

Analysis

What's on this page

Significance Testing

Throughout this report, the word “significant” refers to statistical significance at an alpha level of 0.05 for changes in trend or when comparing differences in rates or ratios. Non-significant changes in trend are described in this report as “stable.” In some instances (such as for survival statistics), statistical significance is assessed using a confidence interval, which accounts for variations and chance errors and represents the frequency with which the interval would contain the average measure 19 times out of 20 (or 95% of the time).

Probability of Developing or Dying From Cancer

The probability of developing or dying from cancer refers to the probability of a newborn child developing or dying from cancer at some point during their lifetime. Lifetime risk calculations are based on current incidence and mortality rates. Therefore, their calculation assumes the current rates within each age group will remain constant during the life of the newborn child.

The probability of developing or dying from cancer was calculated using DevCan software.^[1] The DevCan software program uses life-table methods based on cross-sectional incidence, mortality and population data for 18 age groups to compute the lifetime and age-conditional probabilities of developing or dying from cancer.

Incidence and Mortality

Counts

Incidence counts are the number of new cancer cases diagnosed in a population during a specific period. In this report, this refers to the number of new cancer diagnoses in a calendar year in Ontario. Complete death-cleared incidence data were available up to 2016 at the time of writing.

Mortality counts describe the number of deaths attributed to cancer during a specific period in a specific population. In this report, mortality refers to the number of deaths due to cancer in a calendar year in Ontario. For consistency, this report uses data for the same range of years for incidence and mortality (i.e., 1981 to 2016).

The childhood cancer incidence and mortality statistics included in this report are based on the cohort of children diagnosed with cancer between 1986 and 2016. Childhood cancer incidence statistics are presented for 1986 to 2016, while childhood cancer mortality statistics are presented for 1991 to 2016.

Rates

Incidence and mortality rates are the number of new cancer cases or deaths per 100,000 people in a population during a specific period. This is sometimes called the crude rate because it does not adjust for the age distribution of the population. Rates were calculated using SEER**Stat v.8.3.5.^[2]

For individuals diagnosed with cancer as children 0 to 14 years of age, incidence and mortality rates are the number of new cancer cases or deaths per 1,000,000 people in the population during a specific period. Childhood cancer rates were calculated using Statistical Analysis System (SAS) v.9.4.

Age-standardized rates

Age-standardized rates are weighted averages of age-specific rates using

a standard population. Age-standardized incidence rates (ASIR) and age-standardized mortality rates (ASMR) are adjusted for differences in the age structure of different populations. This permits comparisons of cancer incidence or mortality between different populations. These may be different segments of a population (e.g., different geography) or the same population at different times. Age-standardized rates give the rate that would have occurred if the population had the same age distribution as the standard population.

This report uses direct standardization, which produces artificial rates for the purpose of comparison only. The standard population used is based on the 2011 Canadian population (with 5-year age groups).

Trends in age-standardized rates

Incidence and mortality trends were determined using annual percent change (APC) and average annual percent change (AAPC), calculated using age-standardized rates. APC was used when examining short-term changes in trend. AAPC is a summary measure that describes the change in trend over a longer period with a single statistic.

APC and AAPC were determined using Joinpoint regression software (version 4.2.0.2 and version 4.7.0.0 for childhood cancer).^[3] Joinpoint regression uses piecewise regression to model the change in rates on the log scale. A statistical algorithm finds the optimal number and points in a trend (the joinpoints) where the trend changes. Joinpoint regression software also incorporates a Monte Carlo Permutation test to determine whether the change in trend is statistically significant.^[4]

In general, the model that Joinpoint software found to be the best fit was used. However, for some cancer types, other models were used to best describe the changes in trend for the data. This avoided capturing unstable trends resulting from relatively small numbers of cases in some stratified groups. A maximum of 3 joinpoints and minimum of 5 data

points per segment were allowed.

Projections

Projections of incidence and mortality for the years 2017 to 2020 and the years 2021 to 2030 were estimated using the Canproj projection package^[5] in R software. The Canproj package is a modified version of Nordpred Power 5 package,^[6] which is based on an age-period-cohort Poisson regression model. The Canproj package has enhancements that overcome difficulties in the standard Poisson model and improve projection accuracy.

