

| Article details: 2020-0139                 |  |
|--|--|
| Title                                      | Sociodemographic characteristics of women with invasive cervical cancer in British Columbia, 2004–2013: a descriptive study  |
| Authors                                    | Jonathan Simkin MPH, Laurie Smith RN MPH, Dirk van Niekerk MB ChB, Hannah Caird MPH, Tania Dearden RN BN, Kimberly van der Hoek MSc, Nadine R. Caron MD MPH, Ryan R. Woods PhD, Stuart Peacock DPhil, Gina Ogilvie MD DrPH   |
| Reviewer 1                                 | Ayman Altalib  |
| Institution                                | McGill University, Montréal, Que.  |
| General comments (author response in bold) | A great job.<br><b>Thank you! We appreciate you reviewing our manuscript.</b>  |
| Reviewer 2                                 | Rachel Kupets  |
| Institution                                | Surgical Oncology, Sunnybrook Cancer Center, Toronto, Ont.   |
| General comments (author response in bold) | <p>Thank you for allowing me to review this paper. I have a few comments/ questions</p> <p>Questions:</p> <p>1. Asians had a lower SR that expected, why does conclusion state that non white/ Caucasian had higher rate?<br/> <b>Thank you for your comment. We used the term “non-Caucasian/White” as it captures both visible minorities and Indigenous population groups. In both the ‘Overall’ and ‘Within not a visible minority’ analyses, we found that there were significant differences between observed and expected counts and individually for these groups, the SRs were elevated. You are correct that the South Asian SR was lower than expected. We have changed the conclusion to be more reflective of the groups with elevated SRs:</b><br/> <i>Women who self-identified as a visible minority, Indigenous, current smoker, non-married and from rural areas were over-represented among women with ICC.</i></p> <p>2. Only data on 1315/1705 cases, how would missing data affect results? Sensitivity analysis?<br/> <b>Thank you for your comment. HAF forms were not available for cases that were not seen in consultation at BC Cancer. We believe that cases not seen in consultation were more likely to be from Northern BC, where there is a larger rural/remote and Indigenous population. We examined cases seen in consultation and those who were not in the table below. We included a comment in the limitations regarding the relationship between referral patterns, access to BC Cancer centres, geography, and Indigenous ethnicity:</b></p> |

**Historically, access to cancer services in Northern BC were challenging throughout the study period due to a variety of reasons including to the lack of a regional cancer centre. This changed when the BC Cancer “Centre for the North” opened in 2012 which would have had an impact on referral patterns. Therefore, referral patterns likely were lower for Northern BC people, where the proportion of Indigenous people is larger. Given ethnicity/race information was not available for cases who were not seen at a cancer centre, this group may be over-represented in the Northern and Indigenous populations. This likely underestimates the number of Indigenous ICC cases and SR estimate.**

Table. Referral patterns by region among the referred study cohort and non-referred cases in BC, 2010-2013.

| Health Authority     | Counts   |  | Proportion (%)                                |  | Ratio<br>(ICC cases<br>not seen in<br>consultation<br>over those<br>seen in<br>consultation) |
|----------------------|--|--|---|--|--|
|                      | ICC study<br>cohort (seen<br>in<br>consultation) | ICC cases not<br>seen in<br>consultation | ICC study<br>cohort (seen in<br>consultation) | ICC cases not<br>seen in<br>consultation |  |
| Fraser               | 421  | 93                                       | 34.7  | 25.1                                     | 0.72   |
| Interior             | 174  | 70                                       | 14.3  | 18.9                                     | 1.32   |
| Northern             | 75   | 47                                       | 6.2   | 12.7                                     | 2.05   |
| Vancouver<br>Coastal | 305  | 86                                       | 25.1  | 23.2                                     | 0.92   |
| Vancouver<br>Island  | 229  | 65                                       | 18.8  | 17.6                                     | 0.94   |
| Unknown              | 11   | 9  | 0.9   | 2.4                                      | 2.67   |

3. For the population including Korean, Filipino and Japanese, how long were they living in Canada? Could they have arrived in Canada with prevalent disease? What was the stage distribution of disease? Did the higher risk group outlined in the paper have advanced disease as compared to others?

