BC Cancer Agency
CARE + RESEARCH
An agency of the Provincial Health Services Authority

Re: Differences in Cancer Statistics and Information from Different Sources of Cancer Data

BC cancer statistics can be obtained from many sources including the BC Cancer Agency website, the Canadian Cancer Statistics annual publication, the Statistics Canada on-line data tools and directly by request to BCCA Data Requests. Although the original source of the data for all of these statistics are the BC Cancer Registry, the actual statistics made available from these sources generally vary leading to frequent questions as to the reason for this variation in statistics. The purpose of this note is to provide some description of known sources of this variation in BC cancer statistics. This front page provides a high-level summary of the main differences and subsequent pages provide some additional technical detail and in-depth explanations for those interested.

Main known sources of observed differences in BC cancer statistics:

- *Differences in the way cancers are counted*: national reports and comparative reports of provincial data often convert cancer registry data from different jurisdictions into a common cancer coding system. This system is a common set of rules around how cancers are counted. The system used for the national and pan-Canadian reports generally results in fewer numbers of cases due to differences in how this system counts multiple cancers of the same type for a given person. It is necessary to convert data into this common coding system to make statistics comparable across the country as not all regions have historically reported and registered cancers in the same way.

- *Population Estimates and Projections*: certain statistics produced from different sources of BC population information will vary. In particular, long-term projections of new cancer cases and deaths that use different long-term projections for the population of a region can result in dramatically different projections. Population forecasts are frequently updated by regional government agencies as more data on immigration, economy and other drivers of population growth are made available.

- *Data Currency*: the BC Cancer Agency generally uses and publishes the most recent data we have available. Other national reports however generally have to wait until data have been submitted to and then processed and distributed by the national Canadian Cancer Registry team within Statistics Canada. As such data from the BC Cancer Agency are generally more current and up to date.

- *Population Standard*: Age-standardized cancer incidence and mortality rates produced from different population standards can result in dramatically different statistics. As such, when comparing cancer rates, attention must be paid to the standard used and rates calculated from different standards should not be compared.

If you have further questions about this issue, please email: datareq@bccancer.bc.ca

Further Technical Explanations for Differences in BC Cancer Statistics across Data Sources:

There are several major sources of cancer information for British Columbia that routinely accessed by those within the health care system and the research community. Some of the more common sources include the national Canadian Cancer Statistics annual publication, the BC Cancer Agency's Cancer Statistics section, the Statistics Canada CANSIM tables, the International Agency for Research on Cancer's Cancer Incidence in Five Continents publications, or by request directly to the BC Cancer Agency's Data Requests. Consumers of these data often note that the various sources provide somewhat different results and questions related to why these sources provide different numbers when similar queries are made are frequently received by the BC Cancer Agency. This note is meant to provide some background information about why these different data sources might differ for different types of summaries.

Population Estimates and Projections

To calculate cancer rates, some information about the size and demographic make-up of the population of British Columbia is required. There are two main sources of British Columbia population information that are generally used: those produced by BC Stats in our province or those produced by Statistics Canada. Population estimates are updated regularly and thus different sources can produce different results. Generally national reports or reports that are pan-Canadian will use population information from Statistics Canada whereas provincial cancer statistics or information provided by the BC Cancer Agency will use estimates provided by our provincial statistical agency BC Stats. When comparing reports or statistics, one should examine the source of any population information and the version used in the calculations. Generally, for current or historical cancer incidence or mortality rate calculations the impact of using provincial vs federal population estimates is not significant.

One area where population versions and sources can have a significant impact is with long-term cancer projections. Cancer incidence and mortality projections are produced by extrapolating recent trends in cancer rates and applying these to future population projections. When long-term forecasts for populations are updated, the future population totals and demographic make-up can change quite significantly. Thus, using two different sets of long-term population projections (e.g. the most recent version vs a version made several years ago) can produce very different future cancer case/death totals. Thus it is important to note the source of the population forecast when comparing different projections.

