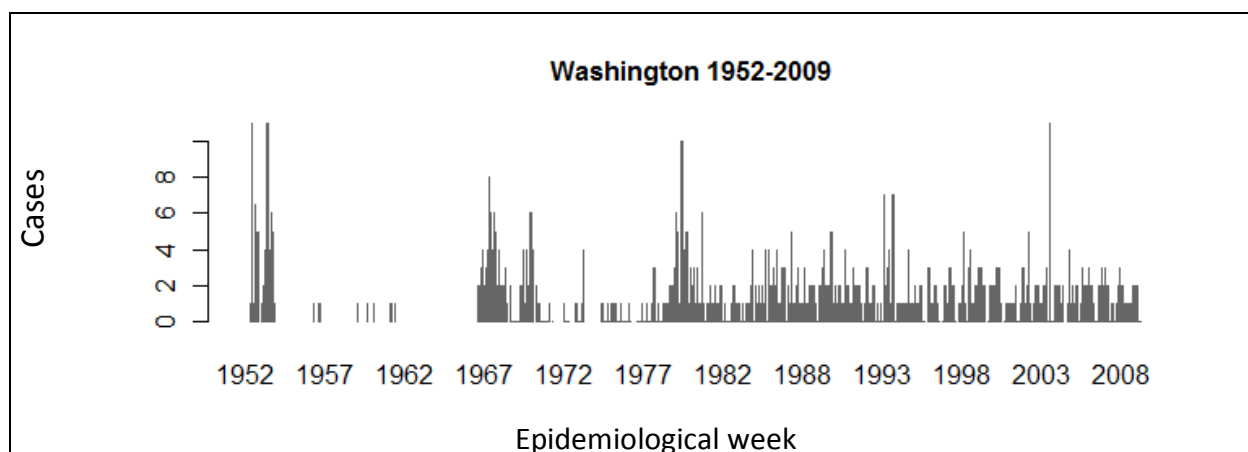*Tycho data report for the state of Washington*

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.
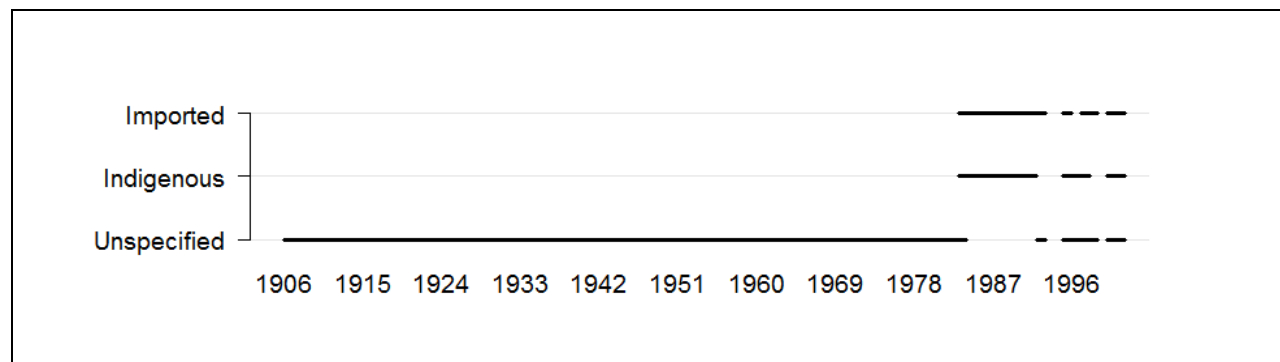
Table D1, Summary information for Malaria (Unspecified)

Indicator	Washington
Report period	1952-2009
Total weeks	1,776
Total cases	1,143
Max. cases per year	70
Year (max)	1970
Max. cases per week	25
Week (max)	1997, wk 45
Average cases per year	24
95%CI	(19-29)
Average cases per week	1
95%CI	(1-1)

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

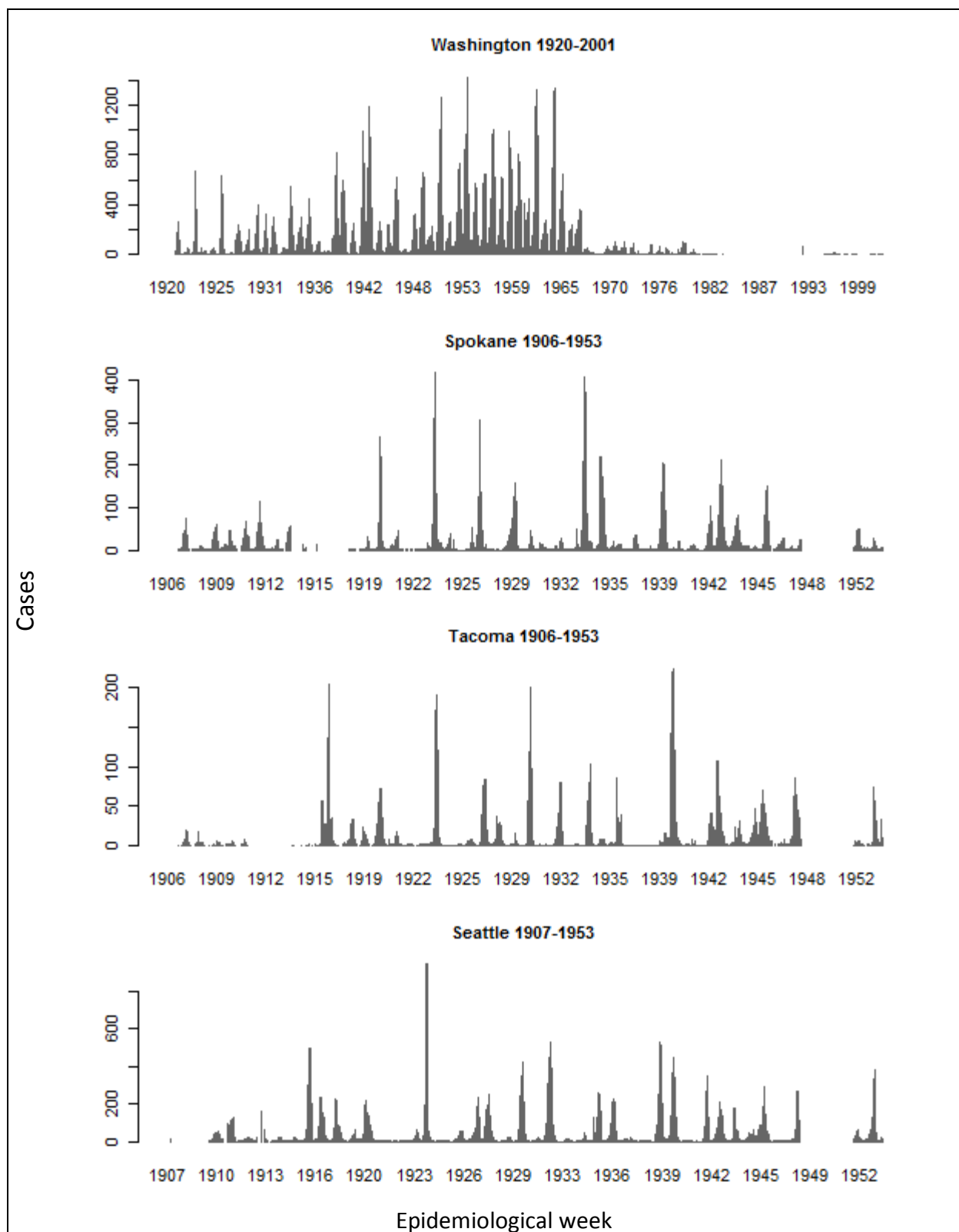*Tycho data report for the state of Washington*

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.
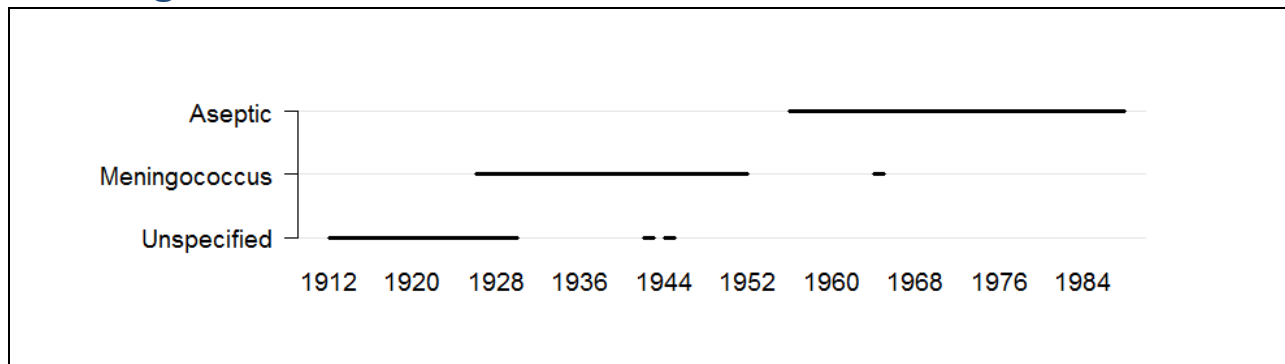
Table D1, Summary information for Measles (Unspecified)