**Thank you, these are important questions. The BC Cancer Registry is a prospectively maintained database that collects information on all tumours among BC residents. Information is obtained from provincial pathology and laboratory reports, surgical reports, the cervical cancer screening program, hospital data, vital statistics, among other sources. Cases reported in this study are all new cases diagnosed in British Columbia among BC residents.**

Country of birth and time since immigration are important and complementary factors to self-identified ethnicity, but they were not available on HAFs. In the discussion, we discuss evidence around cervical cancer incidence, stage at diagnosis, and screening as it relates to population groups characterized by country of birth and immigration history (see discussion section “ethnicity/race, cervical cancer screening and incidence in Canada”). We also included a comment in the limitations:

*Information on factors related to screening, primary care, immigration history, country of birth, socioeconomic status, HIV status and BMI, were not available in HAFs.*

In respect to stage at diagnosis distribution, this is an important question that we do not address in this study. The objective of this paper was to examine the distribution of sociodemographic and health characteristics between the ICC cohort and the general population. Comparing characteristics between population groups within the ICC study cohort was not within the scope of our objectives but we do agree it is important work. We recommend further work in this area in our conclusion:

*The present study suggests that social and sociocultural determinants of health are closely linked to cervical cancer incidence, even in the context of universal healthcare access. Information on indicators of cervical cancer screening by ethnicity/race have not been available in BC and this void is a barrier to informing equitable cancer control and prevention strategies. Efforts are needed to reverse this information gap to better understand and address inequities, and inform culturally relevant and population-based strategies to ensure elimination of cervical cancer for women, trans and non-binary people at risk of this preventable malignancy.*

We also discussed stage of disease among different populations groups in the discussion:

*Few studies have examined ICC incidence and ethnicity/race in Canada. One study revealed significant variations of ICC stage distribution by birth region<sup>18</sup>. Immigrant women from East Asia, Western Europe and America had higher early stage incidence while South Asian women had higher late stage incidence. This study emphasizes that taking into account immigration history and birth region can identify important differences between population groups.*

4. What is the screening history of the patients in the cohort?

Thank you, screening history is an important factor related to ICC incidence. Screening history was not available to us for this study. We included both a comment in the limitations and the conclusion

Limitations:

*Information on factors related to screening, primary care, immigration history, country of birth, socioeconomic status, HIV status and BMI, were not available in HAFs.*

Conclusion:

*Information on indicators of cervical cancer screening by ethnicity/race have not been available in BC and this void is a*

***barrier to informing equitable cancer control and prevention strategies.***

5. How accurate are the age adjusted “expected counts”? what if they over or underestimate the actual cancer rates in each subpopulation being assessed? Doesn’t seem that these expected counts can be validated?

**Thank you for asking about the expected counts. Expected counts were derived using direct age-standardization, an established statistical method used in disease surveillance and epidemiological research. The purpose of the direct method is to adjust a frequency measure for a given population (or populations) relative to the age-distribution of a given population standard. Frequency measures that have been adjusted to the same population standard can be compared while accounting for the effect of age.**

**The accuracy of the expected counts depends on the accuracy of the data source used to derive general population estimates. For analyses on ethnicity/race, language, and CHSA classifications, data were derived from the Census of the population and therefore, we believe the expected counts to be highly accurate. The Census of Canada is mandated by law in Canada to determine the population of Canada every five years. It is the primary source of socioeconomic data for specific population groups. Survey net undercoverage can impact census data accuracy and response rate can impact data quality. The net undercoverage of the 2006 Census was estimated to be 2.8%. The response rate of the 2006 census was 96.5%. The 2011 NHS differed from the 2006 Census in that it was a voluntary survey. The response rate was 77.2%. The Chief Statistician of Statistics Canada released a commentary on the lower response rate of the NHS 2011. In this commentary, he states that several measures were taken to offset data quality risks associated with the move to a voluntary survey – specifically they address variation in response rates at low levels of geography, mitigated risks in sampling error and non-response bias. Estimates that were not of sufficient quality were not released. Ultimately, the Chief Statistician stated that the NHS 2011 produced a rich and robust database of information. More on the strategies used to mitigate sources of error and bias are described here:**

**<https://www.statcan.gc.ca/eng/blog-blogue/cs-sc/2011NHSstory>.**

**We included descriptions of the Census 2006 and NHS 2011 in Table 1 in the methods. We included the response rate of the 2006 Census and 2011 NHS, and also included references for the readership to find response rates and more information on the 2011 NHS survey:**

***In 2011 the mandatory long-form census was replaced with a voluntary survey called the National Household Survey (NHS). The NHS collected similar information as gathered from the Census. The national response rate in was 77.2%<sup>27</sup>. Statistics Canada implemented various methods to account for error and biases related to voluntary nature of the survey<sup>28</sup>. The 2011 NHS PUMF on individuals represents a 2.7% sample of the Canadian population. It contains social, demographic, and economic data<sup>23</sup>.***

**For the Canadian Community Health Survey, 2011/12, the response rate was 68.4%. This was added to table 1 description of the data sources. We also added non-response bias to the limitations:**

***Data sources used for the general population contain potential sampling error and non-response biases.***

In terms of how the expected counts can be validated, the data are derived from data made available from Statistics Canada (public use microdata files for the Census 2006, NHS 2011 and the CCHS 2011/2012). Information released from these surveys are robust and represent sociodemographic and health characteristics on the Canadian population.

