Vermont Subcounty Cancer Maps

– Data Notes –

May 2024

Contents

[Overview](#)

[Data Sources](#)

[Cancer Incidence Analysis Geography](#)

[Calculations](#)

[Suggested Citation](#)

[Acknowledgements](#)

Overview

Cancer is a leading cause of death in Vermont, accounting for nearly one-quarter of all deaths in the state. Nearly four out of 10 Vermonters will develop cancer in their lifetime.

Vermont Cancer Maps are available for seven cancer types: Breast (Female); Colorectal; Lung; Bladder; Prostate; Melanoma; and Non-Hodgkin Lymphoma. These reports display observed cases and age-adjusted expected cases for sub-county geographic areas and provide complete Vermont statewide coverage. Calculated standardized incidence ratios (SIRs) and statistical significance compared to the state are displayed for each analysis area.

SIRs are used to represent the number of newly diagnosed cases of cancer for communities because they are easier to interpret than other measures. An SIR is a ratio of a community's actual cancer occurrence relative to its expected cancer occurrence, based on the state's cancer occurrence. The SIR measure corrects for factors that can vary among communities and affect cancer rates, such as age and size of a community's population.

The statewide benchmark SIR value is 1.00 because the numbers of observed and expected cases are equal. An analysis area with an SIR of 1.25 can be said to have 25% more cases than expected; another analysis area with an SIR of 0.75 can be said to have 25% fewer cases than expected.

Residential location at the time of diagnosis is only one of many possible risk factors for developing cancer. Many forms of cancer can be prevented through good dietary habits and avoiding tobacco. Early screening and detection for certain types of cancer can save lives.

For those instances in which a cancer is due to contact with a carcinogen (a cancer-causing agent), the disease does not develop immediately. Instead, there is often a long period (as much as 30

years) between the exposure to a carcinogen and medical diagnosis of cancer. Cancer typically develops gradually as a result of a complex mix of factors related to lifestyle choices, environment and genetics. Among adults, we know that nearly two-thirds of cancers diagnosed in the U.S. can be linked to preventable behaviors such as tobacco use, diet, obesity, physical inactivity, and UV radiation.

Cancer incidence calculations are based on a person's residence at the time cancer was diagnosed. If a person was exposed to a carcinogen when he or she lived in one location, then moved to different location and later became diagnosed with cancer, the person's residence at diagnosis would not be representative of the place where the person was exposed to the cancer-causing substance.

People with interest in particular cancer types are strongly encouraged to access additional information on the Vermont Department of Health website (http://www.healthvermont.gov) and from other authoritative sources, such as the Centers for Disease Control and Prevention (CDC) (www.cdc.gov/cancer) and the National Cancer Institute (NCI) (https://www.cancer.gov/about-cancer).

Data Sources

The Vermont Cancer Registry (VCR) is Vermont's statewide population-based cancer surveillance system. The registry collects information about all cancers and benign brain tumors diagnosed in Vermont (except non-melanoma skin cancers and non-invasive cervical cancer). The newly diagnosed cancer cases from the VCR are used in the numerator of incidence rate calculations.

The cancer types were coded using the International Classification of Disease for Oncology (*ICD-O*) coding system. Incidence site groupings are also based on *ICD-O-3 / WHO 2008 Definition* (https://seer.cancer.gov/siterecode/icdo3_dwhoheme/).

Each record of cancer diagnosis in the VCR includes the residential address at time of diagnosis along with corresponding latitude and longitude coordinates. The process of assigning coordinates to a street addresses, referred to as "geocoding," is not successful for all addresses. At times, it is not possible to assign coordinates to individual records because of a variety of potential problems such as incorrectly entered street numbers, street names, or town names. For purposes of these Vermont Cancer Maps, a limited number of cases are excluded from the analysis if they cannot be accurately assigned to an analysis area.

Census tracts were assigned to cancer records in the VCR based on the 2000 or 2010 Census, depending on diagnosis year. For diagnosis years 2001-2007, Census 2000 definitions were used; for diagnosis years 2008-2014, Census 2010 definitions were used. The change occurs at 2008, rather than 2010, based on a recommendation from North American Association of Central Cancer Registries (NAACCR) to assign Census Tract 2010 to the 2008-2009 diagnosis years.

