CMAJ OPEN
MEDICAL RESEARCH FOR ALL

The Canadian Forces Cancer and Mortality Study II: A Longitudinal Record-Linkage Study Protocol

Journal:	*CMAJ Open*
Manuscript ID	CMAJOpen-2017-0125
Manuscript Type:	Cohort (retrospective)
Date Submitted by the Author:	14-Sep-2017
Complete List of Authors:	Rolland-Harris, Elizabeth; Department of National Defence; VanTil, Linda; Veterans Affairs Canada, Research Directorate Zamorski, Mark; Department of National Defence, Directorate of Mental Health Boulos, David; Department of National Defence, Directorate of Mental Health Reicker, Alexander; Statistics Canada Masoud, Huda; Statistics Canada Trudeau, Richard; Statistics Canada Weeks, Murray; Department of National Defence, Directorate of Force Health Protection Simkus, Kristen; Department of National Defence, Directorate of Force Health Protection; Veterans Affairs Canada, Research Directorate
Keywords:	Epidemiology, Health services research, Occupational health, Statistics and research methods
More Detailed Keywords:	mortality, neoplasms, military, Veterans, longitudinal studies, cohort studies

Abstract:	Background Military service exposes personnel to unusual situations with unclear health-related implications. In order 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. Cancer and Mortality Study II has been developed with the goal to describe the mortality and cancer experience of approximately a quarter of a million current and former Canadian Armed Forces personnel, to inform health promotion and prevention programs for serving personnel, and services for Veterans after leaving the military. Methods

	Record linkage methods were identified as the most appropriate mechanism to study mortality and cancer in Canada. The Canadian Forces Cancer and Mortality Study II is a retrospective cohort study of serving and released Canadian Armed Forces personnel who enrolled on or after January 1, 1976, in the Regular Force and Reserve Force "C" . Statistics Canada provided a highly secure linkage environment that linked this military cohort file with the Canadian Vital Statistics database that includes cause of death by ICD code for deaths up to December 31, 2012. Linkage to the Canadian Cancer Registry Database is expected in 2018, and will include cancer diagnoses up to December 31, 2013. Linkage to the mortality data used a hierarchical deterministic record linkage with a final linkage rate of 99.7%. Analysis will compare the military cohort to the Canadian general population, using Standardized Mortality/Incidence Ratios. Analysis will also use an internal reference population, such as comparison by military occupation or deployment.

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page No.	Relevant text from manuscript
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	4 (in submitted pdf)	Longitudinal Record-Linkage Study
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	5	
Introduction				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	6-8	
Objectives	3	State specific objectives, including any prespecified hypotheses	8	
Methods				
Study design	4	Present key elements of study design early in the paper	9	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	9	
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	9-10	
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	N/A	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	10	Data elements that were included in the cohort file for submission for linkage to STC included demographic information (name, sex, date of birth, social insurance number

				[SIN]), occupational data (rank, enrolment and release date[s], command [Regular or Reserve C], element [Army, Navy, Air Force]), and deployment and foreign posting data (including location and start and stop dates). Multiple enrolments and releases (if relevant) were also captured.
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	9	
Bias	9	Describe any efforts to address potential sources of bias	10; 15-16	(Pg. 10: Issues around exclusion of Reservists A & B; Pg. 15-16; Changes in ICD-coding over time)
Study size	10	Explain how the study size was arrived at	N/A	Study protocol

Continued on next page

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	N/A	Study protocol
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	N/A	Study protocol
		(*b*) Describe any methods used to examine subgroups and interactions	N/A	Study protocol
		(*c*) Explain how missing data were addressed	N/A	Study protocol
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	N/A	Study protocol
		(*e*) Describe any sensitivity analyses	N/A	Study protocol
Results				
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	N/A	Study protocol
		(b) Give reasons for non-participation at each stage	N/A	Study protocol
		(c) Consider use of a flow diagram	24	Describes building of cohort
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	N/A	Study protocol
		(b) Indicate number of participants with missing data for each variable of interest	N/A	Study protocol
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	N/A	Study protocol
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	N/A	Study protocol
		Case-control study—Report numbers in each exposure category, or summary measures of exposure		
		Cross-sectional study—Report numbers of outcome events or summary measures		
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	N/A	Study protocol
		(*b*) Report category boundaries when continuous variables were categorized	N/A	Study protocol
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A	Study protocol

Continued on next page

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A	Study protocol
Discussion				
Key results	18	Summarise key results with reference to study objectives	N/A	Study protocol
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	N/A	Study protocol
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	N/A	Study protocol
Generalisability	21	Discuss the generalisability (external validity) of the study results	N/A	Study protocol
Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	10	CAF Surgeon General Health Research Fund

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

The Canadian Forces Cancer and Mortality Study II: A Longitudinal Record-Linkage Study Protocol