Age-standardization is an established method used to compare measure of risk between population groups. We provide a reference to the direct age-standardization method (Boyle and Parking, 1991; reference 33 in the revised manuscript). Anyone with the Census microdata files and the ICC study cohort age structure can derive these estimates. Below is an example of how the expected counts were calculated. A description is also provided in the methods of our paper.

***Example of deriving expected counts from the age-standardized proportion***

Referring to Table 2 (previously Table 2a), the age-standardized proportion of Chinese people as reported by the Census was 35.4% (recall we age-standardize the Census population to the same age-distribution as the ICC study cohort to account for differences in age structure). This proportion multiplied by the total number of observed cases (N=320) yields the expected number of cases among Chinese people, relative to the ethnicity distribution in the Census (i.e. the general population).

6. Reasons for elevated risks are different between populations ie: are reasons for elevated risks in First Nations same as those in Filipino, Korean or Japanese populations?

**Thank you for your comment. We do believe that risk differs across different population groups and that Indigenous people experience a unique cancer burden. This is why we discussed the literature related to cervical cancer incidence and screening among Indigenous people in a separate paragraph from visible minorities and immigrant populations are discussed.**

7. The absolute numbers of cases in Table 2a with in visible minority are very small: 118, 72, 38, 21, 17- that is over the almost 10 years of the study which is around 2-12 patients per year. Difficult to interpret data and conclusions from such small numbers.

**Thank you for this comment, small numbers are a common and challenging issue in cancer epidemiology. This is a primary reason why we combined ten years of data – data aggregation is a common strategy to address statistical issues related to small numbers.**

**The statistical test implemented here is a goodness of fit chi-square test. A chi-square goodness-of-fit tests the null of no differences between observed and expected values. This test requires that the expected values in each level of the variable is at least 5, which we satisfy. Significant tests indicated that there are significant variations between**

observed and expected values. We included the null hypothesis in the methods to ensure the readership understood the meaning and purpose of our test:

*A chi-square goodness-of-fit test tested the null of no differences between observed and expected values.*

On its own, rejecting the null hypothesis does not provide very useful information, which is why it is good practice to also present absolute and relative measures for interpretability. Therefore, we also included in our tables the number of observed cases, the age-adjusted expected cases, proportions, and standardized ratios.

For example, there were 320 cases among persons who identified as visible minorities over a ten year period. Relative to the ethnicity distribution of BC females as reported by the Census, we would expect 253 cases. This is an excess of 67 cases. On an annual level, that is an excess of 6 to 7 cases. Annually this may be considered as small, however, cervical cancer is almost entirely preventable and over time, the excess numbers add up.

8. Also when adding together the number of expected cases “with in visible minority”, there are 321 cases vs. the 320 observed. So numbers seem quite similar though distribution between various minorities may be different. For Chinese there were 118 observed vs. 113 expected- for only 5 difference, for South Asian there were more expected than observed, for Filipino there were 14 more extra cases calculated, Japanese 7 extra cases, so it is very difficult to interpret the meaning of these numbers.

**Thank you for your comment. The comparison of 320 observed and 321 expected visible minority cases is incorrect. For the ‘Within visible minorities’ analysis, we break this number down by various population groups and compare the observed to what we would expect relative to Census data. The expected column (highlighted in yellow below) must add up to the total of the observed cases (highlighted in blue below) among visible minorities (i.e. 320 cases). If you add up the expected cases (highlighted in yellow), you are correct that it adds to 321 cases and this is due to rounding – the expected cases are rounded to a whole number for interpretability as opposed to a decimal. As described in the methods, expected counts are derived from multiplying the Census age-standardized weighted proportion (highlighted in green) by the total of the observed cases for that comparison group (N=320). The correct comparison for visible minority observed vs. expected cases is in the first few rows (highlighted in pink below). Specifically, we observed 320 cases, while relative to the Census distribution (general population), 253 were expected. This is an absolute difference of 67 cases or, in a relative sense, we observed 1.3 times the number of cases than expected.**

*Example of deriving expected counts from the age-standardized proportion*

Referring to Table 2 (previously Table 2a), the age-standardized proportion of Chinese people as reported by the Census was 35.4% (recall we age-standardize the Census population to the same age-distribution as the ICC study cohort to account for differences in age structure). This proportion multiplied by the total number of observed cases (N=320) yields the expected number of cases among Chinese people, relative to the ethnicity distribution in the Census (i.e. the general population).

