Methodology for the Survival Estimates

Inclusion/Exclusion Criteria

- Cancer cases are classified according to the *International Classification of Diseases for Oncology - Third Edition (ICDO-3)*
- Disease sites are based on the Canadian Cancer Statistics (CCS) groupings with colorectal separated out into colon (C18 and C19) and rectal (C20). Refer to http://www.bccancer.bc.ca/statistics-and-reports-site/Documents/Cancer%20Definitions.pdf.

Inclusions:

- BC residents at time of diagnosis
- Adult cases only *(age at time of diagnosis between 15 and 99 years, except for conditional survival rates includes age at time of diagnosis between 15 and 74 year)*
- Invasive/malignant cancer cases (includes in-situ bladder)
- All primaries rather than simply first primaries according to the IARC 2004 rules*
- Only the earliest diagnosis for residual multiple primary*

Exclusions:

- Unknown gender or date of birth
- Non-melanoma skin cancers
- Cancers that cannot be classified into one of the CCS cancer groupings
- Diagnosis date precedes date of birth
- In situ or occult stage when calculating survival rates by stage
- Death precedes date of diagnosis
- Death certificate only (DCO) *cases***
- Autopsy only cases with zero survival time**

Multiple Primaries

Patients can have more than one type of cancer diagnosis and can have multiple diagnoses of the same site. The use of IARC 2004 multiple primary rules were applied in order to remove the most subsequent primaries of the same site. However, even with using the IARC rules some subsequent primaries of the same site remain as IARC rules allow for second primaries of the same site if the histology codes are sufficiently different. Inclusion of the same patient multiple times within the same survival analysis goes against the assumption of independence of observations. Therefore, if after applying the IARC rule a patient still has more than one

diagnosis within the same disease site, the patient's earliest diagnosis is included. For example, if after the IARC rule application a patient had a rectal and a colon diagnosis, only the earlier of the two would be included when calculating survival from colorectal cancer. However, the rectal case would be included for survival from rectal cancer and the colon case would be included for survival from colon cancer. Thus, patients can contribute to the survival analysis for more than one cancer site but not multiple times within a cancer site.

IARC rules do not take cancer stage into account. When calculating survival for disease sites by stage, for cancers diagnosed within 30 days of each other (synchronous tumours), the earliest cancer is included, however, the stage of the retained cancer was replaced by the highest stage among all corresponding synchronous tumours. For cancers diagnosed greater than 30 days apart, the earlier cancer is included, and the stage associated with that earlier cancer is used.

***DCO and Autopsy only cases**

DCO cases will have a survival of zero days since these cases are identified from the death certificate. These cases are excluded as they generally did not benefit from any medical interventions and hence they do not reflect anything meaningful about the success of the health care system. Some cases confirmed at autopsy can have a diagnosis that precedes the death date. These cases are included but if the diagnosis date is the same as the death date then these cases are excluded due to similar reasons as per DCO cases.

Calculation of Survival Estimates

Relative survival rates were calculated as the ratio of the observed survival rate after a cancer diagnosis to the expected survival rate of the general population. Relative survival basically measures the excess mortality associated with a cancer diagnosis. Observed survival is calculated from cancer diagnosis and death information available in the BC Cancer Registry. Expected survival is derived from life tables for the entire BC population and can be thought of as being calculated for a cohort of patients from the general population matched to cancer patients by age, sex, and period. The life tables used to calculate relative survival for these reports were based on British Columbia mortality rates, published by Statistics Canada4 [1]. If the life table for the current year is not available, the latest year life table is used with the assumption there would be minimal change over this short period. The Ederer II method [2], which considers matched individuals to be at risk until the corresponding cancer patient dies or is censored, was used to estimate expected survival.

A patient's hazard of death after a cancer diagnosis changes over time, and more rapidly so closer to the diagnosis [3]. Therefore the observed and expected probability of surviving was broken down and calculated for specific intervals of follow-up. Due to the changing hazard, smaller intervals were used earlier on in follow-up. One month intervals were used for up to 6 months following diagnosis, then every 3 months for 6 months to 2 years, then every 6 months for 2 to 5 years, and 1 year intervals after that. The hazard transformation approach [4] was used to calculate these interval specific survivals, and the Delta method [5, 6] was used to compute the variance of the cumulative hazard.