The Ontario populations used for the cancer projections were from population projections up to 2041 derived by the Ontario Ministry of Finance. The methodology and assumptions for the projected populations can be found elsewhere.^[7]

Canproj consists of 3 sub-packages:

- Nordpred model (*adpcproj*: age-drift-period-cohort model)
- Age-Cohort model (*acproj*: age-cohort model)
- Hybrid model (*hybdproj*: age-only model)

Each sub-package can work independently for projections. Canproj has a built-in decision tree to help decide the appropriate projection model out of the 3 models. The package also can replace the Poisson distribution to a negative-binomial distribution when overdispersion is present in the data. Finally, Canproj tests the goodness-of-fit of the chosen model.

Projections for “All cancers” and each individual cancer site were estimated using the Canproj package.

Age-Drift-Period-Cohort Model (Nordpred)

The Norpred Power 5 model is represented as:

$$Case_{ap} \sim \text{Poisson}(\mu_{ap}),$$

$$R_{ap} = \frac{\mu_{ap}}{n_{ap}} = (A_a + P_p + C_c + Dp)^5$$

Norpred equation description

where the symbols represent the following:

- R_{ap} is the incidence rate in age group a in calendar period p .
- μ_{ap} is the mean count of cases in age group a in calendar period p .
- n_{ap} case is the size of the corresponding population.
- A_a is the age component for age group a .
- D is the common linear drift parameter.
- P_p is the non-linear period component of period p .
- C_c is the non-linear cohort component of cohort c .
- p is the calendar period,

Cohorts were calculated as:

with A equaling the total number of age groups (18)

Age-Cohort Model (acproj Model)

Age-cohort model is a reduced form of the Nordpred model selected by Canproj when sparse data exist in the youngest and oldest birth cohorts. Due to sparseness of the data at both extreme cohorts, the remaining cohorts with complete observations are set as reference when age and cohort effects are being estimated.^[5] The expression for the age-cohort model is:

$$Case_{ap} \sim \text{Poisson}(\mu_{ap}),$$

$$R_{ap} = \frac{\mu_{ap}}{n_{ap}} = (A_a + C_c)^5, \quad c = A + p - a,$$

$$p = 1, 2, \dots, P, \quad c = 1, 2, \dots, C \text{ where } C = A + P - 1$$

Age-cohort equation description

Hybrid Models: Age-Only, Common Trend, and Age-Specific Trend Models

Hybrid models are of 3 types: age-only, common trend and age-specific. The models use a combination of 3 methods: average, joinpoint type and, whenever the cohort effects are not significant, Poisson regression methods.^[4,5] The hybrid functions first compare the common trend model with the age-specific model using a chi-square test in the age groups where data exists for the entire periods. When there is no significant difference between the models, the common trend is preferred; otherwise, it is age-specific.

The incidence, death and population data were classified by year of diagnosis, year of death and sex, and grouped by 5-year age groups (0 to 4, 5 to 9, and so on, to 85 and older). For incidence projections, cases meeting the IARC/IACR multiple primary rules from 1981 to 2016 were projected. These were later converted for the period 2010 to 2016 for SEER multiple primary rules, by applying an inflation factor based on the age-specific increase in multiple primary cancers. Projections for all cancers combined were estimated based on the sum of all data from the 23 cancer sites in this report.

Mortality projections were also made with a Canproj package using cancer deaths from 1981 to 2016 divided into 5-year age groups and calendar

year. To get incidence and mortality projections for all cancers combined, projections were calculated by sex and then summed. This is because the projections based only on the data for all cancers combined are not equal to the sum of the projections for both sexes. The lists of models used for all cancers combined and for each individual cancer site by sex are in Table A.6 for incidence projections and Table A.7 for mortality projections.