Indicator	Washington	Spokane	Tacoma	Seattle
Report period	1920-2001	1906-1953	1906-1953	1907-1953
Total weeks	3,292	1,709	1,593	1,771
Total cases	365,952	33,895	18,780	75,744
Max. cases per year	21,574	2,682	2,011	8,296
Year (max)	1954	1939	1939	1924
Max. cases per week	1,507	773	309	1,407
Week (max)	1951, wk 20	1924, wk 03	1939, wk 49	1924, wk 03
Average cases per year				
before 1970	7,057	770	437	1,761
95%CI	(5,407-8,707)	(509-1,031)	(272-602)	(1,169-2,353)
after 1970	303	-	-	-
95%CI	(117-489)	-	-	-
Average cases per week				
before 1970	142	20	12	43
95%CI	(134-150)	(18-22)	(11-13)	(38-48)
after 1970	8	-	-	-
95%CI	(7-9)	-	-	-

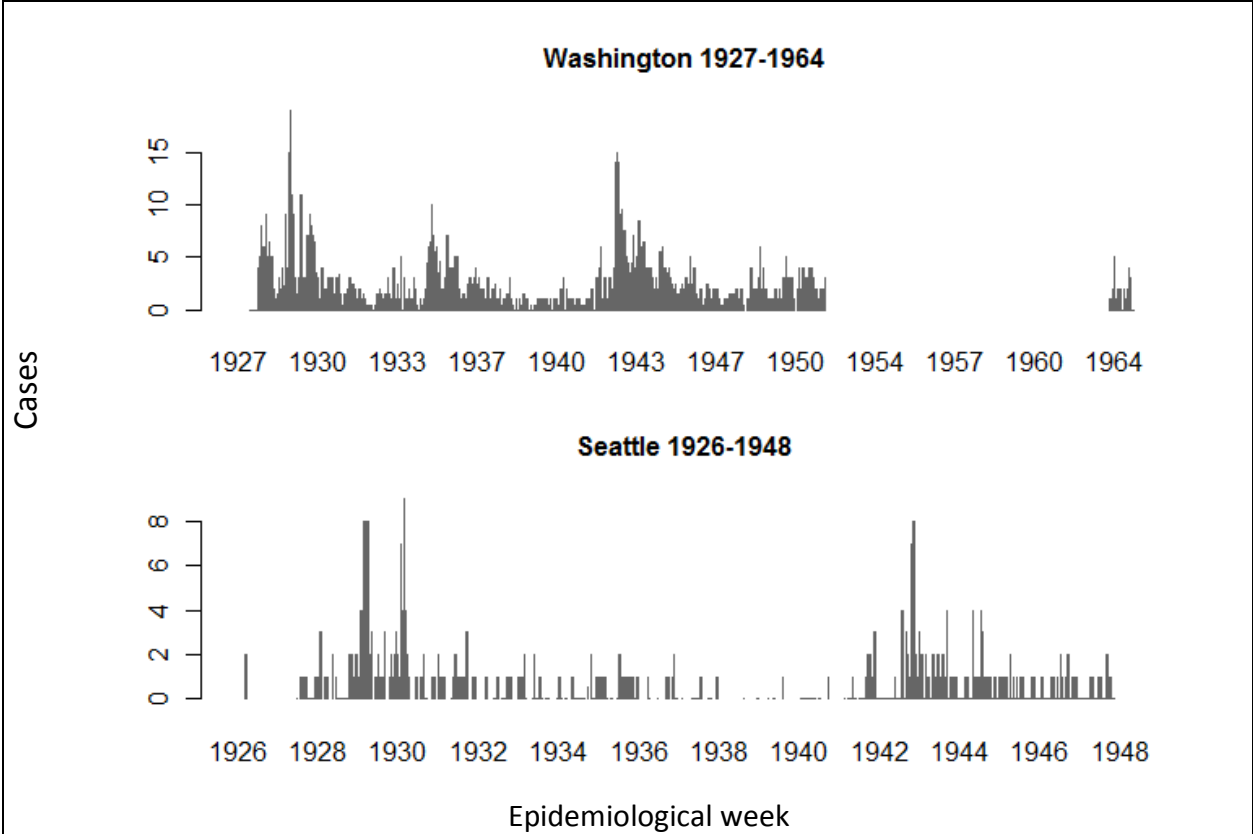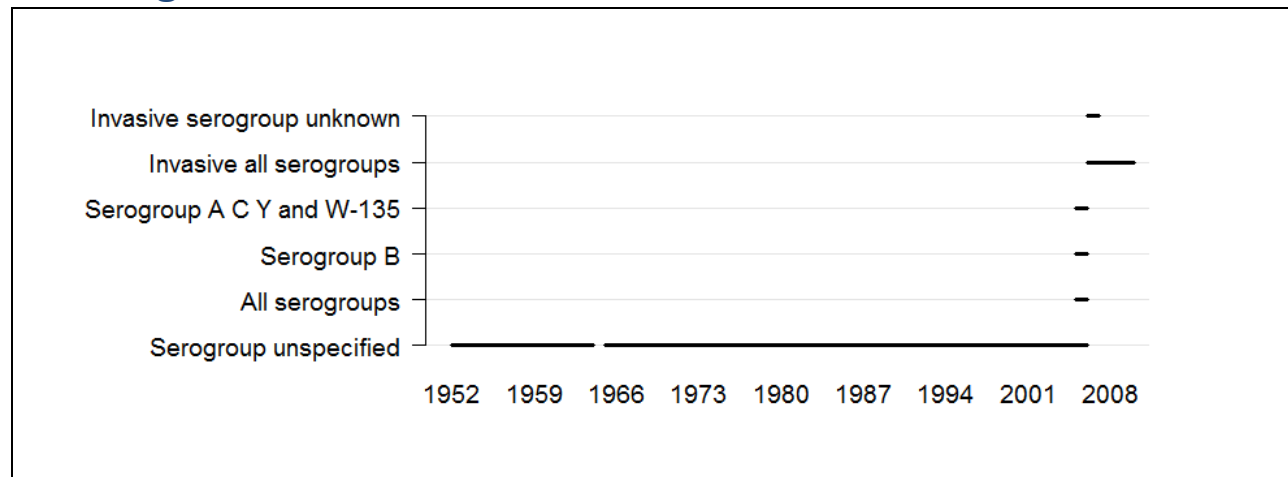**Figure D2,** Number of cases reported for Measles per epidemiological week

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	Washington	Seattle
Report period	1927-1964	1926-1948
Total weeks	1,236	607
Total cases	2,055	409
Max. cases per year	304	87
Year (max)	1943	1929
Max. cases per week	34	14
Week (max)	1929, wk 12	1929, wk 09
Average cases per year	79	18
95%CI	(49-109)	(9-27)
Average cases per week	2	1
95%CI	(2-2)	(1-1)

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

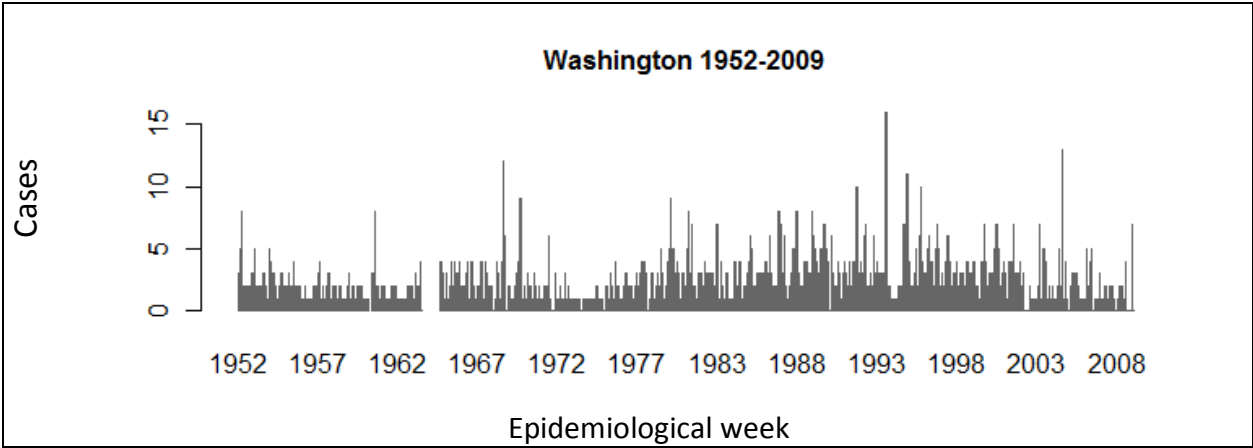*Tycho data report for the state of Washington*

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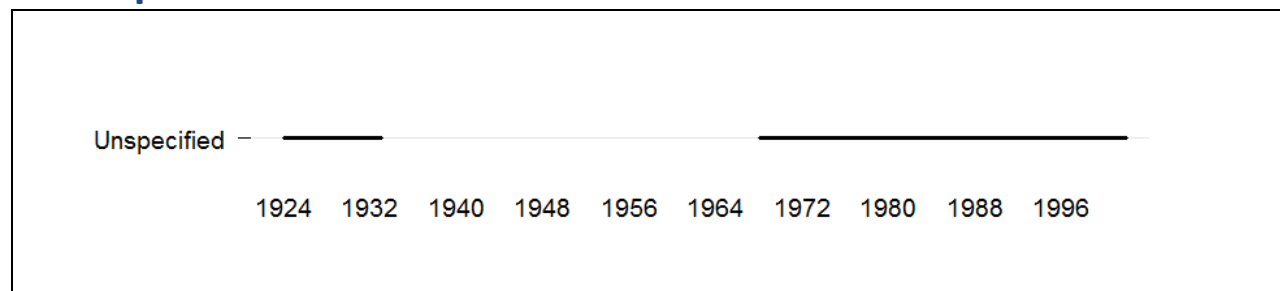
Table D1, Summary information for Meningococcal Disease (Serogroup unspecified and Invasive all serogroups)

