

## The Canadian Forces Cancer and Mortality Study II: a longitudinal record-linkage study protocol

Elizabeth Rolland-Harris MSc PhD, Linda VanTil DVM MSc, Mark A. Zamorski MD MHSA, David Boulos MSc, Alexander Reicker MSc, Huda Masoud MSc, Richard Trudeau BSc, Murray Weeks PhD, Kristen Simkus MPH

### Abstract

**Background:** Military service exposes personnel to unusual situations with unclear health-related implications, and to identify both immediate and delayed risks, part of health surveillance includes examination of mortality and cancer rates that extends beyond periods of military service. The main aim of the Canadian Forces Cancer and Mortality Study II (CFCAMS II) is to describe the mortality and cancer experience of Canadian Armed Forces personnel (serving and released; about 230 000 people), with the further aim of informing health promotion and prevention programs for serving personnel and services for veterans after they leave the military.

**Methods:** This protocol is for a retrospective cohort study of serving and released Canadian Armed Forces personnel who enrolled on or after Jan. 1, 1976 in the Regular Force or Class C of the Reserve Force. To create our cohort, we identified record-linkage methods as the most appropriate mechanism to study mortality and cancer in those with a history of Canadian military service. Statistics Canada will link the CFCAMS II cohort file to the Canadian Vital Statistics (Mortality) and Canadian Cancer Registry databases for outcomes up to Dec. 31, 2014. The linkage will be stored in their highly secure linkage environment. Statistical analyses will be broadly divided into mortality and cancer incidence.

**Results:** We will quantify mortality and cancer morbidity incidence and survival using multiple established methods, as well as age-period-cohort regression models to describe the relation between military service and mortality and cancer outcomes.

**Interpretation:** The findings will represent novel and sound evidence on the risks and protective factors of military life.

As a military organization, the Canadian Armed Forces has the challenge of balancing its defence mandate responsibilities with its duty to protect its personnel. The nature of its operations can be accompanied by unusual exposures, with both expected and unknown risks to personnel. Adverse outcomes, including death, may occur. To identify immediate and delayed risks, the Canadian Armed Forces has included cancer and mortality surveillance as part of its health surveillance strategy for over 15 years. This strategy includes a broad cross-section of methodological approaches and data sources, documenting the various stages of the military life course, as they pertain to long-term health outcomes.<sup>1</sup> This has included cluster investigations, record linkage studies<sup>2,3</sup> and surveillance systems.<sup>4,5</sup>

The Canadian Armed Forces's latest record-linkage study, the Canadian Forces Cancer and Mortality Study II (CFCAMS II), has been developed to address remaining evidence gaps. As such, the underlying hypothesis of the CFCAMS II is that military service has a differential effect on mortality and cancer incidence. Specifically, the study's main

aim is to describe the mortality and cancer experience of Canadian Armed Forces personnel (serving and released). The evidence from this study will be used by the Canadian Armed Forces to inform health promotion and policies for serving personnel, as well as by Veterans Affairs Canada in their provision of care and services for veterans and their families after they leave military service. Specific research questions include: What is the burden of mortality among those with a history of military service? and Are any occupational factors strong predictors of excess mortality or cancer incidence among those with a history of military service?

**Competing interests:** None declared.

This article has been peer reviewed.

**Correspondence to:** Elizabeth Rolland-Harris, Elizabeth.harris@forces.gc.ca

**CMAJ Open 2018. DOI:10.9778/cmajo.20170125**

**Methods**

This protocol is for a retrospective cohort study of serving and released Canadian Armed Forces personnel.

**Participants**

The CFCAMS II is a longitudinal record-linkage study that will describe the mortality and cancer incidence burden among all Regular Force and Class C Reservist personnel enrolled by the Canadian Armed Forces for the first time between 1976 and 2015 inclusive. In other words, the study population is the full population (i.e., this study is a census), and, consequently, no exclusion criteria exist beyond initial enrolment outside this time frame.

**Defining and building the cohort**

The 1976 starting point was chosen because of the availability of administrative information in the Central Computerized Pay System (CCPS), an electronic Department of National Defence data set with information available beginning in 1976. Pre-1976 data exist within the CCPS, but their completeness is questionable. By using 1976 as the cohort start year, cohort completeness (and the resultant census-like nature of the cohort) will be ensured.