Therefore, the total of the expected and observed will be the same (slight variation due to rounding).

| Category                             | Observed cases | ICC cases crude Proportion (%) | Census Age-standardized proportion (Weighted) (%) | 95% CI (%) | Expected Cases | Standardized Ratio | 95% CI (%) | Test Statistics (Chi) | P value |
|--------------------------------------|----------------|--------------------------------|---|------------|----------------|--------------------|------------|-----------------------|---------|
| <b>Ethnicity/ Race</b>               |                |                                |   |            |                |                    |            |                       |         |
| <b>Overall</b>                       | N=1117         |                                |   |            |                |                    |            | 22.8                  | <0.0001 |
| Not a visible minority               | 797            | 74.4                           | 77.3  | 77.2-77.4  | 864            | 0.92               | 0.92-0.92  |                       |         |
| Visible Minority                     | 320            | 28.6                           | 22.7  | 22.6-22.7  | 253            | 1.26               | 1.26-1.26  |                       |         |
| <b>Within not a visible minority</b> | N=797          |                                |   |            |                |                    |            | 55.9                  | <0.0001 |
| Caucasian/White                      | 712            | 89.3                           | 95.1  | 94.9-95.2  | 758            | 0.94               | 0.94-0.94  |                       |         |
| Indigenous                           | 85             | 10.7                           | 4.9   | 4.9-5.0    | 39             | 2.16               | 2.18-2.12  |                       |         |
| <b>Within visible minority</b>       | N=320          |                                |   |            |                |                    |            | 38.8                  | <0.0001 |
| Chinese                              | 118            | 36.9                           | 35.4  | 35.2-35.6  | 113            | 1.04               | 1.04-1.05  |                       |         |
| South Asian                          | 72             | 22.5                           | 35.9  | 35.8-36.1  | 115            | 0.63               | 0.62-0.63  |                       |         |
| Filipino                             | 38             | 11.9                           | 7.4   | 7.4-7.5    | 24             | 1.60               | 1.58-1.62  |                       |         |
| Korean                               | 21             | 6.6                            | 3.7   | 3.6-3.7    | 12             | 1.78               | 1.76-1.80  |                       |         |
| Japanese                             | 17             | 5.3                            | 3.0   | 3.0-3.1    | 10             | 1.77               | 1.74-1.79  |                       |         |
| All other                            | 54             | 16.9                           | 14.5  | 14.4-14.6  | 47             | 1.16               | 1.15-1.17  |                       |         |

9. Can authors use census data to find true denominator of Filipino, Korean and Japanese individuals living in BC to calculate rates of ICC per 100,000?

**Thank you for your question. The Census data only covers Census years 2006 and 2011 – they do not include intercensal annual estimates (e.g. 2007, 2008, 2009...). We used the ethnicity distributions from the 2006 and 2011 Census data files to derive age-adjusted expected counts. It is possible to estimate intercensal population by age and population group and derive crude and age-standardized incidence rates however, for the purposes of this study we believe that a comparison of observed and expected counts was a more informative approach.**

10. Small numbers, estimates for expected vs. observed, no information on screening history or stage of disease, access to primary care information, difficult to interpret this data.

**Thank you for your feedback. While there may be small numbers in some categories, we believe that reporting data by population groups, where information is not readily available, can support cancer prevention and control strategies in BC and elsewhere. We took several steps to mitigate issues related to small numbers. That included aggregation of data over a ten-year period, statistical methods that were appropriate for our sample size, and comparison with key national datasets from Statistics Canada. Our findings were consistent with literature from other regions in Canada, the United States and globally, suggesting that social and sociocultural determinants of health are closely linked to cervical cancer incidence, even in the context of universal healthcare access. We presented both absolute and relative measures to improve interpretability. We agree that screening history, stage of disease and primary care information could help better understand differences in risk between population groups with ICC. However, comparing characteristics between population groups within the ICC study cohort was not within the scope of our objectives. Our objective was to compare the observed and expected cases relative to the general population of women in BC. Again, we agree that this is important work and we recommend further work in this area in our conclusion.**