Indicator	Washington
Report period	1952-2009
Total weeks	2,510
Total cases	3,149
Max. cases per year	326
Year (max)	1953
Max. cases per week	166
Week (max)	1953, wk 02
Average cases per year	
before 1980	53
95%CI	(31-75)
after 1980	58
95%CI	(50-66)
Average cases per week	
before 1980	1
95%CI	(1-1)
after 1980	1
95%CI	(1-1)

Tycho data report for the state of Washington

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

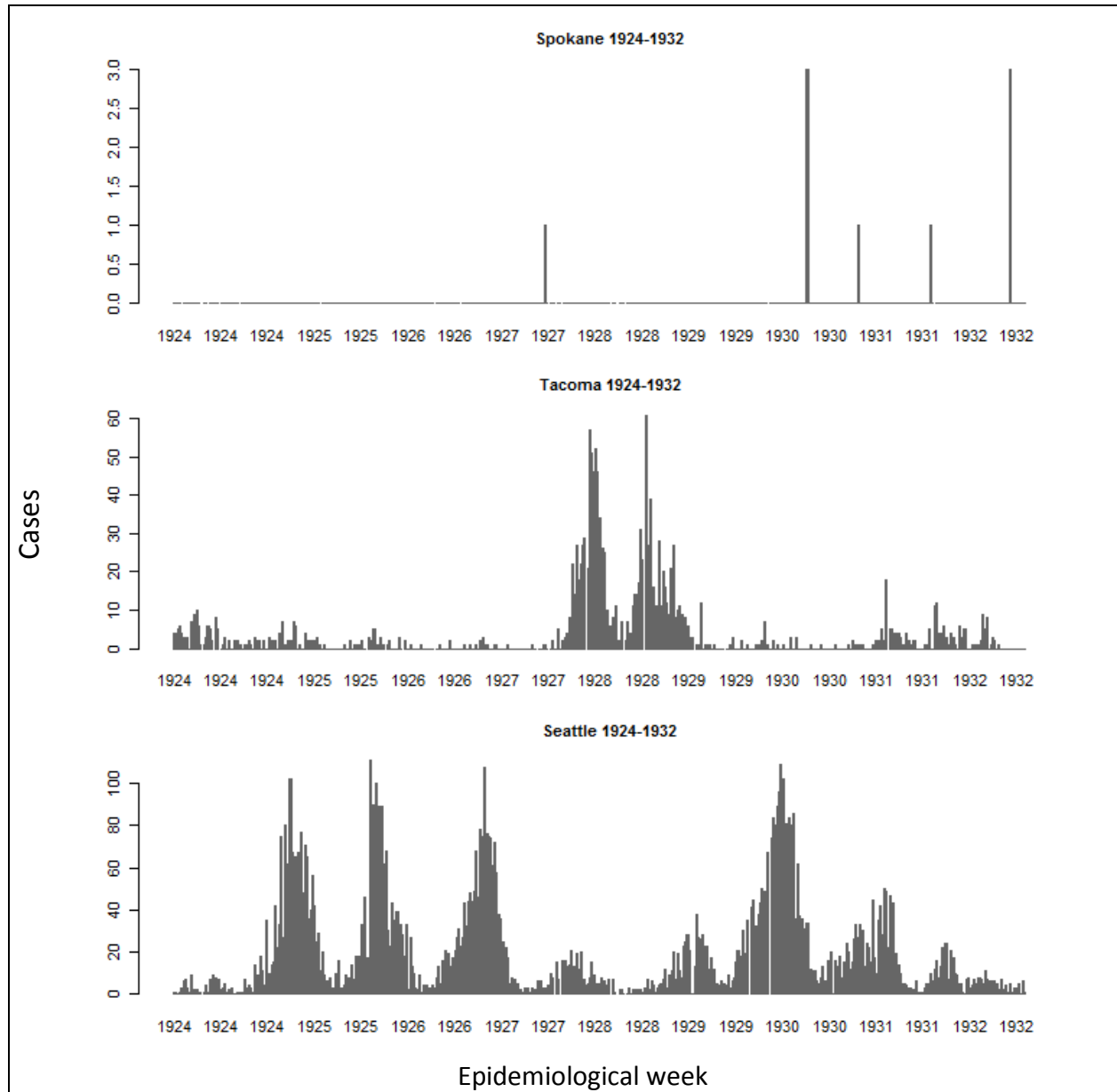*Tycho data report for the state of Washington*

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.
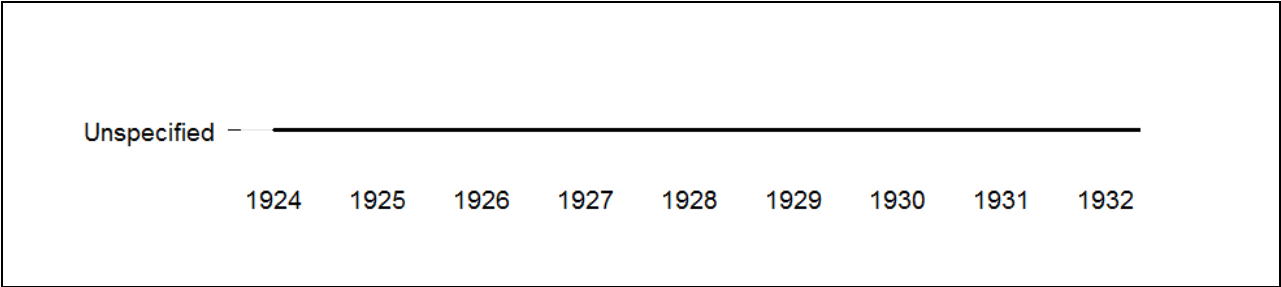
Table D1, Summary information for Mumps

Indicator	Spokane	Tacoma	Seattle
Report period	1924-1932	1924-1932	1924-1932
Total weeks	428	430	443
Total cases	51	1,783	9,178
Max. cases per year	36	1,085	2,078
Year (max)	1926	1928	1930
Max. cases per week	36	72	143
Week (max)	1926, wk 49	1928, wk 21	1926, wk 04
Average cases per year			
before 1980	6	198	1,020
95%CI	(-3-15)	(-63-459)	(482-1,558)
after 1980	-	-	-
95%CI	-	-	-
Average cases per week			
before 1980	0	4	21
95%CI	(0-0)	(3-5)	(19-23)
after 1980	-	-	-
95%CI	-	-	-

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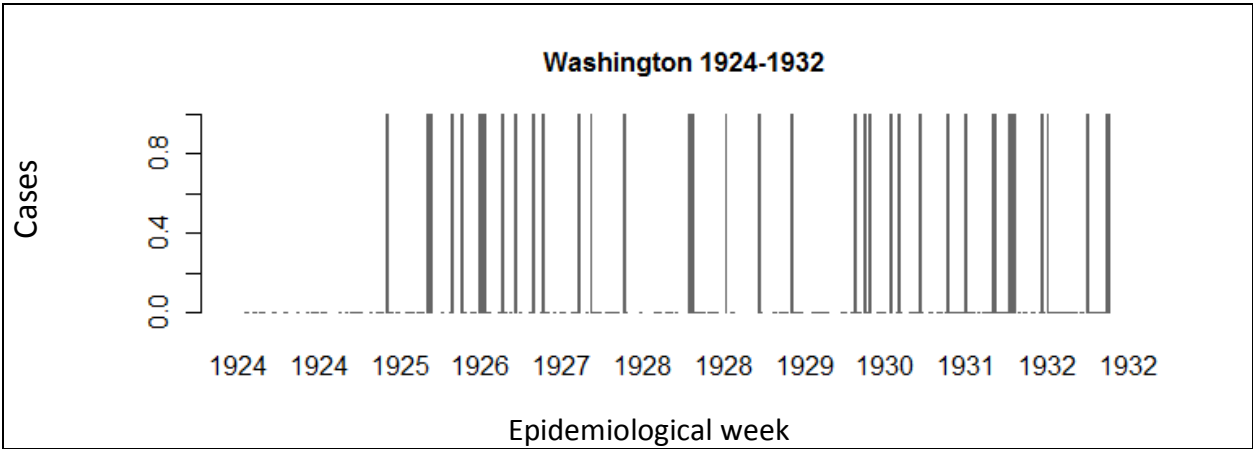
Pellagra

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.
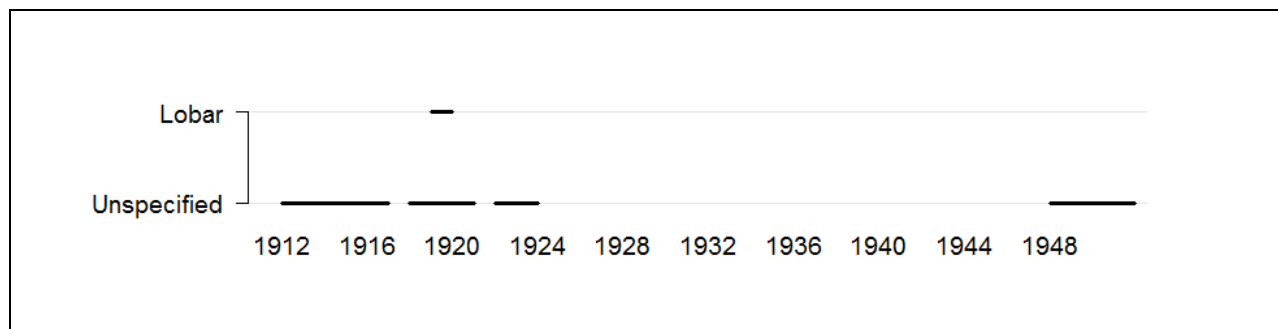
Table D1, Summary information for Pellagra

