<mark>On October 5, 2012 a revised version of this dataset was posted to the Open Data Portal. See **Version History** on page 2 of this document for details.</mark>

Title: Tuberculosis cases and average annual incidence rate, Chicago, 2007-2011.

Brief Description: The annual number of new cases of tuberculosis and average annual tuberculosis incidence rate (new cases per 100,000 residents) with corresponding 95% confidence intervals, by Chicago community area, for the years 2007 – 2011.

Full Description: CDPH calculated the indicators using case reports and laboratory reports received under the notifiable disease rules of the Illinois Administrative Code (see http://www.ilga.gov/commission/jcar/admincode/077/07700690sections.html).

Community area case counts are derived through geocoding of the address of residence of each person diagnosed with active tuberculosis. For individuals with no fixed address (i.e., homeless), the address of residence used is the best estimate of where the patient was residing (e.g., shelter), or if not applicable, the hospital at which the patient received initial tuberculosis care. Geocoding is the process of using location data, such as street address, to determine associated geographic identifiers, such as latitude and longitude, postal code, or community area. Indicators for Chicago as a whole are provided in the final row of the table.

A confidence interval is a range of values used to describe the uncertainty around a measurement (e.g., rate) and serves as a measure of the variability in the data. Confidence intervals are calculated based on the standard error of the rate, which is based on the rate and the number of events (e.g., new cases). Most confidence intervals are calculated as 95% confidence intervals by convention. The 95% confidence interval can best be understood that if the measurement were conducted 100 times, 95 times the true value would be within the calculated confidence interval and 5 times the true value would be either higher or lower than the range of the confidence interval. For example, a tuberculosis incidence rate of 10 per 100,000 with a lower limit of 8 and an upper limit of 12 means that there is a 95 percent chance that the rate was between 8 and 12 per 100,000. Conversely, there is a 5 percent chance that the rate was lower than 8 or higher than 12. Incidence rates are subject to random error, arising from random fluctuations in the number of diagnoses and case reports over time or between different populations. The 95% confidence interval reflects the stability of the rates. A stable rate is one that would be close to the same value if the measurement were repeated, i.e., if the rate did not vary greatly from one year to the next. An unstable rate is one that would vary from one year to the next due to chance alone. Wider confidence intervals in relation to the rate indicate instability. Narrow confidence intervals indicate stability, and large fluctuations from year to year would not be expected. If differences are observed between stable rates (those with narrow confidence intervals), then it is likely that the differences represent true variations, rather than random fluctuations in the number of deaths. In general, if the 95% confidence intervals of two rates do not overlap, they are likely to be significantly different. For a description of the approach used to calculate confidence intervals for the estimates in this dataset, see page 94 of the March 28, 2000 *National Center for Vital Statistics Reports* publication at http://www.cdc.gov/nchs/data/nvsr/nvsr48/nvs48_03.pdf.

The source of annual population estimates by community area was a linear interpolation of counts from the 2000 and 2010 United States Census, using the method described in *Trend Analysis and Interpretation: Key Concepts and Methods for Maternal and Child Health Professionals* (see page 8 of http://mchb.hrsa.gov/publications/pdfs/trendanaylsis.pdf). To generate the denominator used in calculating the average annual tuberculosis case rate for 2007-2011, the 2010 census count was doubled (to cover the years 2010 and 2011) and added to the interpolated population counts for 2007, 2008, and 2009.

Benchmarks: Nationally, 3.6 confirmed new cases of tuberculosis per 100,000 population were reported to CDC in 2010. The Healthy People 2020 target is 1.0 new case per 100,000 population. (See http://www.healthypeople.gov/2020/topicsobjectives2020/objectiveslist.aspx?topicId=23, IID-29)

Disclaimers: The population counts used in the calculations are estimates, and this potential source of error should be taken into account when considering the precision of the indicators. Error can result from geocoding as a result of inaccurate or incomplete source data (e.g., the recording of a person's residential residence does not include "North" or "South") or discrepancies in the reference data that is used to match addresses to their associated geographies (e.g., a particular street segment is excluded or associated to corresponding geographies incorrectly). This potential source of error should be taken into account when considering the precision of the indicators. Rate and confidence interval estimates for the five-year periods in which fewer than 20 cases were reported are unreliable; this instability should also be considered when making comparisons. Methods and data sources may not be identical to those used in CDPH reports published prior to August 2012.

Data Owner: Tuberculosis Control Program, Chicago Department of Public Health (CDPH).

Time Period: 2007-2011

Frequency: Updated as new data becomes available.

Related Applications: N/A

Version History: In September 2012, Epidemiology and Public Health Informatics revised the method used to calculate community area population counts from the 2010 U.S. Census. The revised approach bases population on block-level counts rather than tract-level counts, and better accounts for the population distributions within tracts 8310 and 8439. This revision affects rate calculations for the Logan Square (22), West Town (24), Woodlawn (42), and South Shore (43) community areas. An updated version of the dataset was posted to the Open Data Portal on October 5, 2012 to reflect these changes.