

## Background

The California Department of Public Health (CDPH) maintains a mandatory, passive reporting system for a list<sup>1</sup> of communicable disease cases and outbreaks. Health care providers and laboratories are mandated to report all cases, including suspected cases, of these communicable diseases to their local health department (LHD). LHDs are also mandated to report these cases to CDPH.

These *Technical Notes* describe the definitions, methods, and limitations used to summarize the epidemiology of selected communicable diseases reported to CDPH.<sup>2</sup> The diseases selected for the 2013-2019 *Epidemiologic Summaries of Selected Communicable Diseases in California* are general communicable diseases not covered by CDPH's categorical programs for tuberculosis, sexually transmitted diseases, HIV/AIDS, and vaccine-preventable diseases, all of which produce regular summaries of their diseases.

The distribution of information on the health of the community is a core function and essential service of public health. The data in the *Epidemiologic Summaries* provide important health information on the magnitude and burden of these communicable diseases in California. Bearing in mind their limitations, these data can identify high risk groups that may benefit from public health prevention activities and aid in tracking the effectiveness of control and prevention measures.

## Materials and methods

### ***Case data sources and inclusion criteria***

These *Epidemiologic Summaries* describe incident communicable disease cases that had an estimated illness onset during January 1, 2013 through December 31, 2019. (*Epidemiologic Summaries* for Chikungunya and Zika cover the surveillance period of January 1, 2017 through December 31, 2019, since required reporting of these diseases began on June 1, 2016.) Case data were extracted from California Confidential Morbidity Reports that were submitted to CDPH by May 1, 2020 and entered into the CDPH California Reportable Disease Information Exchange (CalREDIE) system or reported electronically by CalREDIE non-participating LHDs. Cases included met the surveillance case definition per disease (see below).

Data were quality checked, and duplicate case records were removed based on a data-matching probabilistic de-duplication algorithm. For diseases that may occur as acute and chronic conditions—such as coccidioidomycosis and brucellosis—only the first report of the condition in a given patient was included based on an evaluation by the de-duplication algorithm of historical as well as recent surveillance data.

Data on foodborne and waterborne outbreaks with estimated onset dates from 2013 through 2019 were extracted from outbreak report forms submitted to CDPH by May 1, 2020. These reports were the source for the number of outbreak-associated cases per disease, when presented.

### **Population data source**

State of California, Department of Finance population projections and estimations data were used.<sup>3,4</sup>

### **Definitions**

A case was defined as one with laboratory and/or clinical evidence of infection or disease that satisfied the most recent communicable disease surveillance case definition published by the U.S. Centers for Disease Control and Prevention (CDC) or by the Council of State and Territorial Epidemiologists.<sup>5,6</sup> To determine if surveillance case definitions were met, LHDs—and for some diseases, also CDPH—reviewed detailed clinical and laboratory information provided on disease-specific case history forms. Surveillance case definitions of confirmed and/or probable cases included per disease are described in the *Epidemiologic Summaries*.

The estimated date of illness onset for each case was defined as the date closest to the time when symptoms first appeared. For cases for which an illness onset date was not explicitly reported, estimated date of illness onset was selected as the earliest of: date of diagnosis, date the case was reported to or received by CDPH, date of laboratory specimen collection, or date of patient death. For diseases with insidious onset (e.g., coccidioidomycosis), estimated onset was often based on the diagnosis date.

Mutually exclusive race/ethnicity categories were defined as follows: Hispanic/Latino (of any, including unknown, race), and non-Hispanic White, Black, Asian/Pacific Islander, American Indian/Alaska Native, Multiple Race, and Other.