Indicator	Washington
Report period	1924-1932
Total weeks	230
Total cases	42
Max. cases per year	9
Year (max)	1926
Max. cases per week	2
Week (max)	1926, wk 36
Average cases per year	5
95%CI	(3-7)
Average cases per week	0
95%CI	(0-0)

Figure D2, Number of cases reported for Pellagra per epidemiological week
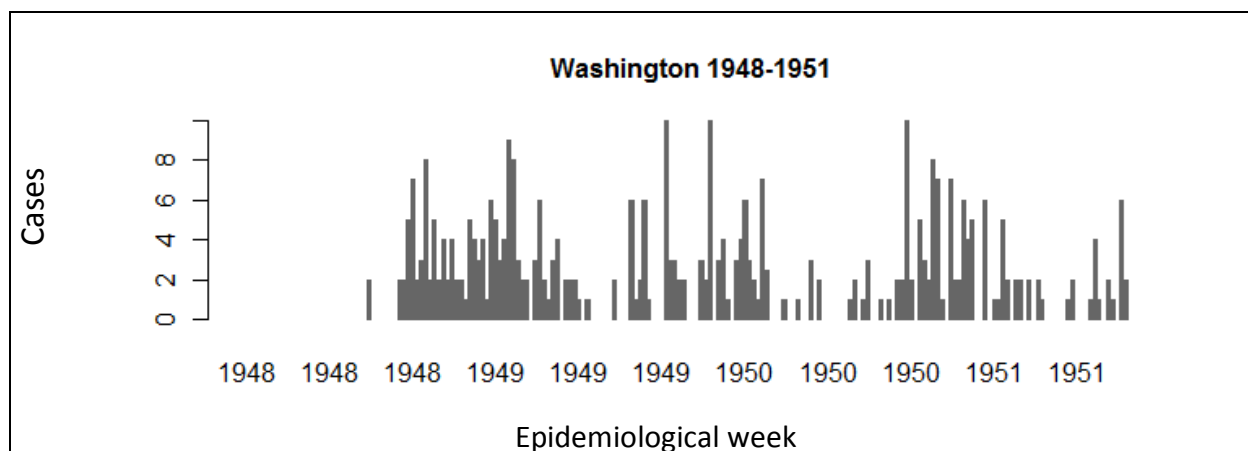
Pneumonia

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.
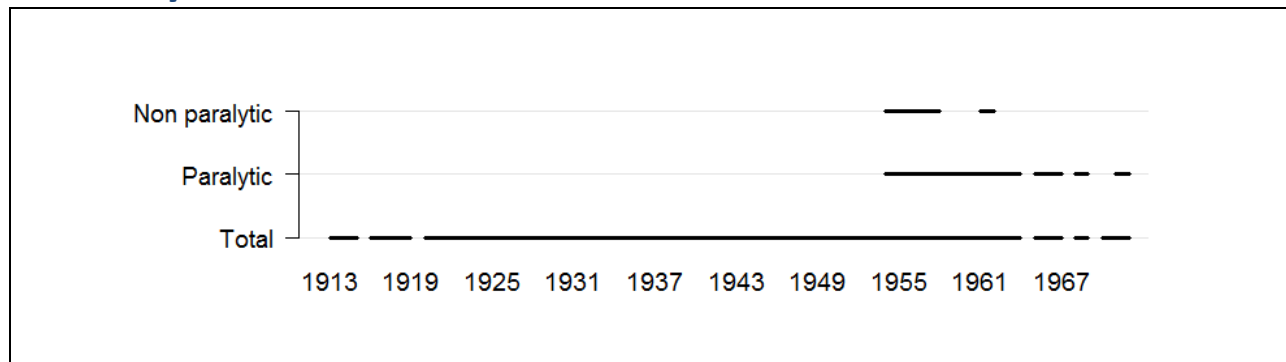
Table D1, Summary information for Pneumonia (Unspecified)

Indicator	Washington
Report period	1948-1951
Total weeks	115
Total cases	430
Max. cases per year	144
Year (max)	1951
Max. cases per week	36
Week (max)	1951, wk 12
Average cases per year	108
95%CI	(44-172)
Average cases per week	4
95%CI	(3-5)

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

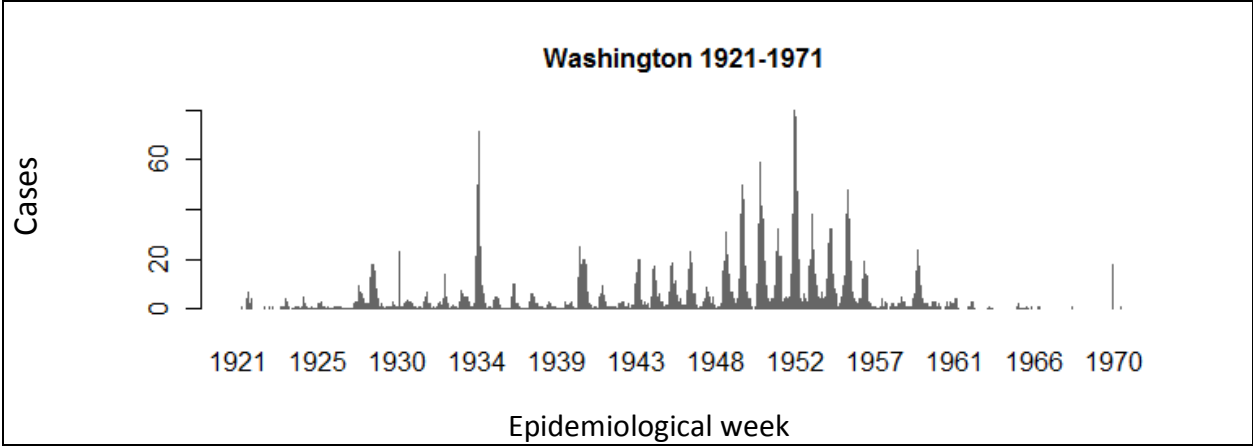*Tycho data report for the state of Washington*

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.

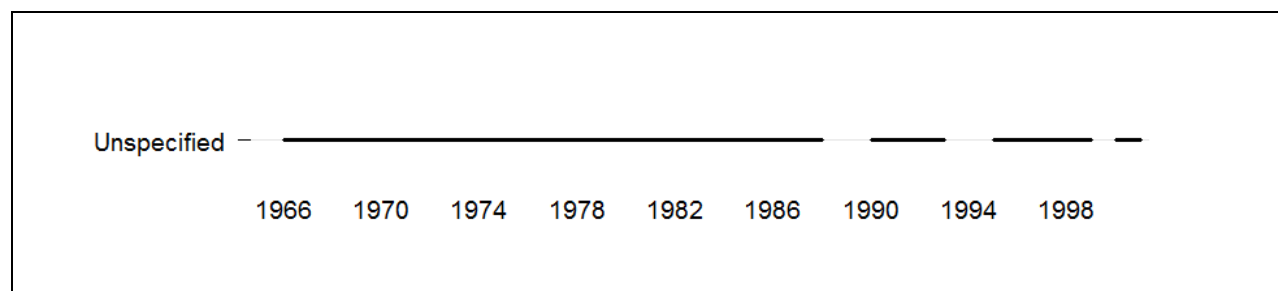**Table D1,** Summary information for Poliomyelitis (Total)

Indicator	Washington
Report period	1921-1971
Total weeks	1,856
Total cases	8,092
Max. cases per year	1,243
Year (max)	1952
Max. cases per week	102
Week (max)	1952, wk 36
Average cases per year	
before 1960	201
95%CI	(120-282)
after 1960	8
95%CI	(-2-18)
Average cases per week	
before 1960	5
95%CI	(5-5)
after 1960	1
95%CI	(1-1)

Tycho data report for the state of Washington

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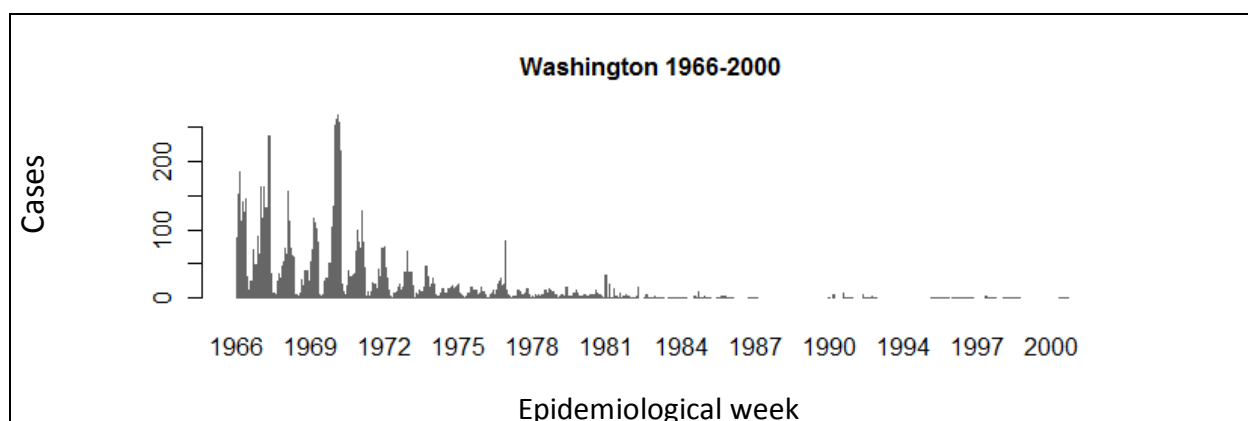
Rubella

