

APPENDIX B: Definitions

1. Definitions for Cancer Incidence Data

The MCSS collects information on all microscopically-confirmed malignant and in situ tumors in Minnesota residents, as well as all microscopically-confirmed intracranial and spinal cord tumors. However, squamous and basal cell carcinomas of non-genital skin sites are not collected.

Data on in situ cancers are not included in this report; and data on nonmalignant tumors of the nervous system are not included in the counts of all cancers.

The primary site and morphologic (cell) type of cancers are coded according to the International Classification of Diseases for Oncology (ICD-O). Cancers included in this report were coded according to the ICD-O, 2nd Edition and then the site codes were translated into the codes used prior to 1992 (ICD-O, 1st Edition) for the purposes of grouping the diagnoses.

Definitions of the specific cancer sites for incidence data included in this report are:

Cancer Type	ICD-O-1 code
Oral Cavity	140.0 - 149.9 (excludes M-959 - M998_)
Esophagus	150.0 - 150.9 (excludes M-959_ - M998_)
Stomach	151.0 - 151.9 (excludes M-959_ - M998_)
Small Intestine	152.0 - 152.9 (excludes M-959_ - M998_)
Colon	153.0 - 153.9, 159.0 (excludes M-959_ - M998_)
Rectum	154.0, 154.1 (excludes M-959_ - M998_)
Liver	155.0, 155.1(excludes M-959_ - M998_)
Gallbladder	156.0 (excludes M-959_ - M998_)
Pancreas	157.0 - 157.9 (excludes M-959_ - M998_)
Larynx	161.0 - 161.9 (excludes M-959- M998_)
Lung	162.2 - 162.9 (excludes M-959_ - M998_)
Trachea and Pleura	162.0, 163.0 - 163.9, 164.2 - 165.9 (excludes M-959_ - M998_)
Bones and Joints	170.0 - 170.9 (excludes M-959_ - M998_)
Soft Tissues	164.1, 171.0 - 171.9 (excludes M-959_ - M998_)
Melanoma	173.0 - 173.9 (M-872_through M-878_)
Breast	174.0 - 174.9, 175.9 (excludes M-959_ - M998_)
Cervix	180.0 - 180.9 (excludes M-959_ - M998_)
Uterus	182.0 - 182.8 (excludes M-959_ - M998_)
Ovary	183.0 (excludes M-959_ - M998_)
Vulva	184.1 - 184.4 (excludes M-959_ - M998_)
Prostate	185.9 (excludes M-959_ - M998_)
Testis	186.0 - 186.9 (excludes M-959_ - M998_)
Bladder	188.0 - 188.9 (excludes M-959_ - M998_)
Kidney	189.0, 189.1 (excludes M-959_ - M998_)
Eye	190.0 - 190.9 (excludes M-959_ - M998_)
Brain	191.0 - 191.9 (malignant behavior; excludes M-953_ and M-959_ - M998_)
Benign Brain and CNS	191.0 - 192.3, 192.8 - 192.9 (benign or uncertain behavior)
Thyroid	193.9 (excludes M-959_ - M998_)
Hodgkin's Lymphoma	M-9650 through M9667
Non-Hodgkin's Lymphoma	M-959_, M-967_through M-970_, M-971_, M-9750
Multiple Myeloma	M-973_
Lymphocytic Leukemia	M-9820 through M-9826
Myelocytic Leukemia	M-9860 through M-9868
All Leukemias	M-9800 through M-9941, M-9951

2. Definitions for Cancer Mortality Data

Cancer mortality data have been obtained from death certificates for Minnesota residents. Death certificates are collected, coded, and computerized by the Minnesota Center for Health Statistics at the Minnesota Department of Health. The underlying causes of death on these death certificates are coded according to the Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death, Ninth Revision (ICD-9).