Elizabeth Rolland-Harris, MSc PhD[1]
Linda VanTil, DVM MSc[2]
Mark A. Zamorski, MD MHSA[3]
David Boulos, MSc[3]
Alexander Reicker, MSc[4]
Huda Masoud, MSc[4]
Richard Trudeau, BSc[4]
Murray Weeks, PhD[1]
Kristen Simkus, MPH[1,2]

[1] Directorate of Force Health Protection, Department of National Defence, 101 Colonel By Drive, Ottawa, ON K1A 0K6, Canada
[2] Research Directorate, Veterans Affairs Canada, 161 Grafton Street, Charlottetown, PE C1A 8M9, Canada
[3] Directorate Mental Health, Department of National Defence, 101 Colonel By Drive, Ottawa, ON K1A 0K6, Canada
[4] Statistics Canada, 150 Tunney's Pasture Driveway, Ottawa, ON K1A 0T6, Canada

Corresponding author:
Dr. Elizabeth Rolland-Harris
Senior Epidemiologist
DFHP/DND
Carling Campus
101 Colonel By Drive
Ottawa, ON K1A 0K6
Canada
Email: Elizabeth.harris@forces.gc.ca
Tel: +1 613-219-4477

ABSTRACT

Background

Military service exposes personnel to unusual situations with unclear health-related implications. In order 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. Cancer and Mortality Study II has been developed with the goal to describe the mortality and cancer experience of approximately a quarter of a million current and former Canadian Armed Forces personnel, to inform health promotion and prevention programs for serving personnel, and services for Veterans after leaving the military.

Methods

Record linkage methods were identified as the most appropriate mechanism to study mortality and cancer in Canada. The Canadian Forces Cancer and Mortality Study II is a retrospective cohort study of serving and released Canadian Armed Forces personnel who enrolled on or after January 1, 1976, in the Regular Force and Reserve Force “C” . Statistics Canada provided a highly secure linkage environment that linked this military cohort file with the Canadian Vital Statistics database that includes cause of death by ICD code for deaths up to December 31, 2012. Linkage to the Canadian Cancer Registry Database is expected in 2018, and will include cancer diagnoses up to December 31, 2013. Linkage to the mortality data used a hierarchical deterministic record linkage with a final linkage rate of 99.7%.

Analysis will compare the military cohort to the Canadian general population, using Standardized Mortality/Incidence Ratios. Analysis will also use an internal reference population, such as comparison by military occupation or deployment.

BACKGROUND

The concept of the healthy worker is a well-documented and generally accepted occurrence (1). However, many occupations and industries place their workers at an increased risk for occupational illness and injury, potentially eroding the healthy worker effect over time, sometimes to the point where employment is directly related to negative health outcomes. For instance, occupation has been causally linked to premature mortality in coal miners (2-4), pulp and paper mill workers (5), farmers (6, 7), firefighters (8, 9) and factory workers (10, 11). While negative health outcomes can include cancer, accidents and other conditions that may result in premature death, the health risks often do not manifest until long after someone has stopped working.

In light of the potential exposure to work-related health risks, employers, especially those maintaining high-risk occupations, have a responsibility to implement health surveillance, and for a number of purposes, including documentation, prevention and control of negative health exposures and outcomes. As high-risk employers, military organisations have a particular obligation to adhere to these tenets of health surveillance, as their personnel have historically been exposed to unusual situations with unclear health-related implications. To date, the broader body of military evidence in this area has focused either on specific causes of death or on specific deployments. In some cases, specific exposures of concern were identified; examples include the use of Agent Orange during multinational Vietnam War deployments (12-16) and on Canadian bases, depleted uranium (17, 18), and fires during the first Gulf War (19). However, in other cases, unidentified and/or non-specific exposures may occur during the tenure of a military member's career, with a link only being made many years after the fact, in large part owing to health surveillance and longitudinal cohort studies. Examples of the latter type of health outcomes to date have included those in U.S. Gulf War Veterans (especially accident-related mortality (20, 21), and lung cancer (22), as well as testicular cancer specifically in electronic equipment repair personnel (23)), New Zealand Vietnam War Veterans (where an increased risk of head, neck, oral, pharynx and larynx cancers, as well as all leukaemias was identified (24)), and Norwegian UN peacekeepers in Kosovo (increased incidence of bladder cancer) (25).

As a military organisation, the Canadian Armed Forces (CAF) has the challenge of balancing its defence mandate responsibilities with its duty to protect its personnel. These two responsibilities can sometimes be at odds, particularly when there are challenges to both national and international security. Military personnel may be involved in peacekeeping and observer missions, post-conflict peace building, humanitarian assistance, and, when required, responses to aggression. The very nature of these operations can be accompanied by unusual exposures with both expected and unknown risks to personnel. Adverse outcomes, including death, may occur. To identify both immediate and delayed risks, the CAF and Veterans Affairs Canada (VAC) require a disease and injury surveillance system that extends beyond periods of military service.