Cases were classified geographically according to the case-patient's county of residence. Cases reported from the City of Berkeley were included in Alameda County, and cases from the cities of Long Beach and Pasadena were included in Los Angeles County. Regions of California were defined by aggregating counties with similar geography, demography, and economic conditions as described by the Public Policy Institute of California.<sup>7</sup> Regions included the Far North (Butte, Colusa, Del Norte, Glenn, Humboldt, Lake, Lassen, Mendocino, Modoc, Nevada, Plumas, Shasta, Sierra, Siskiyou, Sutter, Tehama, Trinity, and Yuba counties); Sacramento Metro (El Dorado, Placer, Sacramento, and Yolo counties); Sierra (Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, and Tuolumne counties); Bay Area (Alameda, Contra Costa, Marin, Napa, San Francisco, San Mateo, Santa Clara, Solano, and Sonoma counties); San Joaquin Valley (Fresno, Kern, Kings, Madera, Merced, San Joaquin, Stanislaus, and Tulare counties); Central Coast (Monterey, San Benito, San Luis Obispo, Santa Barbara, and Santa Cruz counties); Inland Empire (Riverside and San Bernardino counties); South Coast (Los Angeles, Orange, and Ventura counties); and San Diego (Imperial and San Diego counties). Southern California was defined as the counties

comprising the Inland Empire, South Coast, and San Diego regions, while all other counties comprised Northern California.

### **Data analyses**

Case totals and, when the data were sufficient, incidence rates per 100,000 population were reported and stratified by estimated year of illness onset, county, region, sex, and age group.

The formulas used to calculate the incidence rate and relative standard error were:

- Incidence rate (IR) = Number of cases/population x 100,000
- Standard error (SE) =  $IR/\sqrt{\text{number of cases}}$
- Relative standard error =  $SE/IR \times 100$

An incidence rate was considered unstable if the relative standard error was 23 percent or more (a threshold recommended by the National Center for Health Statistics<sup>8</sup>).

To reduce the level of random error when the case number or population was small, the time and geographic range for incidence rates was expanded, and multiple-year average annual incidence rates and region-specific (rather than county-specific) rates were calculated, as needed. Relative standard errors were calculated for all incidence rates.

Because a substantial portion of all case-patients during the surveillance period did not identify their race/ethnicity (28.6%) or identified as non-Hispanic Other race/ethnicity (5.4%), incidence rates by race/ethnicity were not calculated. However, since race/ethnicity can be an important marker for complex social, economic, and political factors that influence health, the racial/ethnic distribution of cases and the statewide population were presented side by side when the data were sufficient. Cases of non-Hispanic Other race/ethnicity were not included in the analysis of race/ethnicity due to lack of population data for this group.

For some rare diseases, the *Epidemiologic Summaries* present a description of the number of cases rather than incidence rates.

Analyses were conducted using SAS software version 9.4, and maps were created using ArcGIS software version 10.7.1.

## **Limitations**

### **Data quality**

For many of the diseases covered in the *Epidemiologic Summaries*, CDPH reviewed case history information to determine whether surveillance case definitions were met. However, for campylobacteriosis, coccidioidomycosis, Creutzfeldt-Jakob disease, cryptosporidiosis, cyclosporiasis, giardiasis, listeriosis (through 2015), salmonellosis, shigellosis, typhoid fever (through 2015), vibriosis (through 2015), and yersiniosis, CDPH relied on LHDs to apply surveillance criteria. It is possible that some cases included in this report did not meet surveillance case definitions and counting criteria.

## ***Deaths***

For some diseases, the number of case-patients who died of their illness was calculated based on date of death and other death-related information reported on the Confidential Morbidity Report or case history. However, deaths might have occurred after the case report was completed (and thus were not included in the calculated numbers). The numbers of deaths and case-fatality ratios reported should be interpreted with caution.

## ***Completeness and timeliness of reporting***

The numbers of cases reported for some diseases in these *Epidemiologic Summaries* are likely to underestimate the true magnitude of disease. Factors that may contribute to under-reporting include delays in notification, limited collection or appropriate testing of specimens, obstacles or impediments to ill persons seeking health care, limited resources and competing priorities in LHDs, and lack of reporting by clinicians and laboratories. Factors that may contribute to enhanced reporting include disease severity, the availability of new or less expensive diagnostic tests, changes in the case definition by CDC or CDPH, recent media attention or public interest, and active surveillance activities.

During the seven-year surveillance period (2013-2019), the CDC and CDPH Infectious Diseases Branch conducted active surveillance of selected diseases in Alameda, Contra Costa, and San Francisco counties through the California Emerging Infections Program (CEIP). CEIP conducted active laboratory-based surveillance of campylobacteriosis, cryptosporidiosis, cyclosporiasis, Shiga toxin-producing *E. coli* (STEC) infection, listeriosis, salmonellosis, shigellosis, *Vibrio* infection, yersiniosis, and pediatric hemolytic uremic syndrome. Therefore, cases of these diseases may be more completely reported in these counties.