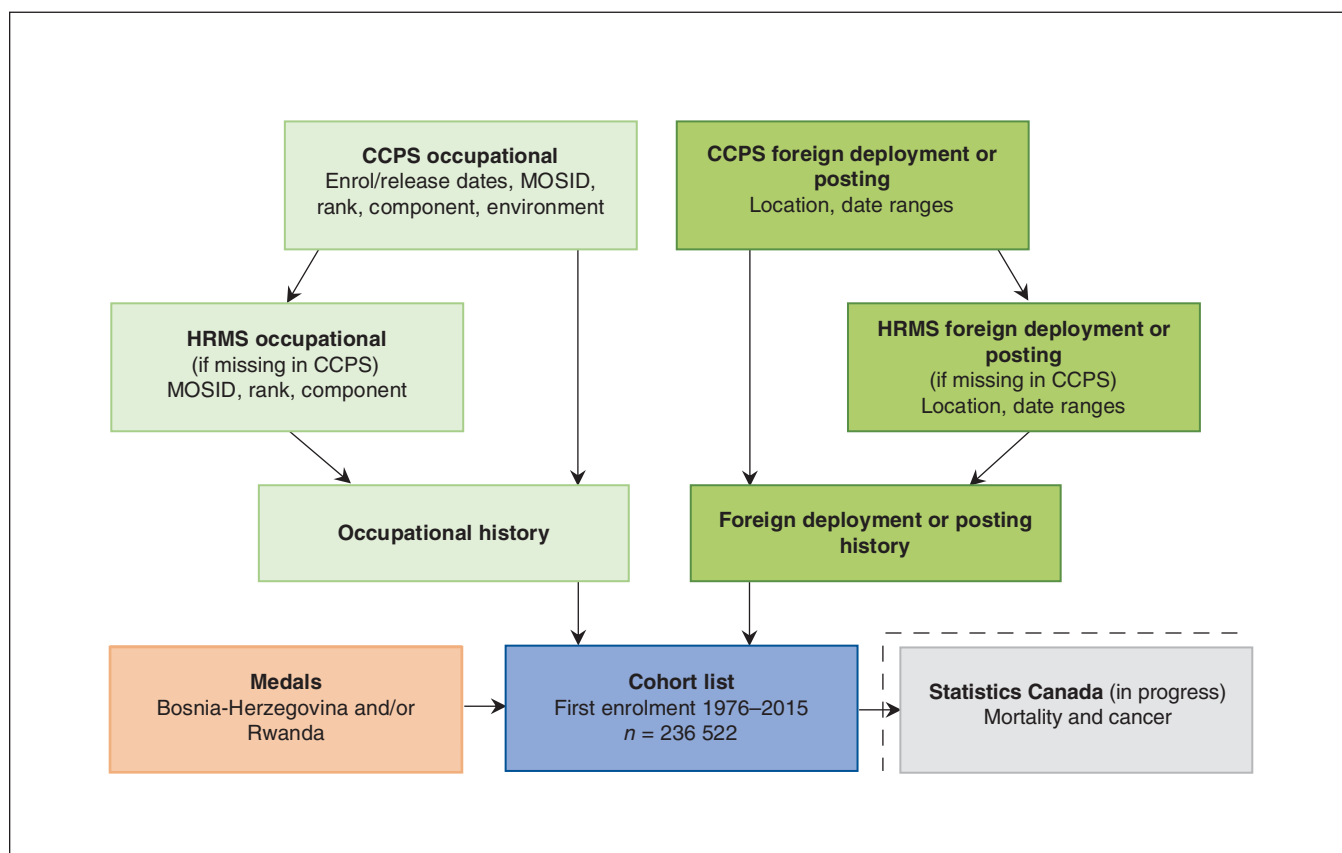
The CCPS includes Regular Force and Class C Reserve Force (reservists on international operations and/or tours of duty) members. Data in the CCPS have an advantage in terms

of accuracy and precision over human resources data when used to identify membership in the CFCAMS II cohort and other information associated with remuneration. They also have a built-in feedback mechanism whereby both service members and the employer are motivated to rectify remuneration errors as soon as possible. This creates an informal data validation process that did not exist for the CFCAMS I.<sup>6</sup>

Class A and B Reserve Force personnel are excluded from this cohort. The rationale is described in more detail in section I of Appendix 1 (available at [www.cmajopen.ca/content/6/4/E619/suppl/DC1](http://www.cmajopen.ca/content/6/4/E619/suppl/DC1)).

Using data extracted from the CCPS, epidemiologists from the Directorate of Force Health Protection created a cohort file (Figure 1). This cohort file was supplemented and validated with the use of human resources data, thereby minimizing missing information and resolving imprecise or apparently anomalous information (e.g., unrealistic birth date). Access to and use of these data files were considered to be consistent with the purpose for which these data were collected and maintained (i.e., “consistent use”) under the Canadian Government Info Source program.

Data elements included in the cohort file submitted to Statistics Canada are catalogued in Table 1. Multiple enrolments and releases (if relevant) are also captured. Data quality considerations with the use of CCPS data to create the cohort file are described in section II of Appendix 1.



**Figure 1:** Flow chart showing the Canadian Forces Cancer and Mortality Study II (CFCAMS II) cohort-building process. Note: CCPS = Central Computerized Pay System, HRMS = Human Resources Management System, MOSID = Military Occupation Structure Identification code.