Table A.6 Canproj models for cancer incidence projections, by cancer type and sex, Ontario

Cancer type	Sex	
	Males	Females
All cancers	adpcproj (NB)	hybdproj(NBags)
Bladder	adpcproj (P)	adpcproj (P)
Brain	acproj (P)	adpcproj (P)
Breast (female)	n/a	adpcproj (NB)
Cervix	n/a	adpcproj (P)
Colorectal	adpcproj (NB)	adpcproj (NB)
Esophagus	hybdproj(ComT)	hybdproj(ComT)
Hodgkin lymphoma	hybdproj(NBags)	hybdproj(Avg)
Kidney	adpcproj (P)	adpcproj (P)
Larynx	adpcproj (P)	adpcproj (P)
Leukemia	hybdproj(NBags)	hybdproj(NBags)
Liver	adpcproj (NB)	adpcproj (P)
Lung	adpcproj (NB)	adpcproj (NB)
Melanoma	adpcproj (NB)	adpcproj (NB)
Myeloma	hybdproj(ComT)	hybdproj(ComT)
Non-Hodgkin lymphoma	adpcproj (NB)	adpcproj (P)
Oral cavity & pharynx	adpcproj (P)	adpcproj (P)
Ovary	n/a	adpcproj (P)
Pancreas	hybdproj(ComT)	hybdproj(Ags)
Prostate	adpcproj (NB) [†]	n/a

Stomach	adpcproj (NB)	adpcproj (P)
Testis	acproj(P)	n/a
Thyroid	adpcproj (NB)	adpcproj (NB)
Uterus	n/a	adpcproj (P)

Abbreviations:

- acproj(P) means age-cohort model with Poisson distribution
- acproj(NB) means age-cohort model with negative-binomial distribution
- adpcproj (NB) means Norpred model with negative-binomial distribution
- adpcproj (P) means Norpred model with Poisson distribution
- hybdproj(NBags) means hybrid model with age-specific and negative-binomial distribution
- hybdproj(Ags) means hybrid model with age-specific and Poisson distribution
- hybdproj(ComT) means hybrid model with common-trend
- hybdproj(Avg) means hybrid model with average method
- n/a means "not applicable"

† Five periods were used to estimate projected incidence. Projected estimates used data starting from the 35 to 39 year age group.

Table A.7 Canproj models for cancer mortality projections, by cancer type and sex, Ontario

Cancer type	Sex	
	Males	Females
All cancers	adpcproj (NB)	adpcproj (NB)
Bladder	hybdproj(Ags)	hybdproj(ComT)
Brain	adpcproj (P)	adpcproj (NB)
Breast (female)	n/a	adpcproj (P)
Cervix	n/a	adpcproj (P)

Colorectal	adpcproj (NB)	adpcproj (P)
Esophagus	hybdproj(NBags)	hybdproj(ComT)
Hodgkin lymphoma	hybdproj(ComT)	hybdproj(ComT)
Kidney	acproj (P)	adpcproj (P)
Larynx	hybdproj(Ags)	adpcproj (P)
Leukemia	adpcproj (P)	hybdproj(Ags)
Liver	adpcproj (P)	adpcproj (P)
Lung	adpcproj (NB)	adpcproj (P)
Melanoma	adpcproj (P)	adpcproj (P)
Myeloma	hybdproj(ComT)	acproj(NB)
Non-Hodgkin lymphoma	adpcproj (P)	adpcproj (P)
Oral cavity & pharynx	hybdproj(ComT)	hybdproj (Ags)
Ovary	n/a	adpcproj (P)
Pancreas	hybdproj(Avg)	adpcproj (P)
Prostate	adpcproj (NB) [†]	n/a
Stomach	adpcproj (P)	adpcproj (P)
Testis	hybdproj(ComT)	n/a
Thyroid	hybdproj(ComT)	hybdproj(ComT)
Uterus	n/a	adpcproj (P)

Abbreviations:

- acproj(P) means age-cohort model with Poisson distribution
- acproj(NB) means age-cohort model with negative-binomial distribution
- adpcproj (NB) means Norpred model with negative-binomial distribution
- adpcproj (P) means Norpred model with Poisson distribution
- hybdproj(NBags) means hybrid model with age-specific and negative-binomial distribution
- hybdproj(Ags) means hybrid model with age-specific and Poisson

distribution

- hybdproj(ComT) means hybrid model with common-trend
- hybdproj(Avg) means hybrid model with average method
- n/a means “not applicable”

† Six periods were used to estimate projected incidence. Projected estimates used data starting from the 45 to 49 year age group.