There are several approaches to calculating cancer survival. The cohort based approach is commonly used. Survival estimates derived from this method are based on patients diagnosed a number of years ago and followed up for the desired period after diagnosis. For example, if there is complete data on cancer incidence and mortality follow-up up to the year 2018, a cohort approach to calculate 5-year survival will use patients diagnosed in 2013. If the cancer type has small incidence, a number of years may need to be combined so it could include patients diagnosed in, say, 2010-2013. At least 5 years follow-up of those patients are then used to estimate the 5-year survival. Thus the survival estimates derived from this cohort is essentially reflecting the level of care from 5 to 8 years ago and may not reflect the survival one would expect for patients diagnosed in more recent years. A better prediction of the 5-year survival for cancer patients diagnosed in more recent years can be obtained using the 'period' method. The period method approach uses the most recent periods of data to estimate the relevant parts of the survival curve. Assume the most recent period for which the data is available is 2018 and a period estimate of 5-year survival is to be derived for 2018 period, then all observations are left truncated at the beginning of 2018 in addition to being censored at the end of 2018. The 5-year period estimate of survival would be obtained from patients diagnosed in 2013–2018 for whom some proportion of 5-year follow-up might have fallen in the 2013–

2018 period. With this approach different parts of the survival function are derived from patients diagnosed in various calendar years. Survival during the first year following diagnosis would be estimated for patients diagnosed in 2017-2018, survival during the second year following diagnosis would be estimated for patients diagnosed in 2016-2017, and so on, until survival experience during the fifth year following diagnosis which would be obtained for patients diagnosed in 2013-2014. These conditional survival probabilities are then combined in the usual way to generate 5-year cumulative survival estimates for the 2018 period. It has been demonstrated this method provides better predictions of the survival of recently diagnosed patients compared to the cohort method [7] and is used more frequently in the recent years[8]. Therefore period method has been utilized for our most recent periods and for consistency purposes of our trend plots, the period method has also been used for earlier periods even though complete follow-up on patients diagnosed in those earlier periods are available.

In order to determine what cases are included in the period approach, the length of the period needs to be specified. The length can vary for each cancer type depending on the cancer incidence. The smaller the incidence, the longer length of period is used in order to obtain more stable survival estimates. The period for the cancer types used in these survival reports were as follows:

- 1-year period window: All cancers, Breast, Colorectal, Lung, Prostate, Colon
- 2-year period window: All other cancers, Bladder, Melanoma (skin), Non-Hodgkin Lymphoma
- 4-year period window: Body of Uterus, Brain, Cervix, Esophagus, Hodgkin Lymphoma, Kidney, Larynx, Leukemia, Liver, Multiple Myeloma, Oral, Ovary, Pancreas, Stomach, Testis, Thyroid, Rectal

A 4-year period window was also used for all cancer types for the calculation of relative survival by stage.

Relative survival rates in these reports have been age adjusted. Age specific relative survival rates were estimated for predefined age groups first and then the International Cancer Survival Standard (ICSS) [9] weights were applied in order to obtain the age-standardised relative survival rates. The ICSS standard used for each specific disease site is listed in Table 1. The weights for the age groups used for each standard are shown in Table 2.

Table 1: ICSS standard weights applied to each disease site

Standard	Cancer Site
1 (For cancer sites with increasing incidence by age	All cancers, All other cancers, Bladder, Body of Uterus, Breast, Colorectal, Esophagus, Kidney, Larynx, Leukemia, Liver, Lung, Melanoma (skin), Multiple Myeloma, Non-Hodgkin Lymphoma, Oral, Ovary, Pancreas, Stomach, Thyroid, Prostate
2 (For cancer sites with broadly	Cervix, Brain

constant incidence by age	
3 (For cancer sites that mainly affect young adults	Testis, Hodgkin Lymphoma

Table 2: Age Group weights for the 3 ICSS standards

Age Group	Standard 1	Standard 2	Standard 3
15—44	0.07	0.28	0.6
45—54	0.12	0.17	0.1
55—64	0.23	0.21	0.1
65—74	0.29	0.2	0.1
75+	0.29	0.14	0.1

Age Group	Prostate only
15—54	0.19
55—64	0.23
65—74	0.29
75—84	0.23
85+	0.06

Survival time was calculated in days, and 1 day was added to the survival time for non DCO and autopsy only cases if the survival time is 0. If the diagnosis date is the first day of the year then 1 day is added to the death date, otherwise 1 day is subtracted from the diagnosis date.