Table 1: Data elements in the Canadian Forces Cancer and Mortality Study II cohort file	
Variable	Description
<b>Demographic information</b>	
Name	These data elements will be used by Statistics Canada to link the study cohort file to the Vital Statistics – Death Database and the Canadian Cancer Registry. Once they are linked, only date of birth and sex will be retained; name and social insurance number will be truncated, and a randomly generated unique identifier will be issued for each unique individual.
Sex	
Date of birth	
Social insurance number	
<b>Occupational information</b>	
Rank(s)	Initial rank and subsequent rank changes, as well as related date(s) of rank changes (promotions and demotions) will be included in the linked cohort file. Each rank change and corresponding date will be captured as their own variables.
Command(s)	Initial command (Regular Force or Class C Reserve Force) and subsequent command change(s), as well as related date(s) of command changes will be included in the linked cohort file. Each command change and corresponding date will be captured as their own variables.
Element(s)	Initial element (Army, Navy or Air Force) and subsequent element changes, as well as related date(s) of element changes will be included in the linked cohort file. Each element change and corresponding date will be captured as their own variables.
Enrolment date(s)	Start of a person's employment as a military member. The first iteration of this variable will be used to ascertain whether a possible participant's first enrolment date corresponds with the cohort's date-related inclusion criteria. As personnel can leave and subsequently reenrol, as well as change between Regular and Reservist service, participants may have > 1 enrolment date. Each enrolment date will be captured as its own variable.
Release date(s)	End of person's employment as a military member. As personnel can leave and subsequently reenrol, as well as change between Regular and Reservist service, participants may have > 1 enrolment date. However, it is possible that a person may have no release date; this means he/she was still in service at the time of study end. Each release date will be captured as its own variable.
Reason(s) for release	For each release date captured in the cohort file, there should be a corresponding reason for release (voluntary, involuntary, medical or unknown). In the event of multiple releases, each reason for release and its corresponding date will be captured as their own variables.
Occupational history	Information on participants' military occupation (Military occupation code or Military Occupation Structure Identification code).
Foreign posting and deployment history	Start and stop date(s) of any international military operation. The information captured as part of this includes the operation name (if applicable), the start and stop date(s) and the location(s) of the posting/deployment.
<b>Derived information</b>	
Person-years of service	Variable derived with the use of enrolment and release dates to calculate cumulative person-years of service for all participants. Will also be used in analyses to control for effect of length of service on outcomes of interest.
Time since release	Variable derived only for participants with a terminal release date (i.e., no subsequent reenrolment date). Quantifies the time between release and death or censorship.
Era of service	Variable derived to control for era effects within the data. Era of service categories used elsewhere, as well as other stratifications based on salient evidence, will be evaluated for suitability and optimal statistical power.
Deployment — dichotomous	Dichotomizes participants into "have deployed" v. "never deployed" on international deployment.
Deployment — cumulative time	Variable derived by calculating a cumulative amount of time on international deployment until release, death or censorship.
Deployment — frequency	Variable derived by counting the number of discrete international deployments a participant will have been on until release, death or censorship.
Deployment — specific deployment flag	For specific deployments of interest (e.g., Afghanistan, Gulf War, Rwanda), a flag may be created to identify those who participated in specific foreign military operations.

### Data linkage

Statistics Canada's Social Data Linkage Environment is a highly secure linkage environment that facilitates the creation of linked population data files for social analysis. At the core of the Social Data Linkage Environment is a Derived Record Depository, a national dynamic relational database containing limited personal identifiers. The depository is created by linking selected Statistics Canada source index files to produce a list of unique individuals. These files, which contain personal identifiers without analysis variables, are brought into the Social Data Linkage Environment, processed and linked to the Derived Record Depository. Updates to these data files are linked to the depository on an ongoing basis. A more detailed explanation of the data linkage process is provided in section III of Appendix 1.

For the CFCAMS II, the cohort file provided by Statistics Canada will be linked to 2 discrete data holdings within the Social Data Linkage Environment: the Vital Statistics – Death Database and the Canadian Cancer Registry. No civilian data from other sources will be used.

The Vital Statistics – Death Database provides mortality information from 1950 onward to the latest available year of data. Information from the database is provided by the International Classification of Diseases (ICD) code version in effect at the time of death. Cause of death information is recorded by vital statistics registries in the provinces and territories, with follow-up for these records taking years to complete. The provinces and territories share their information (under the Statistics Act) with Statistics Canada for national reporting and to allow linkage for approved projects.

The Canadian Cancer Registry contains cancer incident events from 1969 onward to the latest available year of data. All new incident cases of cancer are recorded by cancer registries in the provinces and territories. People can have more than 1 record in the registry depending on how many cancers have been reported or how many registrations have been submitted for a given person. Cancer incidence information in the Canadian Cancer Registry is coded to the ICD code version in effect at the time of registration. This can be years after the date of diagnosis, allowing for the follow-up required to confirm the diagnosis and complete the registration. The registries then share their information (under the Statistics Act) with Statistics Canada for national reporting and to allow linkage for approved projects.

Specific mortality (outcome) variables to be requested are date of birth, sex, province or country of birth, province of residence at time of death, province/country of death, full date of death, underlying cause of death (ICD code), all causes of death (for deaths in 2000 onward owing to Vital Statistics – Death Database holdings) and autopsy code. Malignant tumour variables to be requested are listed in Table 2. Their quality is discussed in section II of Appendix 1, and expected timelines are given in Figure 2.