Survival

Relative survival ratios (RSRs) are estimated by comparing the survival of people with cancer with the expected survival for the general population of Ontarians of the same age and sex during the same period. Relative survival shows the extent to which a diagnosis of cancer shortens a life span. The RSR is usually expressed as a percentage. The closer the value is to 100%, the more similar the survival pattern is to the general population.

Survival analyses were based on first primary cancers. RSR are provided for cases diagnosed in people between 15 and 99 years of age. Cases were excluded from the survival analyses in the following conditions:

- The age of the person was unknown.
- They were diagnosed on the basis of an autopsy only.
- The date of diagnosis and date of death were the same (i.e., death certificate only [DCO] cases where the diagnosis happened at or following death).

See Table A.1 in [Data Sources](#) for details on DCO cases.

Relative and conditional survival analyses were performed using SEER**Stat software (version 8.3.5).^[2] Expected survival proportions were derived using the Ederer II approach^[8] from provincial life tables produced by Statistics Canada. Life tables currently available for calculating

expected survival may not completely reflect all factors contributing to variation in all-cause mortality, such as smoking. Take this into account when interpreting the estimates.

RSRs were estimated by the cohort method when complete follow-up data after diagnosis (e.g., at least 5 years of follow-up to estimate a 5-year ratio) were available. For recently diagnosed cases whose complete follow-up data were not available, the estimates were computed using the period method. Period analysis uses the survival experience of people in a recent interval to estimate survival.^[9] The period method, modeled after period life tables, allows for more up-to-date estimates because it means analysis does not have wait for data on the full follow-up period (e.g., 5 years for a 5-year ratio). Interpret comparisons between cohort and period RSRs with caution because of the 2 different methods used to derive the respective ratios.

RSRs were age-standardized by weighting with the International Cancer Survival Standard (ICSS) weights.^[10] (See Table A.8 for details on weightings.)

Table A.8 International Cancer Survival Standards used for standardizing relative survival ratios, by cancer type and age group

Age groups (years)	Weightings	Cancer types
15 - 44, 45 - 54, 55 - 64, 65 - 74, 75 - 100	60, 10, 10, 10, 10	Testis, Hodgkin lymphoma, acute lymphatic leukemia
15 - 44, 45 - 54, 55 - 64, 65 - 74, 75 - 100	28, 17, 21, 20, 14	Nasopharynx, soft tissues, melanoma, cervix uteri, brain, thyroid gland, bone
15 - 44, 45 - 54, 55 - 64, 65 - 74, 75 - 100	7, 12, 23, 29, 29	All other cancer types except prostate
15 - 54, 55 - 64, 65 - 74, 75 - 84, 85 - 100	19, 23, 29, 23, 6	Prostate

For children between 0 to 14 years of age, 5-year overall survival proportions were estimated based on cancers diagnosed between 1987 and 2016. Overall survival is an estimate of the probability of surviving all

causes of death for a specified interval following a diagnosis of cancer. Unlike relative survival ratios, overall survival proportions do not adjust for the expected survival of the general population in Ontario of the same age and sex during the same period. Overall survival proportions were calculated using the cohort method when complete follow-up data were available (for cases diagnosed historically, 1987 to 2011) and the period method when complete follow-up data were not available (for cases recently diagnosed, 2012 to 2016). Patients were followed until date of death or December 31, 2018 (whichever occurred first).

Prevalence

Prevalence analyses were performed using SEER*Stat software (version 8.3.5).^[2] This report provides person-based limited-duration prevalence; that is, the number of people diagnosed with malignant cancer over a specific period (e.g., 2 years, 5 years or 10 years) who were still alive on the index date. The chosen index date was January 1, 2017.

Only the first primary cancer was included in the prevalence count for all cancers combined, but for individual cancer types, all cases of a given cancer type per individual were counted. For example, a person with a first primary of prostate cancer and a second primary of colorectal cancer would be included once in the prevalence count for all cancers but twice in the individual cancer type counts (i.e., once in the prostate prevalence count and once in the colorectal prevalence count).