Definitions of the specific cancer sites for mortality data included in this report are:

Primary Site	ICD-9 code
Oral Cavity	140.0 - 149.9
Esophagus	150.0 - 150.9
Stomach	151.0 - 151.9
Small Intestine	152.0 - 152.9
Colon	153.0 - 153.9, 159.0
Rectum	154.0, 154.1
Liver	155.0 - 155.2
Gallbladder	156.0
Pancreas	157.0 - 157.9
Larynx	161.0 - 161.9
Lung	162.2 - 162.9
Trachea and Pleura	162.0, 163.0 - 163.9, 164.2 - 165.9
Bones and Joints	170.0 - 170.9
Soft Tissues	164.1, 171.0 - 171.9
Melanoma	172.0 - 172.9
Breast	174.0 - 174.9, 175
Cervix	180.0 - 180.9
Uterus	182.0 - 182.8
Ovary	183.0
Vulva	184.1 - 184.4
Prostate	185
Testis	186.0 - 186.9
Bladder	188.0 - 188.9
Kidney	189.0, 189.1
Eye	190.0 - 190.9
Brain	191.0-191.9
Thyroid	193
Hodgkin's Lymphoma	201.0-201.9
Non-Hodgkin's Lymphoma	200.0 - 200.8, 202.0 - 202.2, 202.8 - 202.9
Multiple Myeloma	203.0, 203.8
Lymphocytic Leukemia	204.0-204.9
Myelocytic Leukemia	205.0-205.9
All Leukemias	202.4, 203.1, 204.0-204.9, 205.0-205.9, 206.0-206.9, 207.0-207.8, 208.0-208.9

3. Minnesota Regional Definitions

Regions and their corresponding counties are as follows:

<u>Northwestern</u>			<u>Northeastern</u>	
Beltrami	Lake of the Woods	Pennington	Aitkin	Koochiching
Clearwater	Mahnomen	Polk	Carlton	Lake
Hubbard	Marshall	Red Lake	Cook	St. Louis
Kittson	Norman	Roseau	Itasca	

<u>West Central</u>		<u>Central</u>		
Becker	Pope	Benton	Kanabec	Stearns
Clay	Stevens	Cass	Mille Lacs	Todd
Douglas	Traverse	Chisago	Morrison	Wadena
Grant	Wilkin	Crow Wing	Pine	Wright
Otter Tail		Isanti	Sherburne	

<u>Southwestern</u>			<u>South Central</u>	
Big Stone	Lincoln	Pipestone	Blue Earth	Nicollet
Chippewa	Lyon	Redwood	Brown	Sibley
Cottonwood	McLeod	Renville	Faribault	Waseca
Jackson	Meeker	Rock	Le Sueur	Watonwan
Kandiyohi	Murray	Swift	Martin	
Lac Qui Parle	Nobles	Yellow Medicine		

<u>Southeastern</u>		<u>Metropolitan</u>	
Dodge	Olmsted	Anoka	Ramsey
Fillmore	Rice	Carver	Scott
Freeborn	Steele	Dakota	Washington
Goodhue	Wabasha	Hennepin	
Houston	Winona		
Mower			

APPENDIX C: Glossary

Age-Specific Rate: The rate of occurrence of a cancer for a specific age group (the number of cancers occurring during a specified period of time in a particular age group divided by the total number of individuals in the age group and time period).

Age-Standardized Rate (age-adjusted rate): Refers to a rate that has been adjusted to control for differences in age distribution between populations. It is a weighted average of age-specific rates, with the proportion of individuals in the corresponding age groups of the standard population functioning as the weights. The 1970 U.S. population is used as the standard in this report.

ALL: Acute lymphocytic leukemia.

Artifact: Any artificial product. In epidemiology, any observation that has been introduced by the methods used for data collection or data analysis.

Ascertainment: The collection of information; the process of finding desired information.

ATSDR: Agency for Toxic Substances and Disease Registries, an agency within the U.S. Centers for Disease Control and Prevention.

Benign: Not malignant, not likely to metastasize.

Biopsy: The removal and examination, usually microscopic, of tissue from the living body, performed to establish precise diagnosis.

Cancer: Diseases characterized by rapid, uncontrolled cell growth, with a tendency to spread throughout the body.

Cancer-Directed Treatment: As defined by SEER, this is therapy specifically undertaken to affect, control, change, remove, or destroy cancer tissue, or to induce remission in leukemias.

Cancer Registry: An ongoing system for the registration and follow-up of patients who develop cancer.