Coding rules

Perhaps one of the more significant sources of variation between different cancer statistics relates to the way cancers are coded and counted. Historically, not all jurisdictions used the same approach to counting new cancers within their regional cancer registries and this has led to some challenges when trying to compare different jurisdictions. There are generally two "coding rules" that have existed in Canada for counting and coding new cancers: the "IARC" rules and the "SEER" rules. The main difference between these coding systems that leads to differences in statistics is the way they count multiple tumours of the same kind for a given patient. For example, under the IARC rules, second or subsequent cancers of the same kind are less often counted as new cancers (e.g. if someone were to be diagnosed

with two breast cancers in their life). However under the SEER rules, multiple cancers at the same body site are counted more frequently. This generally leads to higher numbers of cases when the data are counted according to the SEER rules. As an example, in recent years the difference between the IARC and SEER rules for the total annual new breast cancer cases in British Columbia was about 6%.

In Canada, all cancer registries have now moved to collecting their data under the SEER rules. However, because some regions historically collected data under the IARC rules and only changed very recently, most of the cancer trends (and projections) are calculated using data compiled under the IARC rules. Thus when looking at national reports like the Canadian Cancer Statistics annual publication, the data will most often be presented in IARC rules because the intent is to try and provide comparable statistics across different regions. For these publications, those jurisdictions using SEER rules for data collection have their data "converted" to the IARC rules to make data comparable. Data available from the Statistics Canada CANSIM tables will also use the IARC rules as the intent of these tools is to provide data that can be compared for all of the regions in the summaries available. Data from the BC Cancer Agency website or data that are requested from Data Requests however generally will use the SEER rules data.

Why might the BC Cancer Agency present data under SEER rules when other publications are using the IARC rules? One of the major reasons for this is to most accurately describe the cancer burden for British Columbia. Multiple cancers of the same kind do potentially represent additional demands on cancer diagnosis, treatment and other services. Thus by reporting all cases identified in the population, even if they appear at the same anatomic site, we are providing data that can perhaps better be used to forecast future service demands. Similarly, from a research and clinical point of view, individuals with multiple cancers are important to understand and data coded under SEER rules permit us to see how often patients present with these multiple tumours.

Currency of Data

Projections and incidence rates provided for British Columbia can vary depending on the most recent data included in the calculations. It is common that the data made available by the BC Cancer Agency via web-based publications or through data requests are more current than the data available from national sources (such as the Canadian Cancer Statistics publication or the Statistics Canada website). The main reason for this is that when data are finalized with British Columbia for a given diagnosis year, our Cancer Registry starts a submission process to the national Canadian Cancer Registry. Data submitted to the national cancer registry cannot be released until all provinces have completed submissions and final validated files are produced and released by Statistics Canada (who holds the national registry). As such, the national data set is often one full year behind the provincial data sets.

This can also affect previous years of diagnosis information as in every submission to the national registry provinces submit any new cancers that were reported late from previous years. Thus a submission of 2013 cancer cases might also include another 3-4% of cases from 2012 and even some from previous years. Thus, until data from the most recent year are uploaded into the national data set, there is still some potential that previous years can differ from provincial counts as well.

Population Standards

For calculation of age-standardized cancer incidence or mortality rates one must select a population to use as the "standard". In Canada, up until about 2016, most publications used the 1991 Canadian population as the standard; starting in 2016, many statistical agencies started using the 2011 Canadian population as the new standard for health statistics. If a given publication chooses to use a different population standard, the results (meaning the rate statistics provided) can change and the change can be substantial. Incidence rates calculated from the 1991 and 2011 Canadian populations can be more than 30% different for some cancers and thus it is important to compare only statistics produced using the same population standard.

Common standards that are used to generate incidence and mortality rate statistics for British Columbia include the 1991 Canadian Population, the 2011 Canadian Population, the 1960 World Population and the 2000 US Population.

Table 1: Summary of Differences for Common Sources of BC Cancer Registry Data

Sources of Variation in Statistics	**BC Cancer Agency (website or Data Requests)**	**Statistics Canada CANSIM online tables and other reports**	**Canadian Cancer Statistics Publication**	**WHO: IARC Cancer Incidence in Five Continents**
Population Source for Rate and/or Projection Calculations	BC Stats	Statistics Canada	Statistics Canada	Statistics Canada
Coding Rules for Data Used in Statistics	SEER	IARC	IARC	IARC
Currency of Data	Generally most current available for BC	Generally ~1-2 years behind BC publications	Generally ~1-2 years behind BC publications	Generally 3-5 years behind BC publications
Population Standard for Rate Calculations	2011 Canadian*	2011 Canadian*	2011 Canadian*	1960 World
* Note: Older publications may have standardized using the 1991 Canadian population and thus attention should be paid when reviewing publications written prior to 2016.				