### Living cohort

As the Canadian Armed Forces military cohort continues to grow over time, the addition of new records for new enrollees, as well as updates to existing cohort records held in the Social

Data Linkage Environment are planned. In turn, these will periodically be relinked to the Vital Statistics – Death Database and Canadian Cancer Registry, providing more person-years of observations, more outcome data and the ability to study the possible impact of new military actions on health outcomes.

### Sample size and power calculations

One of the main strengths of the CFCAMS II is that the study population is a census. This eliminates sampling error and the corresponding need for sample size calculations.<sup>7,8</sup> Given that a priori power calculations imply that sample size can be modified in response to poor power, it is of poor limited usefulness in this census-like context. However, outcome projection models were built and are presented in section IV of Appendix 1, and post hoc power calculations will be derived when answering any of the study questions.

### Preanalysis data manipulation

Before data analysis can begin, some study data editing/manipulation will be required. These considerations are described in more detail in section V of Appendix 1, and the proposed derived military service variables are presented in the latter half of Table 1.

### Statistical analyses

Statistical analyses will be broadly divided into mortality and cancer incidence. An overview of the proposed incidence and/or prevalence, survival and regression methods within these 2 subsets is presented in Table 3. Analyses will be conducted with Stata 14 (StataCorp).

### Considerations

- All analyses assume sufficient power to respond to the research question(s).
- Whenever possible, analyses will be:
  - Stratified by sex or by serving status. If a research question justifies it, analyses may also be stratified by component (Regular Force only, Class C Reserve Force only, combination of Regular Force and Class C Reserve Force service).
  - Age- and sex-adjusted.
- Cancer analyses will include only invasive malignant neoplasms.
- For cancer incidence data, multiple primary cancers will be identified with the use of the International Agency for Research on Cancer rules,<sup>9</sup> per Statistics Canada practice.<sup>10</sup>

### Mortality

For the purpose of our analyses, mortality will be classified in 3 ways: 1) all-cause mortality, 2) cause-specific mortality and 3) population- or deployment-specific mortality.

### Incidence

Incidence will be presented as follows:

- Incidence rates per 100 000 observed person-years.
  - The use of observed person-years as a denominator is a more exact measure of population time at risk,<sup>11</sup> particularly in a dynamic cohort.

<b>Table 2 (part 1 of 2): Requested Canadian Cancer Registry data elements</b>	
Variable	Description
<b>Patient variables</b>	
Patient record type	Type of record (new, update, delete)
Sex	
Date of birth	
Province or country of birth	
Province of diagnosis	
Date of death	
Underlying cause of death	
Autopsy confirming cause of death	Code indicating whether cause of death from official death certificate takes into account autopsy findings
<b>Derived patient variables</b>	
Vital status	
No. of tumours	No. of tumour records belonging to patient record
<b>Tumour variables</b>	
Tumour reference number	Unique tumour identification number
Method of diagnosis	
Date of diagnosis	
ICD-9 cancer code	
Source flag classification	Indicates classification system used to code topography, histology and behaviour of tumour
ICD-O-2/3 topography	Site of origin of neoplasm – ICD-O-2/3 coding
ICD-O-2 histology	Histological description of neoplasm – ICD-O-2 coding
ICD-O-2 behaviour	Behaviour associated with histological description of neoplasm – ICD-O-2 coding
Laterality	Site-specific localization of tumour in paired organs or on side of body on which tumour originated (right, left, bilateral)
ICD-O-3 topography	Site of origin of neoplasm – ICD-O-3 coding
ICD-O-3 histology	Histological description of neoplasm – ICD-O-3 coding
ICD-O-3 behaviour	Behaviour associated with histological description of neoplasm – ICD-O-3 coding
Grade, differentiation or cell indicator	Describes system used to identify type of grade/differentiation/cell indicator
Method used to establish date of diagnosis	Code that specifies method by which date of diagnosis of tumour was established
Diagnostic confirmation	Most accurate diagnostic confirmation
CS tumour size	Largest dimension/diameter of primary tumour in millimetres
CS extension	Primary tumour growth within organ of origin or its direct extension into neighbouring organs
CS tumour size/ext-eval	Code indicating how “CS tumour size” and “CS extension” were determined and diagnostic methods used
CS lymph nodes	Site-specific code identifying regional lymph nodes involved with cancer at time of diagnosis
CS reg nodes eval	Code indicating how “CS lymph nodes” code was determined and diagnostic methods used