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	Washington
Report period	1966-2000
Total weeks	1,223
Total cases	19,730
Max. cases per year	4,801
Year (max)	1970
Max. cases per week	455
Week (max)	1970, wk 19
Average cases per year	
before 1970	2,920
95%CI	(1,402-4,438)
after 1970	205
95%CI	(66-344)
Average cases per week	
before 1970	60
95%CI	(52-68)
after 1970	5
95%CI	(4-6)

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

50

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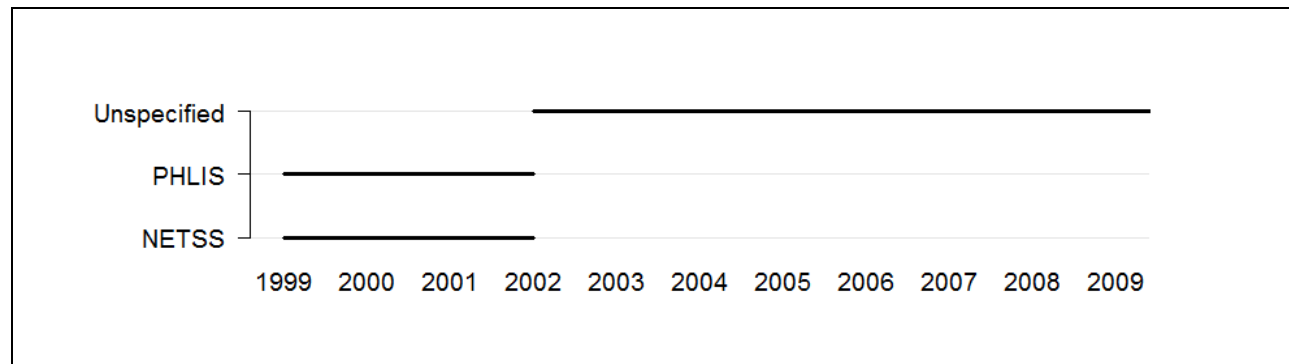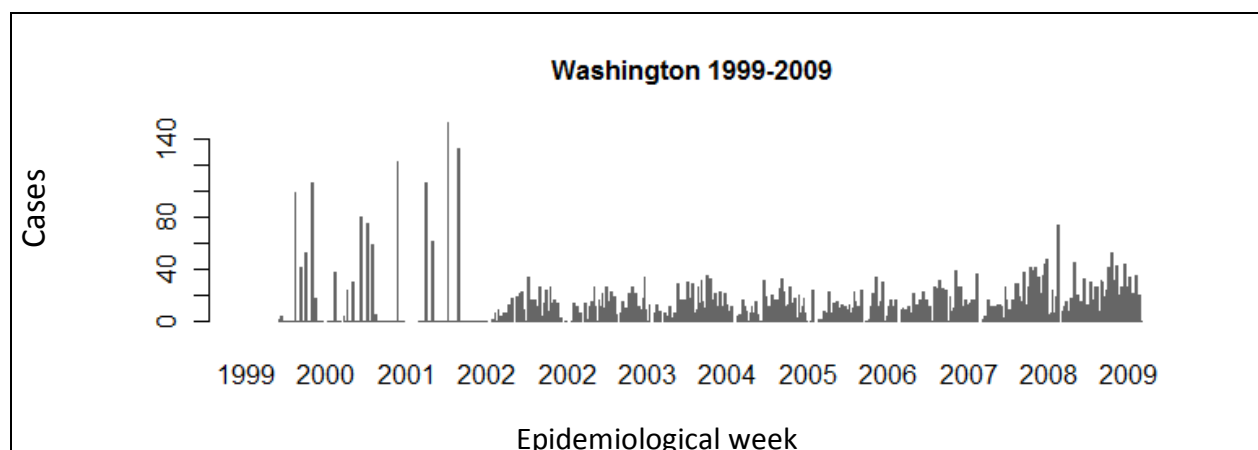
Table D1, Summary information for Salmonellosis (PHLIS and Unspecified)

Indicator	Washington
Report period	1999-2009
Total weeks	500
Total cases	6,964
Max. cases per year	1,029
Year (max)	2009
Max. cases per week	236
Week (max)	1999, wk 24
Average cases per year	633
95%CI	(504-762)
Average cases per week	14
95%CI	(12-16)

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

51

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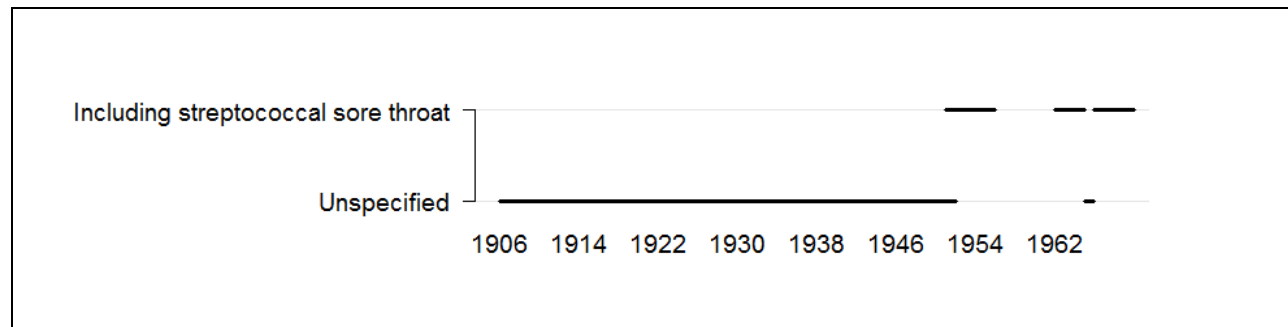
Table D1, Summary information for Scarlet Fever (Including streptococcal sore throat and Unspecified)

Indicator	Washington	Spokane	Tacoma	Seattle
Report period	1920-1969	1906-1953	1906-1953	1907-1953
Total weeks	2,207	1,797	1,707	1,797
Total cases	187,407	9,512	6,056	13,700
Max. cases per year	20,579	828	1,137	1,090
Year (max)	1969	1909	1944	1944
Max. cases per week	1,240	80	88	233
Week (max)	1968, wk 48	1909, wk 19	1944, wk 04	1944, wk 13
Average cases per year	4,259	216	141	319
95%CI	(2,450-6,068)	(153-279)	(80-202)	(248-390)
Average cases per week	85	5	4	8
95%CI	(78-92)	(5-5)	(4-4)	(8-8)
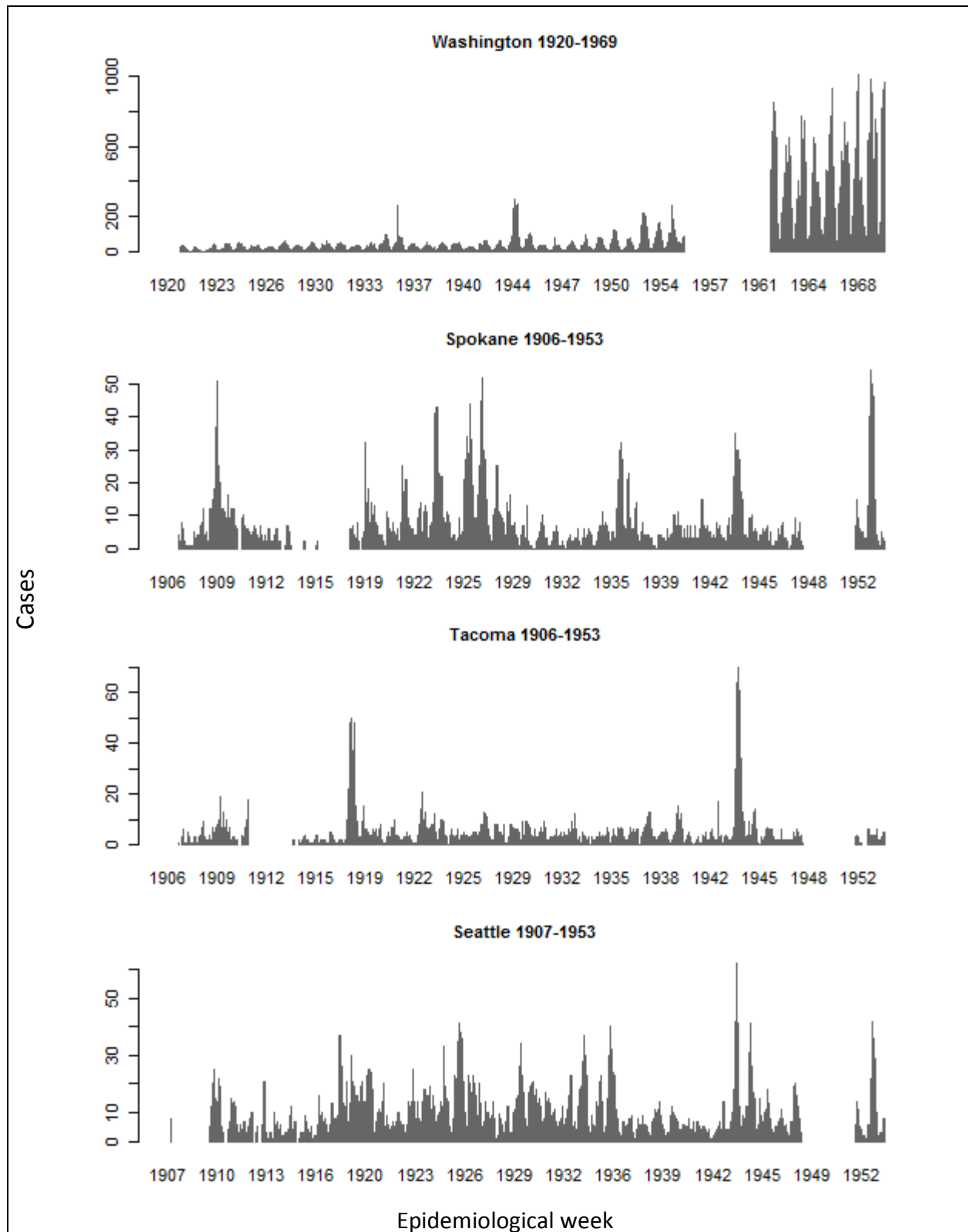
Figure D2, Number of cases reported for Scarlet Fever per epidemiological week