Population data are used in conjunction with VCR case data to calculate standardized incidence ratios. Census demographic data, including population totals by age and sex for each of the individual geographic analysis areas and the State, are used in the denominator of the incidence rate calculations. The Census year used depends on the year of diagnosis being analyzed. For diagnosis years 2001-2007, Census 2000 populations were used; for diagnosis years 2008-2016, Census 2010 populations were used.

Cancer Incidence Analysis Geography

These Vermont Cancer Incidence Reports attempt to display cancer data at the smallest appropriate geographic scale. Geographic scales displayed in these reports include census tracts, combined census tracts, towns, paired towns, and groups of towns. The goals that guided the development of the sub-county identification (SCID) analysis geography included minimizing areas requiring data suppression (<6 cases) and minimizing the hi-low range of total populations across statewide analysis areas. In addition, initial testing focused on optimizing statistical power, which led to a decision to limit the total number of analysis areas to approximately 100 to provide a good balance of geographical specificity (areas as small as possible) vs. numerical sensitivity. In sparsely populated areas of the state, combining towns into larger areas increases the number of cases in the areas. This improves the numerical sensitivity (the chance of detecting true differences).

The 107 SCID areas used for Vermont SIR Cancer Maps include 89 areas that use town boundaries as the smallest geographical unit and 18 sub-town areas for select cities. In this set of 107 areas, the total population of individual analysis areas was designed to range from 3,500 to 10,000 – although a small number of exceptions fall outside this range.

Census tracts are displayed in many high-population cities and towns (Bennington, Colchester, Hartford, Rockingham, Rutland, South Burlington, and St. Johnsbury) where geocoding accuracy suffices and where census tract boundaries reflect substantially distinctive areas. In some cities (Burlington, Essex, Montpelier), tracts are combined into sub-areas such as the “South End” of Burlington or the “Essex Junction” portion of Essex Town. For other cities and towns with populations higher than 3,500, calculations are displayed for the undivided city or town area.

Towns are combined into logical groupings for more sparsely-populated areas of the state. These combined-town areas always consist of towns located in the same county. In certain unique situations, towns are paired (Hinesburg – St. George; Newport Town – Newport City; Vernon – Guilford).

When viewing a Vermont Cancer Map, it is possible to view town boundaries on the report map. This is intended to provide viewers with the ability to determine which analysis area any particular town is part of. Town labels accompany town boundaries when zoomed in on a single county.

Calculations

Vermont Cancer Maps display the following calculations for each analysis area: Age-Adjusted Expected Cases; Standardized Incidence Ratio (SIR); and “Statistical Comparison to State.”

Intermediary calculations not displayed in the reports include: cancer rate by age and sex for 18 age groupings; Byar’s Z and Byar’s P-values for each calculated SIR; and 2-stage False Discovery Rate (FDR) calculations for each SIR that correct for the problem of determining statistical significance when making multiple comparisons.

Below are calculation details relating to the calculation of expected cases:

Age-Adjusted Expected Cases are calculated based on 36 population classes (Male/Female for 18 age groupings). The 18 age groupings (in years) are: Under 5, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84, and Over 84. Before calculating expected cases for each analysis area, statewide rates are first calculated for each of the 36 population classes. The statewide rate is the number of cases in the state in all years divided by the weighted average population for one year. For example, the prostate cancer rate for males aged 60-64 equals the total number of prostate cancer cases in males 60-64 divided by the total

statewide population of males aged 60-64. The total expected cancer cases by cancer type are tallied across the 36 population classes for each analysis area by multiplying the total number of people within a class by the above-calculated rate for the class.

SIR values (observed cases/expected cases) are calculated for each analysis area, with the expected number of cases based on the statewide rates corrected for the local age structure.

It is important to determine how likely an SIR value is the result of random chance, especially when making many comparisons, such as comparing each of the sub-county analysis areas to the Statewide SIR at the same time. When analyzing a single geographic area, a P-value calculation typically implies a "confidence level" when testing for statistical significance. For example, a P-value of 0.10 would be characterized as having a "90% confidence level" that the SIR result is not the result of random chance. In other words, we are 90% certain that the difference between the area's SIR and the statewide SIR is not due to chance alone.