**Table 2 (part 2 of 2): Requested Canadian Cancer Registry data elements**

Variable	Description
Regional nodes examined	Total no. of regional lymph nodes that were removed/ examined by pathologist
Regional nodes positive	Exact no. of regional nodes examined by pathologist and found to contain metastases
CS mets at dx	Code identifying distant site(s) of metastatic involvement at time of diagnosis
CS mets eval	Code indicating how “CS mets at dx” code was determined and diagnostic methods used
AJCC clinical T	Site-specific code evaluating primary tumour clinically (T) and reflecting tumour size and/or extension
AJCC clinical N	Site-specific code identifying absence/presence of clinical regional lymph node (N) metastasis; describes extent of regional lymph node metastasis as recorded
AJCC clinical M	Site-specific code identifying absence/presence of clinical distant metastasis (M)
AJCC pathologic T	Site-specific code evaluating primary tumour pathologically (T) and reflecting tumour size and/or extension
AJCC pathologic N	Site-specific code identifying absence/presence of pathological regional lymph node (N) metastasis; describes extent of regional lymph node metastasis as recorded
AJCC pathologic M	Site-specific code identifying absence/presence of clinical pathological metastasis (M)
AJCC clinical TNM stage group	Site-specific code identifying anatomic extent of disease based on clinical T, N and M elements as recorded in TNM clinical T, N and M fields
AJCC pathological TNM stage group	Site-specific code identifying anatomic extent of disease based on pathologic T, N and M elements as recorded in TNM pathologic T, N and M fields
AJCC TNM stage group	Site-specific code identifying stage group when clinical/ pathologic T, N, M values are incomplete and do not lead to a clinical/pathologic T, N, M group
AJCC edition number	Identified edition of cancer staging manual used to stage case

Note: AJCC = American Joint Committee on Cancer, ICD-9 = International Classification of Diseases, 9th Revision, ICD-O = International Classification of Diseases for Oncology.

- Standardized mortality ratios<sup>12</sup> (see section V of Appendix 1 for more details).
- Lexis diagrams:<sup>13</sup>
  - Present incidence data where age–period–cohort effects are likely to influence mortality rates over a long follow-up period.
  - Particularly useful in visually communicating temporal and population subset differences in incidence.

**Survival**

Bakker and colleagues<sup>14</sup> stated that “the purpose of survival analysis ... is to estimate a patient’s chances of survival as a function of time,” and it does so using both censored and uncensored observations<sup>15</sup> (which traditional regression models cannot do). Survival methods will include:

- Kaplan–Meier survival analysis:
  - Primarily to study survival trends of more frequent causes of death.
  - We will estimate multiple survival curve differences using log-rank tests, except where earlier and later separations between survival curves might be differentially weighted, in which case alternative tests (e.g., Wilcoxon) will be considered.
- Cox proportional hazards regression:
  - Used in instances where variables may contribute to differences between survival curves<sup>16</sup> or where time-dependent variables are potentially important to the ascertainment of risk.<sup>17</sup>
  - Inclusion of time-dependent variables may resolve issues of nonproportionality.