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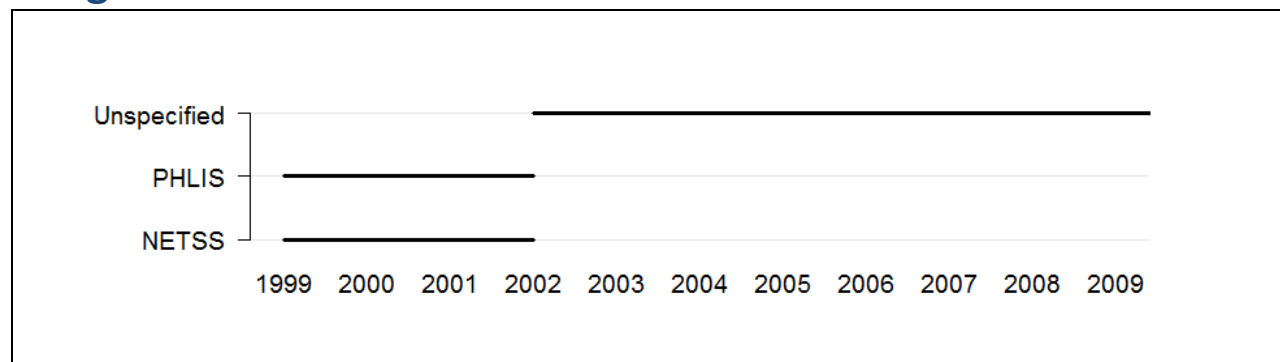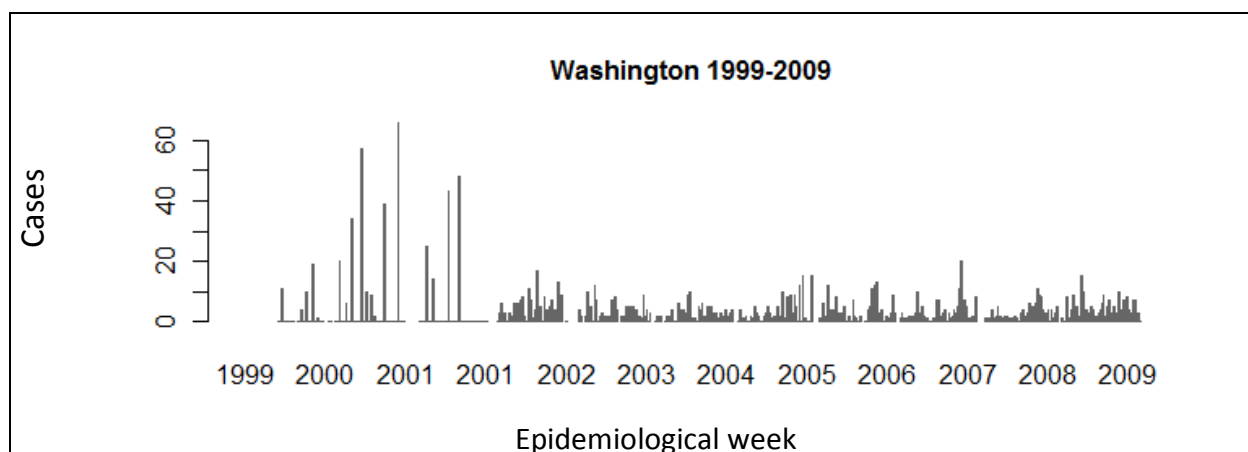
Table D1, Summary information for Shigellosis (Unspecified and PHLIS)

Indicator	Washington
Report period	1999-2009
Total weeks	492
Total cases	2,273
Max. cases per year	471
Year (max)	1999
Max. cases per week	426
Week (max)	1999, wk 33
Average cases per year	207
95%CI	(125-289)
Average cases per week	5
95%CI	(3-7)

Figure D2, Number of cases reported for Shigellosis per epidemiological week
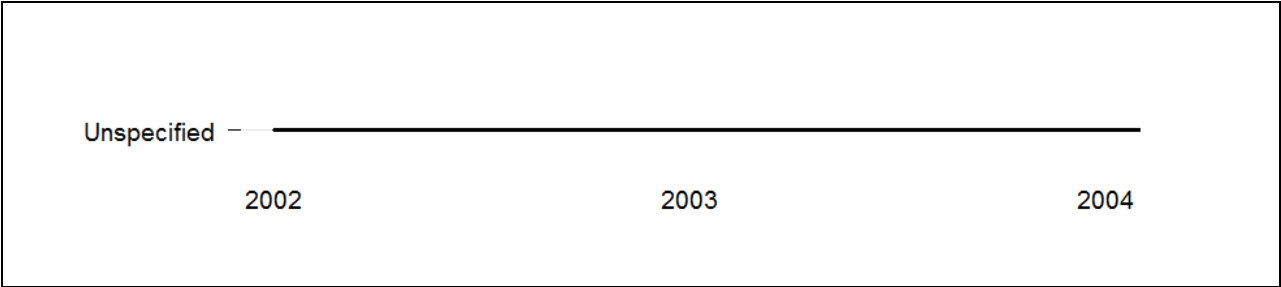
54

Streptococcal Disease, Invasive Group A

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 Disease, Invasive Group A

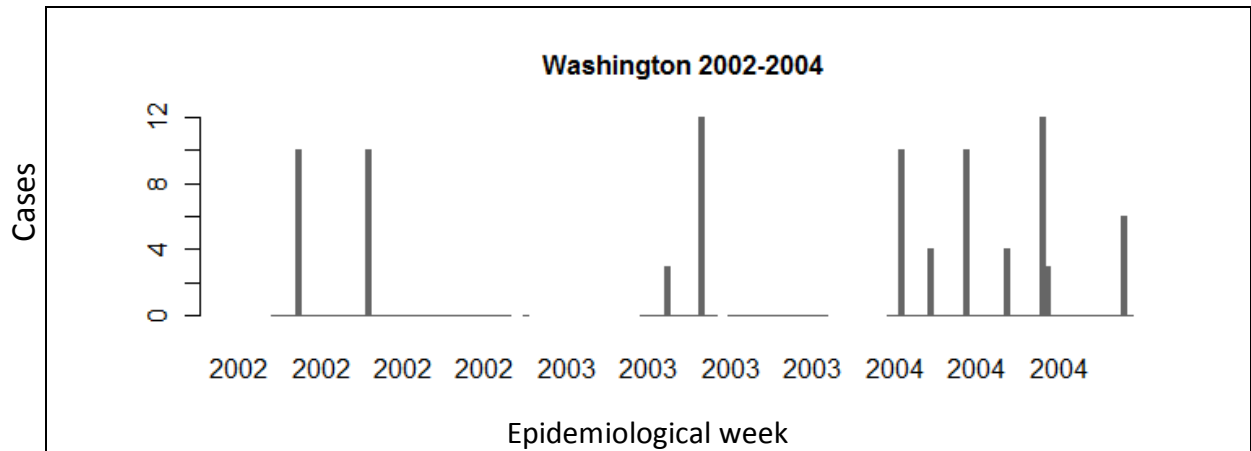| Indicator	Washington
Report period	2002-2004
Total weeks	117
Total cases	145
Max. cases per year	49
Year (max)	2002
Max. cases per week	29
Week (max)	2002, wk 34
Average cases per year	48
95%CI	(45-51)
Average cases per week	1
95%CI	(0-2)

