

Background

The California Department of Public Health (CDPH) maintains a mandatory, passive reporting system for a list of communicable disease cases and outbreaks¹. Health care providers and laboratories are mandated to report cases or 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 CDPH². In particular, these selected communicable diseases come from the general communicable diseases not covered by the 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 communicable diseases in California. Bearing in mind their limitations, these data can contribute toward identifying high risk groups needing preventive actions and tracking the effectiveness of control and prevention measures.

Materials and methods

Case data sources and inclusion criteria

We extracted data on communicable disease cases with an estimated onset date from 2001 through 2008 from California Confidential Morbidity Reports that were submitted to CDPH by May 8, 2009 and which met the surveillance case definitions (see below). Because of inherent delays in case reporting and depending on the length of follow-up clinical, laboratory and epidemiologic investigation, cases with eligible onset dates may be added or rescinded after the date of this report. Therefore, **data for 2008 contained in this report are provisional** and may differ from data published in future reports.

CDPH reviewed detailed clinical and laboratory data provided on disease-specific case history forms to determine if surveillance case definitions were met. LHDs applied surveillance criteria for diseases that did not require a case history form by regulation (campylobacteriosis, coccidioidomycosis, cryptosporidiosis, giardiasis, salmonellosis, and shigellosis).

We extracted data on foodborne and waterborne outbreaks with estimated onset dates from 2001 through 2008 from outbreak report forms submitted to CDPH by July 1, 2009. These reports were the source for the number of outbreak-associated cases for each disease.

Population data source

We used projections for state, county, and age-specific population totals that were published in: State of California, Department of Finance, *Race/Ethnic Population with Age and Sex Detail, 2000–2050*. Sacramento, CA, July 2007.

Definitions

In general, we defined a case as laboratory and/or clinical evidence of infection or disease in a person that satisfied the most recent communicable disease surveillance case definition published by the United States (US) Centers for Disease Control and Prevention (CDC) or by the Council of State and Territorial Epidemiologists (CSTE)³. Surveillance case definitions are described in individual disease summaries. By California regulation, an animal case was one that was determined, by a person authorized to do so, to have rabies or plague.

We defined the estimated onset date for each case as the date closest to the time when symptoms first appeared. Because date of onset may not be recorded, the estimated date of onset can range from the first appearance of symptoms to the date the report was made to CDPH. For diseases with insidious onset (for instance, coccidioidomycosis), estimated onset was more frequently drawn from the diagnosis date. We defined the surveillance period as 2001 through 2008.

We defined single race-ethnicity categories as follows: Hispanic (of any, including unknown, race); White, non-Hispanic; Black, non-Hispanic; Asian/Pacific Islander, Native American; and Other or multi-race. Cases with unknown race and ethnicity were listed as unknown.

We defined regions of California by collapsing counties with similar geography, demography, and economic conditions as described by the Public Policy Institute of California⁴. 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*); Sierras (*Alpine, Amador, Calaveras, Inyo, Mari-
posa, 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*). We defined Southern California as the counties comprising the Inland Empire, South Coast, and San Diego regions. All other counties comprised Northern California.

We defined a rate as unreliable if the relative standard error was 23 percent or more (a threshold recommended by the National Center for Health Statistics). The formulas used to calculate the 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 = $SE/IR \times 100$

Data analyses

We reported case totals and rates per 100,000 population (unless otherwise indicated) stratified by estimated year of onset, age, and geographic residence. We calculated geographic-based rates by county, region, and bisection of the State (Northern or Southern California). 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.

To reduce the level of random error, we expanded the time and geographic range for incidence rates when few cases or small populations were identified. We produced multiple-year average rates and region-specific (rather than county-specific) rates, as needed. We calculated relative standard errors for all county-specific rates.

Because a substantial portion of race/ethnicity data was missing (disease-specific range: 12 to 50 percent), we did not calculate incidence rates. However, because race/ethnicity can be an important marker for complex social, economic, and political factors that influence health, we presented the distribution of single race/ethnicity categories among cases with complete information.