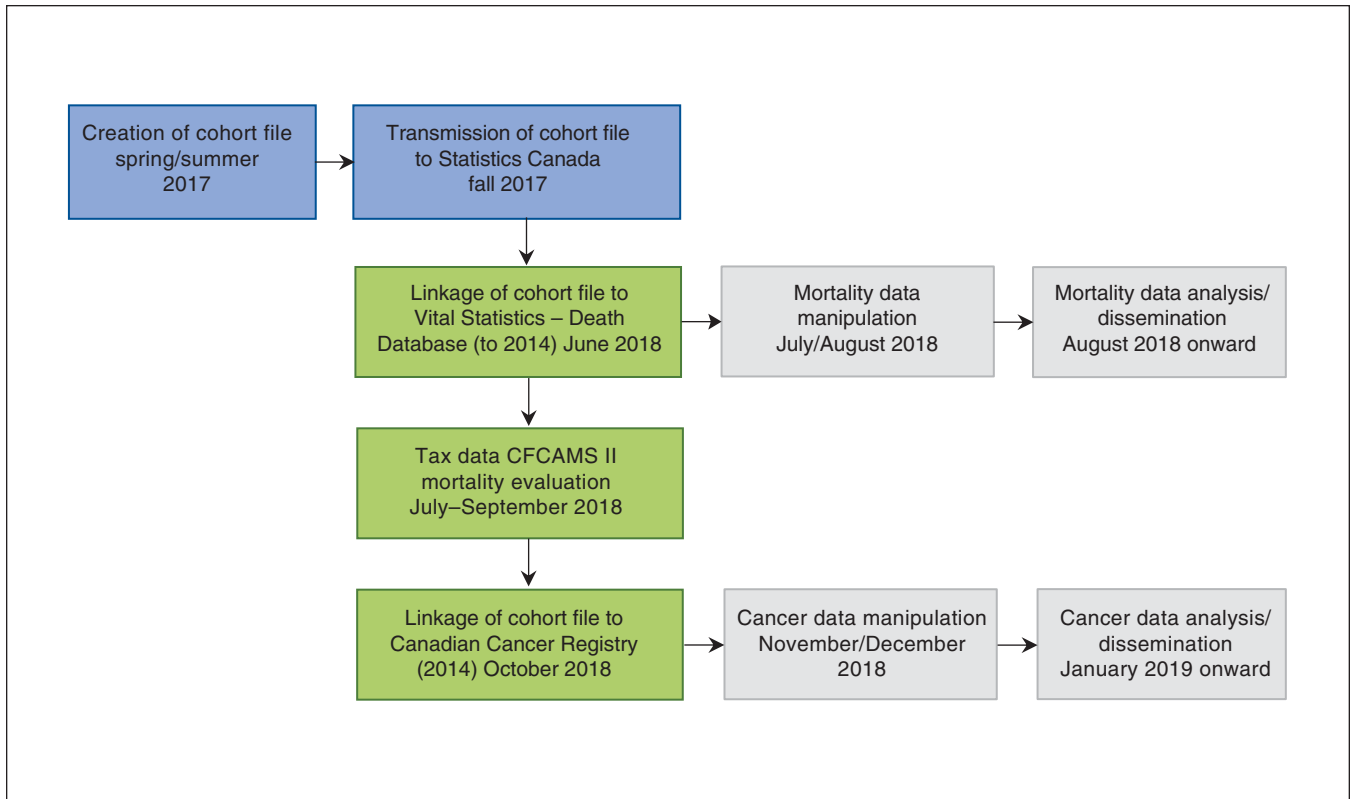


Figure 2: Expected timelines for the Canadian Forces Cancer and Mortality Study II (CFCAMS II).

Table 3: Overview of proposed analyses, by data type		
Measure	Mortality data	Cancer data
Univariate/bivariate analyses	Frequencies Measures of central tendency (where appropriate) Cross-tabulations	
Incidence	Age- and sex-adjusted rates per 100 000 person-years of observation Standardized mortality ratio Lexis diagrams	Age- and sex-adjusted rates per 100 000 person-years of observation, new cases Standardized incidence ratio Standardized mortality ratio* Lexis diagrams
Prevalence	NA	Person-based prevalence, per 100 000 person-years of observation† Tumour-based prevalence, per 100 000 person-years of observation†
Survival	Kaplan–Meier survival analysis Cox proportional hazards regression	
	NA	Pohar-Perme estimator of net survival
Regression	Poisson regression Joinpoint models Age–period–cohort models	
Note: NA = not applicable. *Standardized mortality ratio of death recorded in cancer data specifically. †Period or point prevalence, depending on research question.		

Competing hazards models will not be included, given limited historical data on underlying causes of death.<sup>18</sup>

### Regression models

Age–period–cohort modelling will also be used, as it also controls for secular (“period”) influences, as well as era (“cohort”) influences on rates over time.<sup>19</sup> With the use of the Stata “apcfit” command,<sup>20</sup> this approach can also be used with Lexis diagrams.<sup>21</sup>

Other regression approaches (Poisson, logistic) have been considered, but published evidence suggests that adjusted parameters result in very similar results across methods<sup>22,23</sup> or less precise estimates.<sup>23</sup>

### Cancer incidence data

Cancer incidence analyses share many methodological approaches with mortality incidence analyses. Only methods specific to cancer incidence are described below.

#### Incidence

Incidence is one of the most common descriptive measures of cancer burden. Das<sup>24</sup> described cancer incidence as “the number of newly diagnosed cancers of a specific site/type occurring in a specified population during a defined period.” We will quantify incidence using 1) cancer incidence rates per 100 000 observed person-years, for primary tumours only and/or all instances of multiple primary tumours, depending on the research question, sample size considerations and clinical suitability, and 2) standardized incidence ratios, which are analogous to standardized mortality ratios for first primary tumours but are focused on incident cancer cases.<sup>25</sup> Standardized incidence ratios can also be used to quantify subsequent primary tumours, which can be particularly useful in describing the risk within a cohort of the development of secondary primary cancers.<sup>26</sup>

#### Prevalence

Both person-based and tumour-based prevalence will be generated, for all cancers and for leading cancers separately.<sup>10,27</sup> The time frame(s) of the prevalence estimates will be driven by the data themselves and have therefore not yet been finalized.

#### Survival

The survival methods described for mortality analyses will also be used to measure cancer survival, despite debate as to the superiority of nonparametric versus parametric methods<sup>28</sup> in generating efficient,<sup>29</sup> unbiased survival estimates.<sup>17</sup>

Net survival is sometimes also used as a secondary measure of survival. Broadly, this approach produces a survival estimate that “might occur if cancer was the only cause of death.”<sup>30</sup> Cause-specific net survival is strongly dependent on the completeness and consistency of registry records. In the case of the Canadian Cancer Registry, completeness is not a concern, but consistency may be, as records are collated from different provincial and territorial registries, where different cause of death selection and coding rules may prevail. Consequently, we will also calculate relative net survival, when relevant,

using the Stata “stns” command,<sup>31</sup> which employs the Pohar-Perme estimator approach to calculating net survival.<sup>32</sup> There is substantial evidence that this approach provides a less biased estimation of survival than other established approaches,<sup>30</sup> although this varies with the length of follow-up.

Published evidence suggests that the inclusion of subsequent primary tumours in survival analyses differentially reduces survival estimates; the magnitude of this reduction depends on both the type of cancer in question and the proportion of multiple primary tumours.<sup>33</sup> As with any decisions requiring clinical and epidemiological considerations, we will seek external expertise before analysis.