Standard errors illustrate the degree of variability associated with a survival estimate. Large standard errors indicate high variability and interpretation needs to be made with caution. Therefore, only survival estimates with a standard error less than 0.05 were presented. If the estimated standard error was between 0.05 and 0.1 the 95% confidence interval for the survival rate was provided, if the standard error was greater than 0.1 the survival rate was suppressed [10].

Conditional relative survival rates were estimated as the ratio of observed survival to the expected survival after having survived for a certain length of time after diagnosis, for example, 5 more years survival after having survived 2 years since initial diagnosis. The same methods used to compute relative survival were used to compute conditional relative survival except patients were restricted to those that had survived a certain length of time since diagnosis. In addition, cancer patients aged 75+ were excluded due to long term survival rates not being as stable and with them more likely to have died from causes other than cancer [10, 11].

Smoothed Trend Plots

Smoothed curves were fit to the long-term trends of 1-, 3- and 5-year relative survival estimates in order to reduce the effect of year-to-year random fluctuation. The smoothing process was fulfilled using the LOESS procedure in SAS. The default golden section search of the LOESS procedure was used to locally minimize the bias corrected Akaike information criteria (AICC)

over the range (0,1] and linear interpolating polynomials was utilized for blending local polynomial fits at the vertices of k-dimensional tree cells.

If the relative survival was not estimable for the first or last period of the trend plot, the relative survival estimate of the next or the previous period, respectively, was used. If, for any other period on the trend plot the relative survival was not estimable then the LOESS predicted value was used.

References:

[1]. Statistics Canada. Life Tables, Canada, Provinces and Territories (84-537-X) [Web resource]. Ottawa, ON: Statistics Canada;[Available at: http://www5.statcan.gc.ca/olc-cel/olc.action?objId=84-537-X&objType=2&lang=en&limit=0; cited5 June 2018]

[2]. Ederer, F., and Heise, H. (1959), "Instructions to Ibm 650 Programmers in Processing Survival Computations," Technical, End Results Evaluation Section, National Cancer Institute.

[3]. Woods, Laura M., et al. "Full dates (day, month, year) should be used in population - based cancer survival studies." International Journal of Cancer131.7 (2012): E1120-E1124.

[4]. Dickman, Paul W., Enzo Coviello, and M. Hills. "Estimating and modeling relative survival." Stata J 15.1 (2015): 186-215.

[5]. Breslow, Norman E., and Nicholas E. Day. Statistical methods in cancer research. Ed. W. Davis. Vol. 2. Lyon: International Agency for Research on Cancer, 1987.

[6]. Oehlert, Gary W. “A Note on the Delta Method.” The American Statistician, vol. 46, no. 1, 1992, pp. 27–29. www.jstor.org/stable/2684406.

[7]. Ellison LF. An empirical evaluation of period survival analysis using data from the Canadian Cancer Registry. Annals of Epidemiology 2006; 16: 191-6.

[8]. Ellison, Larry F. "Estimating relative survival for cancer: An analysis of bias introduced by outdated life tables." *Health reports* 25.2 (2014): 13.

[9]. Corazziari, Isabella, Mike Quinn, and Riccardo Capocaccia. "Standard cancer patient population for age standardising survival ratios." European Journal of Cancer 40.15 (2004): 2307-2316.

[10]. BRYANT, H. et al. Conditional survival in Canada: adjusting patient prognosis over time. Current Oncology, [S.l.], v. 19, n. 4, p. 222-224, aug. 2012. ISSN 1718-7729. Available at: <http://www.current-oncology.com/index.php/oncology/article/view/1148/994>. Date accessed: 12 may 2016. doi:http://dx.doi.org/10.3747/co.19.1148.

[11]. Janssen-Heijnen, Maryska LG, et al. "Clinical relevance of conditional survival of cancer patients in Europe: age-specific analyses of 13 cancers." Journal of Clinical Oncology 28.15 (2010): 2520-2528.