Methodology of Projecting Cancer Incidence for the State of Illinois and Illinois Counties
(V25 June 2021)

Introduction

Projections of future cancer incidence in Illinois are calculated annually when incidence data are updated at the state level. Cancer incidence is projected for all sites combined and each of 22 specific sites for males and females of all races. For each of the state's 102 counties, cancer incidence projections are estimated for all sites combined, colon and rectum, and lung and bronchus for all races, both sexes, and for female breast (invasive), cervix (invasive), and prostate among all races. The state projection is for four future years ahead of the most recent data (e.g., with 2018 data as being most recent, the four future years are 2019, 2020, 2021, and 2022). The projection at the county level is for four future time periods, each consisting of five years (e.g., 2015-2019, 2016-2020, 2017-2021, and 2018-2022). Projections at both the state level and the county level are only reported for all races.

Methods

The age-specific rate method is utilized to project cancer incidence in Illinois. The basic rationale for this method is straightforward in that it involves the multiplication of the most recent age-sex-race-specific cancer incidence rates for a group by the age-, sex-, and race-specific population projections for that group. The method recognizes and relies on an assumption that population size and composition determine cancer incidence.

The age-, sex-, and race-specific population projection is obtained by modeling population growth as an exponential function of time. Specifically, the following model is constructed with population numbers supplied by the U.S. Census Bureau:

$N_t = N_0 . e^{r \bullet t}$ where N_t is population at year t.

By logarithm transformation, the model becomes $\log(N_t) = \log(N_0) + r \bullet t$, which can be estimated by linear regression. Note this model is the same as the growth function in Excel or Quattro Pro spreadsheets. The projection for new years is then made by extrapolation of the model. To ensure the model estimates are stable and capture enough information, the population data used for model construction must include the two most recent census years (e.g., 2000 and 2010), all years in between (e.g., 2001-2009), and all years since the last census year (e.g., 2011-2021).

Obviously, cancer projections can never be as accurate as actual cancer registrations. Revisions will vary due to the dynamic nature of the registry database and changing population projections. Because the numbers may vary considerably, particularly for less common cancers and for small counties, the use of these estimates to track year-to-year changes in cancer incidence is discouraged. For the purpose of tracking incidence changes, actual cancer incidence data collected by the Illinois State Cancer Registry should be used.

State Projections: The most recent Illinois annual cancer incidence rates for each sex are calculated for the standard 19 age groups. Population estimates by individual year for the same age-sex-race groups are available for Illinois from the U.S. census. It should be noted that population data for all races reflect the sum of estimates for Whites, Blacks, and others rather than an extrapolated value for the all races population estimates.

Expected incidence is calculated by multiplying the age-, sex-specific for all races cancer incidence rate by the population for that subgroup and then summing expected incidence for all 19 age-sex subgroups to arrive at a total expected cancer incidence estimate for males and females. This procedure is applied to cancer incidence for all sites combined and each of the 22 selected sites. Projected cancer incidence for these sites is calculated for each of the projected years (e.g., 2019, 2020, 2021, and 2022) by applying the same age-sex-race-specific rates to the extrapolated population estimates of that year.

County Projections: A slight modification of the procedures for state projections is used to project cancer incidence from all sites combined and selected sites for each of the state's 102 counties. At the county level, it is necessary to aggregate data collected for five-year periods in order to protect confidentiality and produce more stable, interpretable data that may be viewed with reasonable confidence. Therefore, cancer incidence data for five-year time periods (e.g., 2018-2022) are used for calculating county-level projections.

Age-sex-specific cancer incidence rates for all races are calculated for the 19 age groups. Expected cancer incidence is then calculated by multiplying the age-sex-specific cancer incidence rate for the most recent five-year period by the population estimate for the new period and for that age-sex specific group, and then summing expected incidence for all age and sex groups to arrive at a total expected cancer incidence estimate. This procedure was applied using county-level cancer incidence data among all races for cancers from all sites combined, colon and rectum, lung and bronchus for both sexes, female breast (invasive), cervix (invasive), and prostate. Note projected incidence for both sexes is based on combining separately projected sex-specific projections.

Projected cancer incidence for those respective sites are calculated for the five-year periods, each rolling ahead by one year, with a total of four five-year periods (e.g., 2015-2019, 2014-2018, 2017-2021, and 2018-2022). Like the state projection methodology, it is necessary to extrapolate county population estimates for one or more years involved in the five-year period. The county population estimates are individually extrapolated for each of the four future years ahead of the most recent data year (e.g., with 2018 data as being most recent, the four future years are 2019, 2020, 2021, and 2022), using the exponential growth function as previously described. Once individual population extrapolations are in place, population estimates for five-year periods are produced by simply aggregating appropriate year's totals for each of the age-sex-specific groups.

State Mortality Projections

The methodology behind the projection of cancer incidence at the state level is applied to the most current mortality rates to obtain projections for cancer mortality. Again, because the most current age-, sex-, and race-specific rates are applied to population estimates that are derived through model extrapolation, these cancer mortality projections, like cancer incidence projections, are less precise than if actual census bureau population estimates were used to calculate them and should be viewed as such. Further, because a fixed rate is applied to all future years, the projection reflects only changes associated with population size and not changes in risk factor distributions among populations (e.g., changes in cancer mortality rates).