### Regression models

In addition to the age–period–cohort regression models described in the mortality analysis section, we will consider Joinpoint models<sup>34</sup> when building predictive cancer incidence models, given that they are commonly used<sup>35</sup> by the National Cancer Institute Surveillance, Epidemiology, and End Results (SEER) Program<sup>36</sup> and the International Agency for Research on Cancer.

### Missing values and sensitivity analyses

Although missing data can be challenging, the CFCAMS II will rely on administrative data that are collected with the use of stringent data-completeness standards. This, combined with the high quality of the pay data (used to create the study cohort), means that broad, systematic missing data requiring a management protocol are not likely to be an issue.

A more likely source of missing data will be deaths and cancer diagnoses that took place outside of Canada. To address the mortality gap, we have contracted Statistics Canada to use tax records to identify people who died outside of Canada. Although no cause of death will be available, this will allow us to determine the magnitude of the issue and to use this information to conduct sensitivity analyses on reported mortality. We will also conduct cancer incidence sensitivity analyses, using the proportion of cancer incidence to deaths and the proportion of reported to nonreported deaths. In both cases, different scenarios will be evaluated to provide a range of levels of sensitivity.

### Dissemination

The results from the study will be disseminated within the involved organizations as well as to the general public, through publication and dissemination in peer-reviewed journals.

### Ethics and privacy

The CFCAMS II protocol was reviewed and approved by an external institutional review board (Quorum Review, reference no. QR#31460CDN/1). A consent waiver was requested and was approved by the review board, which allows the study to be conducted without the individual consent of all participants (per Article 3.7 of the *Canadian Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans*<sup>37</sup>). Additional redundant mechanisms will be implemented to ensure that study population privacy and data are rigorously protected (see section VI of Appendix 1).



## Interpretation

The CFCAMS II aims to describe the mortality and cancer experience of Canadian Armed Forces personnel using a longitudinal record-linkage design. The study will supplement existing Canadian Armed Forces and Veterans Affairs Canada health surveillance capacity to describe the burden of mortality and cancer incidence in all Regular Force and Class C Reserve Force personnel who began serving in the forces between 1976 and 2015, regardless of their current military status, and begin to investigate potential links between excess burden and military service.

The nearly 40 years of follow-up time may provide sufficient lead-up time for conditions with delayed expression following exposure (e.g., certain cancers). In addition, the large sample (about 230 000 people contributing more than 5 million person-years) may provide sufficient statistical power to investigate less common outcomes. The living nature of this study is also an important strength, as new Canadian Armed Forces members, new occupational data on existing cohort members and new outcome data for all will be periodically appended to the current cohort.

## Conclusion

This study has the potential to provide novel and sound evidence on the risks and protective factors of military life. The strengths of this study make it potentially groundbreaking, particularly as it relates to evaluating the relation between military service and adverse health outcomes. The body of evidence emanating from this study will allow for the development of effective policies and programs for promoting, protecting and caring for the health of Canada's airmen, airwomen, soldiers and sailors throughout their life courses and will provide evidence that may also benefit our allied military forces.