Because outbreak-related case reports were not always identified as such on the Confidential Morbidity Report, it was not possible to ascertain the proportion of outbreak-related cases that were also reported as individual cases. Additionally, case definitions used to classify probable outbreak-related cases may not meet the specific surveillance criteria required for individual case reporting. Thus, some outbreak-related cases may not be included in the total number of cases reported for each disease.

## ***Small numbers and rate variability***

All rates, even those based on full population counts, are subject to random error. Random error may be substantial when the number of cases is small (e.g., less than 20) and can make it impossible to distinguish random fluctuations from true changes in the underlying risk of disease. Rates and proportions based on small numbers should be interpreted with caution.

## ***Count and Rate comparisons***

Incidence rate comparisons between geographic areas and over time should be made with caution. The limitations previously listed (especially the completeness of reporting

and random variability of rates) should be considered when interpreting and comparing incidence rates.

Data presented in the *Epidemiologic Summaries* may differ from previously published data due to delays inherent to case reporting, laboratory reporting, and epidemiologic investigation.

Updated by Kirsten Knutson, Yanyi Djamba, and Allyx Nicolici – Infectious Diseases Branch, December 2021

## References

<sup>1</sup> [Reportable Diseases and Conditions: Reporting to the Local Health Authority, 17 CCR § 2500 \(2021\)](#).

<https://govt.westlaw.com/calregs/Document/I5849DB60A9CD11E0AE80D7A8DD0B623B>

<sup>2</sup> [Epidemiologic Summaries of Selected Communicable Diseases in California. California Department of Public Health website](#). Accessed December 27, 2021.

<https://www.cdph.ca.gov/Programs/CID/DCDC/Pages/EpiSummariesofCDsCA.aspx>

<sup>3</sup> State of California, Department of Finance, P-3 State and County Population Projections by Race/Ethnicity and Detailed Age, 2010–2060. Sacramento, California, January 2020.

<sup>4</sup> State of California, Department of Finance, E-4 Population Estimates for Cities, Counties, and the State, 2011–2019, with 2010 Census Benchmark. Sacramento, California, May 2019.

<sup>5</sup> [Surveillance Case Definitions for Current and Historical Conditions. National Notifiable Diseases Surveillance System \(NNDSS\), U.S. Centers for Disease Control and Prevention website](#). Accessed September 22, 2020.

<https://wwwn.cdc.gov/nndss/conditions/>

<sup>6</sup> U.S. Centers for Disease Control and Prevention. [Case definitions for infectious conditions under public health surveillance](#). MMWR 1997;46(No.RR-10):1-55.

[https://wwwn.cdc.gov/nndss/document/1997\\_MMWR\\_RR\\_Case\\_Definitions\\_PH\\_Surveillance\\_rr4610.pdf](https://wwwn.cdc.gov/nndss/document/1997_MMWR_RR_Case_Definitions_PH_Surveillance_rr4610.pdf)

<sup>7</sup> Johnson, H. [A state of diversity: Demographic trends in California's regions](#). *California Counts: Population Trends and Profiles*. 2002 May; 3(1). San Francisco, CA: Public Policy Institute of California, 2002. Accessed September 22, 2020.

[http://www.ppic.org/content/pubs/cacounts/CC\\_502HJCC.pdf](http://www.ppic.org/content/pubs/cacounts/CC_502HJCC.pdf)

---

<sup>8</sup> Xu JQ, Murphy SL, Kochanek KD, Bastian B, Arias E. [Deaths: Final data for 2016](https://www.cdc.gov/nchs/data/nvsr/nvsr67/nvsr67_05.pdf). *National Vital Statistics Reports*. 2018; 67(5). Hyattsville, MD: National Center for Health Statistics, 2018. Accessed September 22, 2020.  
[https://www.cdc.gov/nchs/data/nvsr/nvsr67/nvsr67\\_05.pdf](https://www.cdc.gov/nchs/data/nvsr/nvsr67/nvsr67_05.pdf)