- **Hospital-Based Cancer Registry:** A cancer registry that uses hospital records as the primary data source for identification of cases.
- **Pathology-Based Cancer Registry:** A cancer registry that uses pathology laboratory records as the primary data source for identification of cases.
- **Population-Based Cancer Registry:** A cancer registry that attempts to collect information on at least 95 percent of the incident cancers occurring in the individuals residing within a defined geopolitical region.

Carcinoma: A malignant tumor of epithelial origin.

Case-Control Study: A study in which individuals with a particular condition such as cancer (referred to as cases) are selected for comparison with individuals in whom the condition is absent (controls). Cases and controls are compared with respect to past exposures, risk factors, or attributes thought to be relevant to the development of the condition under study.

Cell Type: See Histologic Type.

Central Nervous System (CNS): Brain, meninges, spinal cord and cranial nerves.

Clinically Diagnosed: Refers to cancers which are not histologically confirmed, but are instead diagnosed through other means—for example, through imaging procedures such as CT scans. Cancers which are only clinically diagnosed and have no microscopic confirmation are not collected by MCSS.

Cohort: In this report, cohort refers to a group of people with one or more common characteristics. Researchers follow cohorts over time, often to see if certain members are at higher risk than others of developing a disease like cancer.

Completeness: In the context of cancer surveillance, it is the ascertainment of all newly diagnosed cases of cancer occurring in Minnesota residents.

Death Clearance: A quality control activity that links MCSS' database of incident cancers with Minnesota cancer-related death certificates. Any death certificates that do not have a corresponding match in the MCSS database indicate a cancer that may have been missed. MCSS staff members follow up each of these cases to see if the cancer should have been included in the database.

Demographic Data: Descriptive information such as name, social security number, address, age, and sex, that is useful in identifying individuals or their geographical location of residence.

Epidemiology: The study of health conditions (e.g., cancers, injuries, etc.) by looking for patterns of occurrence by time, place, or person in the hopes of finding causes or identifying control measures for the condition.

Etiology: The study or theory of the causation of any disease; the sum of knowledge regarding causes.

Expected Number of Cases: The number of cases (of a cancer) expected in a given population in a given time period if the incident rates for that cancer were the same as the rates in a comparison population, adjusting for age differences of the two populations.

Histologic Type: “Histo” refers to tissue, and histologic type refers to the type of tissue in which a tumor originated, e.g., glandular tissue, connective tissue, etc.

Histologically Confirmed: Refers to a tumor of which at least a piece has been examined microscopically and diagnosed by a pathologist or other specialist.

Hospital-Based: See Cancer Registry.

In Situ: Preinvasive cancer; a cancer that is diagnosed before it penetrates too deeply.

Incidence Rate: The rate at which new events (in this case, cancers) occur in a population. It is usually expressed as a number per 100,000 persons per year.

Incident: A newly-diagnosed cancer from a defined population, within a specified period of time.

Invasive: The tendency to spread to adjacent healthy tissues. Technically, “invasive” means the carcinoma has penetrated the basement membranes and is close to blood vessels.

Lifetime Risk of Cancer: An approximate measure of the chance of developing cancer in an individual’s lifetime.

Malignant: Tending to become progressively worse, to spread, and invade other tissues.

MCHS: Minnesota Center for Health Statistics.

MCSS: Minnesota Cancer Surveillance System.

MDH: Minnesota Department of Health.

Metro (Metropolitan): In this report the metro area refers to Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington counties.

Minnesota Resident: Defines the population of individuals on whom cancer occurrence information is being collected by the MCSS.

Mortality Rate: A measure of the rate at which deaths occur in a population (the number of deaths occurring in a defined period of time divided by the total number of people in the population during that period of time).

NCI: National Cancer Institute.

Non-Metro (Non-metropolitan): In this report, non-Metro refers to all Minnesota counties except Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington.

Non-Identified File: A set of electronic records from which the identity of any one individual cannot be deduced.

Observed Number of Cases: The actual (also called crude) number of cases of a cancer recorded for a given population for a given time period.

Pathology: The branch of medicine that studies the essential nature of disease, especially the structural changes in tissues or organs associated with disease.

Pathology-Based: See Cancer Registry.