Figure D2, Number of cases reported for Group A Streptococcal Disease per epidemiological week
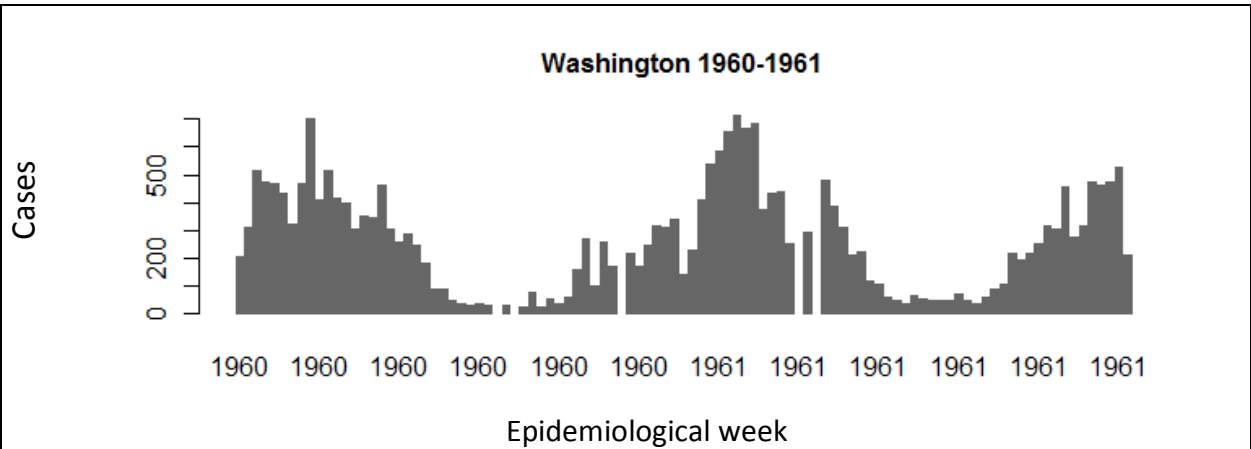
Streptococcal Sore Throat

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	Washington
Report period	1960-1961
Total weeks	99
Total cases	28,383
Max. cases per year	15,199
Year (max)	1961
Max. cases per week	1,428
Week (max)	1960, wk 50
Average cases per year	14,192
95%CI	(1,390-26,994)
Average cases per week	287
95%CI	(241-333)

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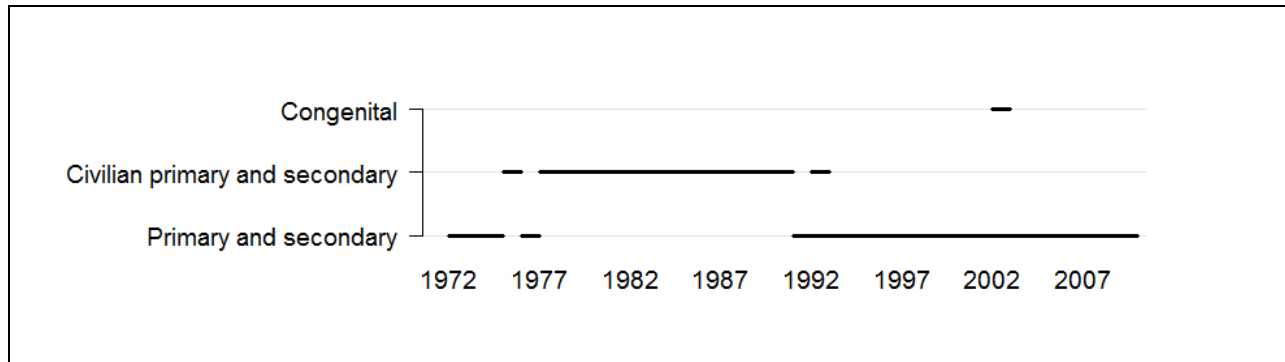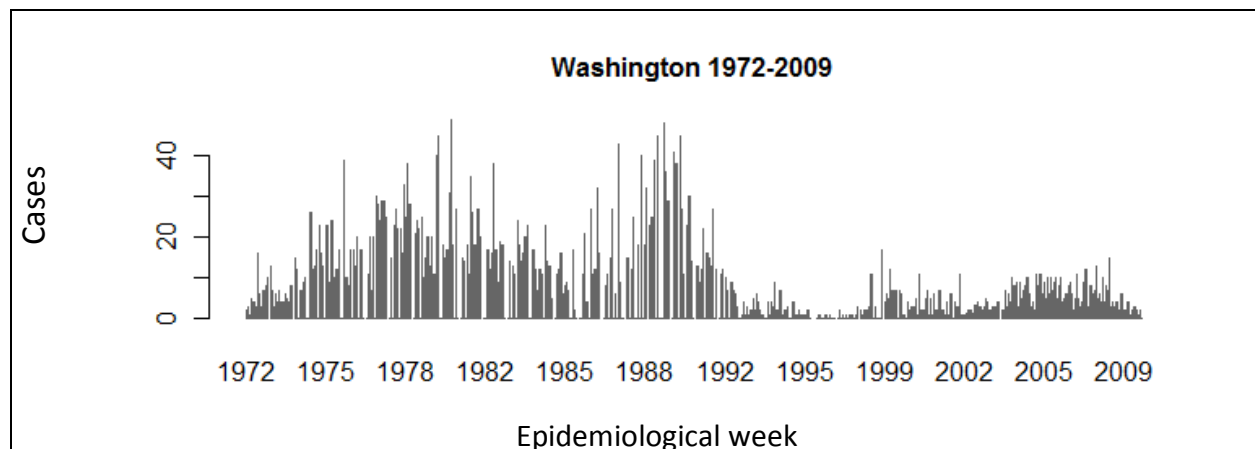
Table D1, Summary information for Syphilis (Primary and secondary and Civilian primary and secondary)

Indicator	Washington
Report period	1972-2009
Total weeks	1,680
Total cases	5,117
Max. cases per year	440
Year (max)	1989
Max. cases per week	116
Week (max)	1989, wk 32
Average cases per year	135
95%CI	(106-164)
Average cases per week	3
95%CI	(3-3)

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

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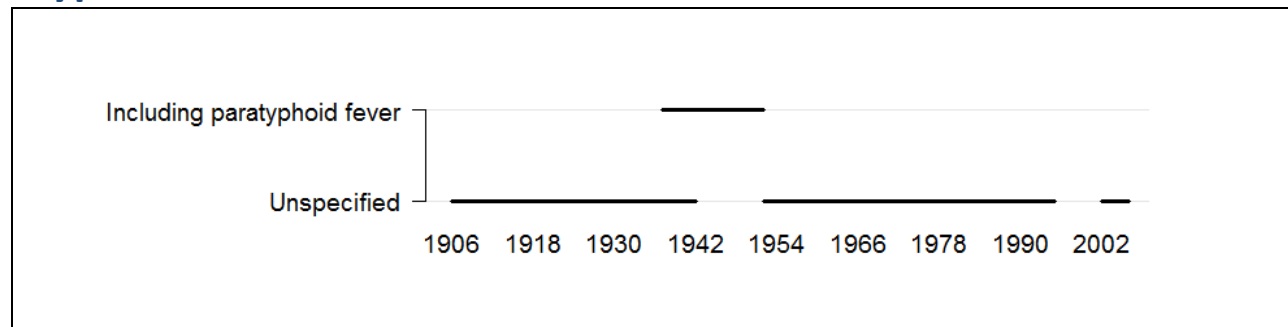
Table D1, Summary information for Typhoid Fever (Unspecified and Including paratyphoid fever)

Indicator	Washington	Spokane	Tacoma	Seattle
Report period	1920-2005	1906-1953	1906-1948	1907-1953
Total weeks	2,978	1,499	1,441	1,528
Total cases	3,337	1,249	742	1,549
Max. cases per year	210	199	112	156
Year (max)	1953	1911	1909	1910
Max. cases per week	107	24	23	23
Week (max)	1953, wk 02	1911, wk 37	1909, wk 49	1910, wk 32
Average cases per year				
before 1950	91	30	18	38
95%CI	(72-110)	(16-44)	(10-26)	(27-49)
after 1950	11	2	-	2
95%CI	(2-20)	(-11-15)	-	(-4-8)
Average cases per week				
before 1950	2	1	1	1
95%CI	(2-2)	(1-1)	(1-1)	(1-1)
after 1950	0	1	-	1
95%CI	(0-0)	(1-1)	-	(1-1)
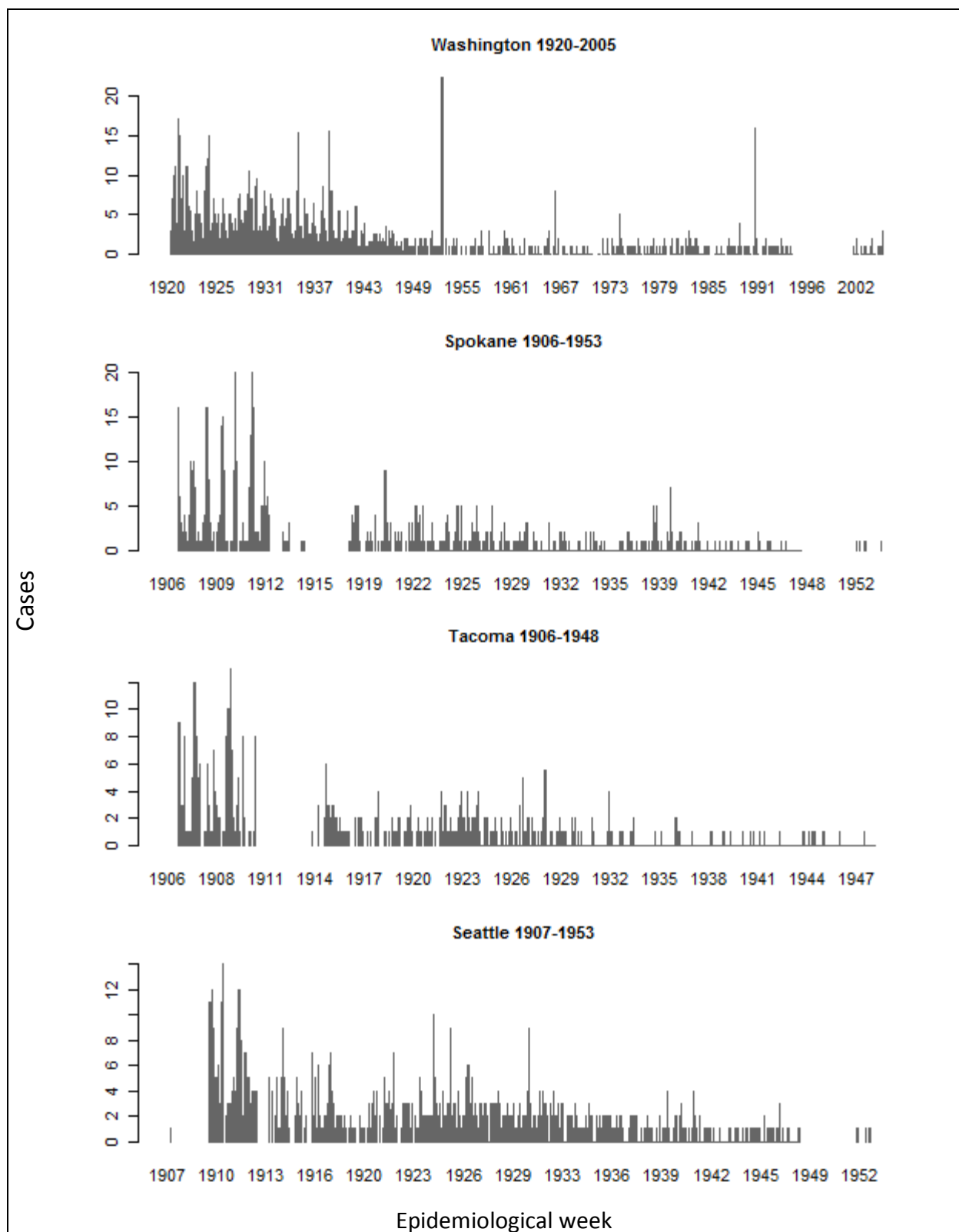
Figure D2, Number of cases reported for Typhoid Fever per epidemiological week