We evaluated the temporal trends in incidence rates for selected diseases using Poisson regression models. Values of $p < 0.05$ were considered statistically significant. Analyses were conducted using SAS Release 9.1 (SAS Institute, Inc, Cary North Carolina) and maps were created using ArcGIS version 9.3 (ESRI, Inc, Redlands, California).

Limitations

Data quality

CDPH relied on LHDs to apply surveillance and counting criteria for campylobacteriosis, coccidioidomycosis, cryptosporidiosis, giardiasis, salmonellosis, and shigellosis. It is possible that some cases did not meet surveillance case definitions or counting criteria.

Deaths

We presented the number of cases reported to CDPH as having died with their disease. There is no standardized method for determining whether a communicable disease caused or contributed to the death for the purposes of reporting here. Deaths may have occurred after the report was filed (and thus not reported). The numbers of deaths and case-fatality ratios reported should be interpreted with caution.

Completeness of reporting

The numbers of disease cases in this report are likely to underestimate the true magnitude of disease. Among factors that may contribute to underreporting are: delays in notification, limited collection or appropriate testing of specimens, health care seeking behavior among ill persons, limited resources and competing priorities in LHDs, and

lack of cooperation of clinicians and laboratories. Among factors that may contribute to increased reporting are disease severity, the availability of new or less expensive diagnostic tests, changes in the case definition by CDC or CDPH, recent media or public attention, and active surveillance activities.

During the surveillance period, CDC and CDPH conducted active surveillance in Alameda, Contra Costa, and San Francisco Counties through the California Emerging Infections Program (CEIP). CEIP conducted active laboratory-based surveillance for *Salmonella*, *Shigella*, *Campylobacter*, *Escherichia coli* O157, Shiga toxin-producing *E. coli* (STEC) non-O157, *Listeria monocytogenes*, *Yersinia*, *Vibrio*, *Cryptosporidium*, and *Cyclospora* infection and active physician-based surveillance of pediatric hemolytic uremic syndrome (HUS) through a network of nephrologists in the catchment area.

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 reported as individual cases in the passive reporting system. Additionally, case definitions used to classify probable outbreak-related cases may not meet the more specific criteria required for individual case reporting. Therefore, outbreak-related cases may not be included in the total number of cases reported for each disease and outbreak-related cases reported in the probable classification may not meet surveillance reporting criteria.

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.

Rate comparisons

Incidence rate comparisons between geographic entities and over time should be done with caution. Because not all LHDs reported age data, the rates in this report are not age-adjusted. Additionally, the limitations previously listed (especially the completeness of reporting and random variability of rates) should be considered when interpreting and comparing incidence rates.

Acknowledgements

We thank California Communicable Disease Controllers and program staff for providing these data and for their continued diligence in communicable disease control. We thank CDPH subject matter experts for their expertise and collaboration in producing the disease summaries.

References

- ¹California Code of Regulations, Title 17, Sections 2500, 2501, 2502, 2503, 2505
<http://www.cdph.ca.gov/HealthInfo/Pages/ReportableDiseases.aspx>
- ²Epidemiologic Summaries of Selected General Communicable Diseases in California, 2001 - 2008.
<http://www.cdph.ca.gov/data/statistics/Pages/EpiSummariesCDsCA-01-08.aspx>
- ³Centers for Disease Control and Prevention. National Notifiable Diseases Surveillance System, Case Definitions for Infectious Conditions Under Public Health Surveillance
http://www.cdc.gov/ncphi/diss/nndss/casedef/case_definitions.htm
- ⁴Johnson, H. 'A State of Diversity: Demographic Trends in California's Regions', *California Counts*, Vol 3, No 1, Public Policy Institute of California, San Francisco, California 2002
http://www.ppic.org/content/pubs/cacounts/CC_502HJCC.pdf

Last updated 8/17/2009

Prepared by Kate Cummings, MPH, and Farzaneh Tabnak, PhD, Infectious Diseases Branch