Population: All the inhabitants of a given area considered together; the number of inhabitants of a given area.

Population-Based: Pertaining to a population defined by geopolitical boundaries; this population is used as the denominator in calculating rates. For the MCSS, this is the State of Minnesota. (Also see Cancer Registry.)

Primary Site (Site): The place in the body where the cancer first arose.

Quality Control: The steps taken to avoid making errors and to find and correct errors before the data are added to the master database.

Rapid Ascertainment: The process by which cases are reported within a shorter time than through the routine process.

Record Linkage: The process of comparing two records from different sources, deciding if the records correspond to the same individual or entity, and then taking some action based on that decision.

Risk Factor: An attribute or exposure that is associated with an increased probability of developing a condition or disease, but does not necessarily imply cause and effect.

SEER (Surveillance, Epidemiology and End Results): An ongoing, population-based cancer surveillance system sponsored by the National Cancer Institute that monitors cancer incidence, treatment, and follow-up in nine or eleven U.S. regions comprising approximately 10 or 14 percent of the U.S. population depending on which years of data are examined.

Stage (of a tumor): Stage at diagnosis classifies how far a cancer has progressed, in order to determine the best course of treatment and to predict a patient's prognosis. Although there are various staging systems in use, two of the most well-known are:

- General Summary Stage (GSS), developed for the National Cancer Institute's SEER Program, which categorizes tumors as in situ, localized, regional, distant (see specific terms for further definitions); and
- The American Joint Committee on Cancer's TNM system, which incorporates information on the size of the tumor; which (if any) lymph nodes are affected; and whether the tumor has spread to distant organs, and then assigns a cancer-specific stage (e.g., Stage IIA breast cancer).

Certain cancers, such as leukemias, may be staged differently. Clinicians use prognostic factors specific to these cancers to determine the appropriate course of action.

Surveillance: The systematic collection, analysis, and interpretation of data pertaining to the occurrence of specific diseases (in this report, cancer).

- **Active Surveillance:** The reporters of disease are contacted at regular intervals and specifically asked about the occurrence of the disease under surveillance. This is considered the most ideal and complete form of surveillance.
- **Passive Surveillance:** Reporting of the disease in question is initiated by the reporting source.

Tumor: A mass resulting from the abnormal growth of cells. Tumors may either be benign (with little tendency to spread throughout the body) or malignant (with a tendency to spread throughout the body). Malignant tumors are synonymous with cancer.

Underlying Cause of Death: The disease or injury that initiated events resulting in death.

APPENDIX D: Statistical Methods

Estimated Annual Percent Change (EAPC)

The EAPC was calculated using the same method as employed by the NCI's SEER program. A regression line was fit to the natural logarithm of the rates (r) using the calendar year as the independent variable. That is, $y = mx + b$ where $y = \ln(r)$, x = the calendar year, and m is the slope of the line. The EAPC was estimated as $100(e^m - 1)$. The determination of whether the EAPC was different than zero was made by testing whether the slope of the regression line was statistically different than zero.

Regression

There are many models for creating a regression line that “best fits” empirical data. The purpose of modeling the data is to smooth out random variation from an underlying relationship and to enhance the parsimonious interpretation of that relationship.

Least square regression used in the report is one of the methods used to model data (e.g. cancer incidence rates as a function of calendar year). A straight line is estimated that minimizes the square of the difference between the observed and expected values. In this context, the best fitting straight line is the one that minimizes this difference. Once the characteristics of the best fitting line are determined, analytic parameters, such as slope and intercept required for specific estimates, can be easily defined.

Standard Error of Age Standardized Rates

Age-standardized rates are computed from weighted averages of the age-specific rates. The weights are traditionally calculated from the 1970 U.S. census as the proportion of the total census that the specific age group represents. Age-standardized rates are then considered age-adjusted in that differences in age distributions of two populations will not distort the comparison of the (directly) age-standardized rates.

The statistical inference whether the rates are different requires consideration of the variability (standard error) of the age-standardized rates. Keyfitz ([Human Biology 26:301-7, 1966](#)) developed estimates of the standard error using the Poisson probability distribution. The larger the population and the resultant number of cases, the smaller the standard error of the estimated rate.