On May 16, 2014 a revised version of this dataset was posted to the Open Data Portal. See **Version History** on page 4 of this document for details.

Title: Selected underlying causes of death in Chicago, 2006 – 2010

Brief Description: This dataset contains the cumulative number of deaths, average number of deaths annually, average annual crude and adjusted death rates with corresponding 95% confidence intervals, and average annual years of potential life lost per 100,000 residents aged 75 and younger due to selected causes of death, by Chicago community area, for the years 2006 – 2010. A ranking for each measure is also provided, with the highest value indicated with a ranking of 1.

Description: CDPH calculated the indicators using geocoded annual death certificate datasets supplied by the Illinois Department of Public Health (IDPH). When the ICD-10 codes listed below were recorded as the underlying cause, a death was classified as being due to the corresponding cause.*

Cause	ICD-10 codes
Alzheimer's disease	G30
Assault (homicide)	X85-Y09, Y87.1
Breast cancer in females	C50
Cancer (all sites)	C00-C97
Colorectal cancer	C18-C21
Coronary heart disease	I20-I25
Diabetes-related*	E10-E14
Firearm-related	W32-W34, X72-X74, X93-X95, Y22-Y24, Y35.0
Injury, unintentional	V01-X59, Y85-Y86
Liver disease and cirrhosis	K70, K73-K74
Lung cancer	C33-C34
Prostate cancer in males	C61
Stroke (cerebrovascular disease)	I60-I69
Suicide (intentional self-harm)	X60-X84, Y87.0

*A death was considered diabetes-related if a diabetes-associated ICD-10 code was recorded as an underlying cause or a contributing cause.

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. Approximately 1% of deaths per year (all causes) in the death certificate datasets are classified as Chicago residents but are not geocoded to a particular community area; these records are included in citywide counts and analyses only. Deaths of any cause classified as Chicago but for which the census tract of residence is located outside the city limits were excluded; this resulted in the exclusion of approximately one out of every 25,000 deaths. Age-adjustment was done using a slight modification of the direct method to the 2000 U.S. standard million population. (Specifically, using

standard distribution 1, the *less than 1 year old* and *1 – 4 years old* groups were combined.) Rates are expressed per 100,000 population. Rates are presented with the lower and upper confidence intervals computed at the 95 percent level using the Tiwari method. Years of potential life lost are calculated using an end-point age of 75 years and are expressed as the number of years per 100,000 population. (Death at age 74 is considered one year of potential life lost.) The source of age-specific 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). Indicators for Chicago as a whole are provided in the final row of the table.

Mortality measures

The **underlying cause of death** is the main disease, accident, or injury that caused the death. Death certificates are generally completed by a funeral director, physician, coroner, or medical examiner, and then are sent to a state's department of health or other agency. At the state agency, automated coding software is used to translate the submitted information into *International Classification of Diseases, Tenth Revision* (ICD-10) codes, and according to set algorithms, an underlying cause of death is determined. Other factors that played a role in the death but are not considered the main cause are classified as "contributing causes." For more information, see http://library.ahima.org/xpedio/groups/public/documents/ahima/bok1_043992.hcsp?dDocName=bok1_043992.

The **crude death rate** is the total number of deaths to residents in a specified geographic area (state, county, city, etc.) divided by the total population for the same geographic area for a specified time period. Crude death rates are not appropriate for comparison of different populations or geographic areas due to the significant impact of age in mortality data and differing age-distributions between populations.

The **age-adjusted death rate** is a death rate that controls for the effects of differences in population age distributions. The age-adjusted rate is hypothetical, and is useful **only** for comparing populations, either over time, by geographic area, by sex or by racial/ethnic subgroups. Age-adjusted rates do not reflect the absolute frequency of the event in a population in the way crude death rates do. The age-adjusted death rate is the weighted average of the age-specific death rates observed in a population, with the weights derived from the age distribution of an external population standard, such as the U.S. 2000 standard population. Different standard populations have different age distributions and the choice will affect the resulting age-adjusted rate. When comparing age-adjusted rates, ensure the rates were calculated using the same standard population. For more information, see http://www.cdc.gov/nchs/data/statnt/statnt20.pdf.

The years of potential life lost rate (YPLL rate) is a summary measure of years lost due to premature death per 100,000 population at or below age 75. In contrast to mortality measures, YPLL emphasizes the effect of premature mortality on a population. YPLL is the sum of the differences between a predetermined end point (e.g., age 75 or average life expectancy) and the ages of death for those who died before that end point, divided by the total population at or below that end point, and multiplied by 100,000. For example, people who die before age 75 are defined as having lost some

potential years of life. Because of the way in which YPLL is calculated, this measure gives more weight to a death the earlier it occurs. For instance, although most deaths occur at advanced ages, more years of potential life are lost for deaths among younger age groups, especially for certain causes, such as HIV/AIDS, accidents, homicide and suicide. For more information, see http://www.naphsis.org/Pages/StatisticalMeasuresandDefinitions.aspx.

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., deaths). 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, an age-adjusted mortality rate of 10 deaths 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 deaths per 100,000. Conversely, there is a 5 percent chance that the rate was lower than 8 or higher than 12.

In the case of mortality data, death rates are subject to random error, arising from random fluctuations in the number of deaths 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 significantly different. For a description of the approach used to calculate confidence intervals for the crude rate estimates in this dataset, see pages 26-28 of *Public Health Data: Our Silent Partner, Module 2* at http://www.cdc.gov/nchs/data/training/module2.pdf. For more information on the Tiwari method of confidence interval calculation for age-adjusted rates, see http://www.idph.state.il.us/cancer/11/county_rpt/County_Appendix_C_Formulas_for_Rates.pdf.

Benchmarks

Healthy People 2020 is a set of 10-year national objectives for improving the health of all Americans. Healthy People 2020 is comprised of approximately 600 objectives and over 1300 measures. Many of the objectives focus on interventions that are designed to reduce or eliminate illness, disability, and premature death among individuals and communities. Additional objectives are aimed to eliminate health disparities, address social determinants of health, improve access to quality health care, strengthen public health services, and improve the availability and dissemination of health-related information. For additional information on Healthy People 2020, see http://www.healthypeople.gov/2020/about/default.aspx.

Cause	Baseline	2020 Target
Alzheimer's disease	22.7	-----

Assault (homicide)	6.1	5.5
Breast cancer in females	22.9	20.6
Cancer (all sites)	178.4	160.6
Colorectal cancer	17.0	14.5
Coronary heart disease	126.0	100.8
Diabetes-related*	73.1	65.8
Firearm-related	10.2	9.2
Injury, unintentional	40.4	36.0
Liver disease and cirrhosis	9.1	8.2
Lung cancer	50.6	45.5
Prostate cancer in males	23.5	21.2
Stroke (cerebrovascular disease)	42.2	33.8
Suicide (intentional self-harm)	11.3	10.2

Disclaimers: IDPH specifically disclaims responsibility for any analysis, interpretations, or conclusions. When fewer than 20 deaths attributed to this cause during the period of study were recorded (cumulatively, not annual average), the rate, confidence interval, and years of potential life lost estimates are unreliable; this instability should be considered when making comparisons. 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. Methods and data sources may not be identical to those used in CDPH reports published prior to April 2012.

Data Owner: Epidemiology and Public Health Informatics, Chicago Department of Public Health (CDPH).

Time Period: 2006 - 2010

Frequency: Updated as new data becomes available

Related Applications: N/A

Version History: The version of the dataset posted on May 16, 2014 incorporates data from 2006 - 2010 and replaces the previous version, which covered 2005 - 2009.