However, errors due to random chance are more likely when considering multiple areas simultaneously; therefore, the simple P-value methodology has been expanded in these Vermont Cancer Maps. VCR has elected to perform a statistical correction for this multiple comparisons problem by calculating a 2-stage False Discovery Rate (FDR) Yes/No determination of statistical significance.

The following description of the False Discovery Rate (FDR) method is listed on the http://en.wikipedia.org/wiki/Multiple_comparisons webpage: "…if a study is viewed as exploratory,… control of the false discovery rate is often preferred. The FDR, defined as the expected proportion of false positives among all significant tests, allows researchers to identify a set of "candidate positives," of which a high proportion are likely to be true. The false positives within the candidate set can then be identified in a follow-up study."

Below are calculation details relating to statistical significance testing for a 0.1 significance level (or a 90% confidence level):

First a Z-value is computed based on the observed and expected numbers of cases. Then the Z-value is converted to a P-value.

If the number of observed cases exceeds expected cases then Byar's Z = SQRT(9 x Observed) x (1-1/9/Observed - (Expected/Observed)^(1/3))

Otherwise, Byar's Z = SQRT(9 x (Observed+1)) x (1-1/9/(Observed+1) - (Expected/(Observed+1))^(1/3))

Byar's P = 2 x (1-1/(1+ EXP(-0.07056 x ABS(Byar's Z)^3 – 1.5976 x ABS(Byar's Z))))

Before calculating the 2-stage FDR testing process, all areas are ranked based on Byar's P-values from low to high.

FDR Stage 1: If (Byar's P < (Rank / Total Number of SCIDs x 0.1)) then "Yes", Otherwise "No".

FDR Stage 2: If (Byar's P < (Rank x 0.1 / (1 + 0.1) / (Total Number of SCIDs – (sum of FDR Stage 1 "Yes" results)) then "Yes", Otherwise "No".

Final significance testing algorithm: If FDR Stage 1 are all "No" then "No"; If FDR Stage 1 are all "Yes" then "Yes"; If FDR Stage 2 is "Yes" then "Yes"; Otherwise "No".

The “Statistical Comparison to State” has one of three possible descriptive categoric values that characterize the incidence of an analysis area as compared to the state as a whole. Each analysis area that is not suppressed (>5 observed cases) is displayed as “Statistically High;” “Statistically Low;” or “Not Different” compared to the state as a whole.

The following algorithm is used to assign the appropriate descriptive categoric value: If SIR > 1.0 and FDR significance = “Yes” then “Statistically Higher”; If SIR < 1.0 and FDR significance = “Yes” then “Statistically Lower”; Otherwise “Not Different”.

For more information on the problem of multiple comparisons, refer to authoritative statistical theory sources or online summaries such as that found at the above-mentioned Wikipedia page.

Vermont Cancer Registry (VCR) Background

All health care facilities and providers diagnosing or treating cancer in the State of Vermont are required to report cancer cases to the VCR, according to the Cancer Registry Law, 18 VSA §§ 151-157. The law requires that all information collected by the VCR be kept confidential and privileged. This specifically includes identifying information regarding individual patients, health care providers and health care facilities.

The VCR’s aims are: (1) to make the most accurate and complete cancer incidence data available to public health professionals, researchers and the public, in the timeliest way possible; and (2) to determine the incidence of cancer in the Vermont population, monitor cancer incidence and mortality trends among state residents, identify high risk populations, report findings to health care professionals and the public and contribute data for cancer prevention, control and treatment programs.

The Vermont Cancer Registry can be contacted at (802) 863-7644.

Suggested Citation

Vermont Department of Health, Vermont Cancer Maps and Environmental Public Health Tracking, 2024.

Acknowledgements

We acknowledge the Centers for Disease Control and Prevention, for its support of the Vermont Cancer Registry and the Vermont Environmental Public Health Tracking Program. The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Centers for Disease Control and Prevention.