## References

1. Rolland-Harris E, Maher M. Surveillance of mortality and cancer morbidity in the Canadian Armed Forces: a life-course approach [oral presentation]. STO-HFM-254 Military Health Surveillance Symposium, Paris, 2015 Oct. 12–14.
2. *The Canadian Persian Gulf cohort study: summary report*. Ottawa: Statistics Canada; 2005. Cat no 82-580-XWE.
3. Birkett N, Brodsky L. *Gulf War Veterans study protocol*. Ottawa: Department of National Defence; 2001.
4. Rolland-Harris E. *2018 report on suicide mortality in the Canadian Armed Forces (1995 to 2017)*. Ottawa: Department of National Defence; 2018.
5. Rolland-Harris E. More than just counting deaths: the evolution of suicide surveillance in the Canadian Armed Forces. *Mil Med*. In press.
6. Statistics Canada. Canadian Forces Cancer and Mortality Study: causes of death. Cat no 82-584-X. Ottawa: Minister for Industry; 2011. Available: <https://www150.statcan.gc.ca/n1/pub/82-584-x/2011001/inf-eng.htm> (accessed 2018 Nov. 26).
7. Israel G. Determining sample size. Gainesville (FL): University of Florida; 1992.
8. *Statistics: power from data!* Cat no 12-004-X. Ottawa: Statistics Canada; 2001.
9. Working Group Report. International rules for multiple primary cancers (ICD-0 third edition). *Eur J Cancer Prev* 2005;14:307-8.
10. Ellison LF, Wilkins K. Cancer prevalence in the Canadian population. *Health Rep* 2009;20:7-19.
11. Alexander LK, Lopes B, Ricchetti-Masterson K, et al. Calculating person-time. In: *UNC ERIC Notebook*. 2nd ed. Chapel Hill (NC): University of North Carolina; 2015:1-3.
12. Goldman A, Brender JD. Are standardized mortality ratios valid for public health data analysis? *Stat Med* 2000;19:1081-8.
13. Rau R, Bohk-Ewald C, Muszynska MM, et al. *The Lexis diagram. Visualizing mortality dynamics in the Lexis diagram*. New York: Springer; 2018.
14. Bakker B, Heskes T, Neijt J, et al. Improving Cox survival analysis with a neural-Baysian approach. *Stat Med* 2004;23:2989-3012.
15. Compton CC, Byrd DR, Garcia-Aguilar J, et al., editors. *Cancer survival analysis. AJCC cancer staging atlas, 2nd edition: a companion to the seventh editions of the AJCC cancer staging manual and handbook*. New York: Springer; 2012.
16. Bewick V, Cheek L, Ball J. Statistics review 12: survival analysis. *Crit Care* 2004;8:389-94.
17. Xue X, Hoover DR. Statistical methods in cancer epidemiological studies. *Methods Mol Biol* 2009;471:239-72.
18. Tsiatis AA. Nonidentifiability aspect of the problem of competing risks. *PNAS* 1975;72:20-2.
19. Clayton D, Schifflers E. Models for temporal variation in cancer rates I: age-period and age-cohort models. *Stat Med* 1987;6:449-67.
20. Rutherford MJ, Lambert PC, Thompson JR. Age-period-cohort modeling. *Stata J* 2010;10:606-27.
21. Carstensen B. Age-period-cohort models for the Lexis diagram. *Stat Med* 2007;26:3018-45.
22. Royston P, Lambert PC. *Flexible parametric survival analysis using Stata: beyond the Cox model*. College Station (TX): Stata Press; 2011.
23. Callas PW, Pastides H, Hosmer DW. Empirical comparisons of proportional hazards, Poisson, and logistic regression modeling of occupational cohort data. *Am J Ind Med* 1998;33:33-47.
24. Das A. Cancer registry databases: an overview of techniques of statistical analysis and impact on cancer epidemiology. *Methods Mol Biol* 2009;471:31-49.
25. Breslow NE, Day NE, editors. *Statistical methods in cancer research: volume II — the design and analysis of cohort studies*. Lyon (FR): International Agency for Research for Cancer; 1987.
26. Crocetti E, Buiatti E, Falini P. Multiple primary cancer incidence in Italy. *Eur J Cancer* 2001;37:2449-56.
27. de Oliveira C, Weir S, Rangrej J, et al. The economic burden of cancer care in Canada: a population-based cost study. *CMaj Open* 2018;6:E1-10.
28. Miller RGJ. What price Kaplan-Meier? *Biometrics* 1983;39:1077-81.
29. Meier P, Karrison T, Chappell R, et al. The price of Kaplan-Meier. *J Am Stat Assoc* 2004;99:890-6.
30. Roche L, Danieli C, Belot A, et al. Cancer net survival on registry data: use of the new unbiased Pohar-Perme estimator and magnitude of the bias with the classical methods. *Int J Cancer* 2013;132:2359-69.
31. Clerc-Urmès I, Grzebyk M, Hédelin G. Net survival estimation with stns. *Stata J* 2014;14:87-102.
32. Perme MP, Stare J, Estève J. On estimation in relative survival. *Biometrics* 2012;68:113-20.
33. Rosso S, De Angelis R, Ciccolallo L, et al. Multiple tumours in survival estimates. *Eur J Cancer* 2009;45:1080-94.
34. Kim HJ, Fay MP, Feuer EJ, et al. Permutation tests for Joinpoint regression with applications to cancer rates. *Stat Med* 2000;19:335-51.
35. Arnold M, Sierra M, Laversanne M, et al. Global patterns and trends in colorectal cancer incidence and mortality. *Gut* 2017;66:683-91.
36. Joinpoint Regression Program. Washington: National Cancer Institute; 2018 [updated 2018 Apr. 10]. Available: <http://surveillance.cancer.gov/joinpoint/> (accessed 2018 Nov. 26).
37. *Tri-Council policy statement: ethical conduct for research involving humans*. Ottawa: Canadian Institutes for Health Research, Natural Sciences and Engineering Research Council of Canada, Social Sciences and Humanities Research Council of Canada Interagency Secretariat on Research Ethics; 2010.

**Affiliations:** Directorate of Force Health Protection (Rolland-Harris, Weeks, Simkus) and Directorate Mental Health (Zamorski, Boulos), Department of National Defence, Ottawa, Ont.; Research Directorate (VanTil, Simkus), Veterans Affairs Canada, Charlottetown, PEI; Statistics Canada (Reicker, Masoud, Trudeau), Ottawa, Ont.

**Contributors:** Elizabeth Rolland-Harris participated in the design of the study and drafted the manuscript. Linda VanTil, David Boulos, Mark Zamorski, Alexander Reicker, Huda Masoud, Richard Trudeau, Murray Weeks, and Kristen Simkus participated in the design of the study and helped to draft the manuscript. All of the authors gave final approval of the version to be published and agreed to be accountable for all aspects of the work.

**Supplemental information:** For reviewer comments and the original submission of this manuscript, please see [www.cmajopen.ca/content/6/4/E619/suppl/DC1](http://www.cmajopen.ca/content/6/4/E619/suppl/DC1).