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.
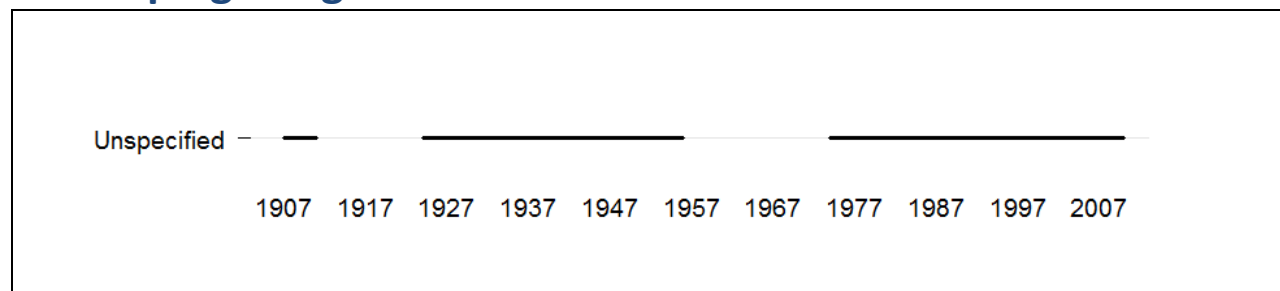
Table D1, Summary information for Whooping Cough

Indicator	Washington	Spokane	Seattle
Report period	1924-2009	1910-1953	1909-1953
Total weeks	3,093	1,195	1,286
Total cases	42,349	4,400	17,247
Max. cases per year	2,845	552	1,692
Year (max)	1941	1934	1934
Max. cases per week	769	45	115
Week (max)	2005, wk 50	1934, wk 24	1925, wk 16
Average cases per year			
before 1960	1,014	157	595
95%CI	(766-1,262)	(103-211)	(387-803)
after 1960	275	-	-
95%CI	(171-379)	-	-
Average cases per week			
before 1960	20	4	13
95%CI	(19-21)	(4-4)	(12-14)
after 1960	7	-	-
95%CI	(6-8)	-	-

Tycho data report for the state of Washington
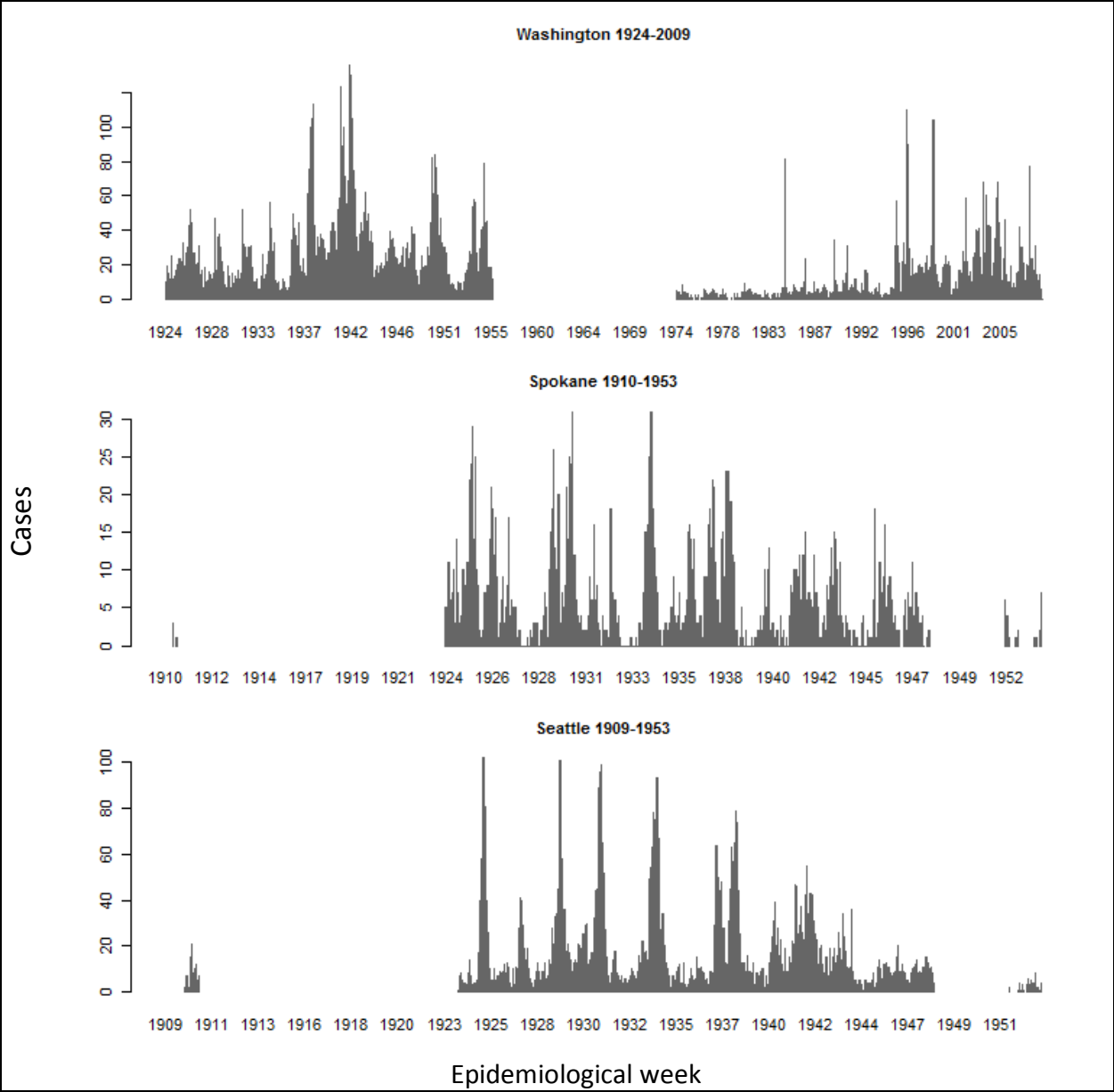
Figure D2, Number of cases reported for Whooping Cough per epidemiological week

Diseases with fragmented data over time

For a number of diseases, only fragmented data was available for Washington. 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 Washington

Disease	Report type	Report period	Number of reports	Total cases
Anthrax	State	1942-1945	143	0
Brucellosis [undulant fever]	City	1952-1952	2	2
Brucellosis [undulant fever]	State	1943-1981	550	2694
Dengue	City	1924-1925	4	0
Dysentery	State	1942-1948	217	68
Encephalitis	City	1941-1953	1346	31
Encephalitis	State	1942-1994	1057	947
Leprosy	State	1942-1988	217	1
Malaria	City	1952-1952	1	18
Malaria	State	1952-2009	1367	711
Psittacosis	State	1956-1961	86	116
Rocky mountain spotted fever	State	1942-2009	426	98
Tetanus	State	1962-1981	227	19
Trichiniasis	City	1953-1953	1	1
Trichiniasis	State	1952-1955	9	23
Tularemia	State	1942-1994	733	234
Typhus fever	City	1924-1952	25	1
Typhus fever	State	1942-1960	200	10
West nile disease	State	2004-2009	97	24

Tycho data report for the state of Washington

Project Tycho

This report provides preliminary data for the state of Washington 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. 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