Health Surveillance Capacity in the CAF

In recognition of the importance of health surveillance in military personnel, the CAF has included cancer and mortality surveillance as part of its health surveillance strategy for over 15 years. The evolution of this strategy is ongoing, but its approach has been shifting to a focus on health surveillance from a life-course perspective, whereby the importance of what precedes enrolment in the CAF is not diminished or ignored, and where the long-term effects of military service beyond the end of career are not discredited or devolved to the civilian sector. The sources of data within the CAF, their focus, and how they fit within a military individual's life course are described in more detail elsewhere (26).

Mortality and cancer incidence are standard health indicators that are regularly analysed and published by national organisations such as Statistics Canada (STC) and the Canadian Institute for Health Information (CIHI). However, neither the Department of National Defence (DND) nor VAC maintain comprehensive information on deaths or cancers among CAF personnel throughout their lifespan, limiting the ability of these departments to identify both the immediate and delayed effect of various risk exposures. In Canada, the most complete records of death and cancer are maintained by provincial and territorial authorities in their vital statistics and cancer registries; however, these data sources do not include unique identifiers for individuals with current or past military service. This necessitates the use of record linkage methods to associate military service exposure with such outcomes.

Cluster Investigations

The need for record linkage studies arose in response to limitations in the availability of data for occupational health investigations. Since 2000, DNDs Directorate of Force Health Protection (DFHP) has conducted three cancer cluster investigations that responded to CAF personnel's concerns about possible occupational exposures.
The first cluster involved a group of personnel deployed to Camp Doha (Persian Gulf) in April 1991, to assist in the clean-up and rebuilding after active fighting ended. The second CAF cluster investigation responded to concerns of elevated Chronic Lymphocytic Leukemia among air traffic controllers (in 2002), while the third and final cluster investigation focused on the Non-Hodgkin's Lymphoma among aircrew of the Aurora aircraft (in 2004).

In all three cases, these investigations made use of all data available at DND but were limited their ability to capture all cancer cases among serving personnel and in the lack of a method for monitoring cancer among individuals who left the CAF.

Record Linkage Studies

To date, two record linkage studies have been conducted by DND, in conjunction with STC and VAC. The first study, the Gulf War Veterans Record Linkage Study, was conducted to ascertain the 9-year mortality and cancer experience of 5,100 CAF personnel deployed during the first Gulf War conflict (27). The overall findings demonstrated that, compared with a random sample of non-deployed CAF personnel,

there was no significant difference in the overall risk of death, or in deaths due to suicide or motor vehicle accidents. In addition, there was no significant difference in having been diagnosed with cancer. A preceding feasibility study had predicted that the number of events expected during the follow-up period would be small due to the young age of the deployed military personnel, the relatively small size of the cohort and the short follow-up time (28).

In 2010-2011, STC, DND and VAC conducted the CF Cancer and Mortality Study (CF CAMS), a record linkage study that looked at mortality outcomes for all CAF personnel enrolled between 1972 and 2006. Preliminary results suggested that, overall, both women and men who had served and/or were still serving were at significantly lower risk of dying compared to their civilian counterparts (Male all-cause mortality [ACM] Standardized Mortality Ratio [SMR]: 0.64 [0.62, 0.66]; Female ACM SMR: 0.67 [0.59, 0.75]). Exceptions to this pattern were noted amongst men dying as a result of air space accidents (SMR: 2.59 [2.1, 3.16]). However, due to important methodological limitations directly related to quality issues with the data used to build the original CF CAMS cohort (Human Resources Management System [HRMS]), in particular with the definition and identification of Reserve Force personnel and in the calculation of years of service and time since release, further analyses were not conducted.

CF Cancer and Mortality Study II

The CAF's latest record linkage study [CF Cancer and Mortality Study II (CF CAMS II)] has been developed to address the gaps outlined above and its goal is to describe the mortality and cancer experience of CAF personnel (serving and released) in order to inform:

- Health promotion and the DFHP's policies and programs for serving personnel
- The need for services that provide care for Veterans and their families after leaving military service.

The specific objectives are to:

- Describe the overall and cause-specific mortality and cancer incidence rates among CAF personnel (serving and released);
- Evaluate the overall and cause-specific mortality and cancer incidence risk among CAF personnel (serving and released) relative to the general Canadian population;
- Determine if overall and cause-specific mortality and cancer incidence patterns vary by subject characteristics such as age, sex, rank, military occupation, deployment history, employment history, element (Army, Navy, Air Force), duration of employment, duration since release, and identified occupational exposures.

This study protocol describes the methods used to conduct this record linkage study linking the Canadian Vital Statistics Database (CVSD) and Canadian Cancer Registry Database (CCR) to pay data for all Regular Force and Class C Reservist personnel enrolled by the CAF between 1976 and 2015, inclusive, in partnership with STC and VAC. The study was funded by the CAF Surgeon General Health Research Fund.

METHODS/DESIGN

1. Defining the Cohort

The CF CAMS II is a retrospective cohort study of serving and released CAF personnel who enrolled in the CAF on or after January 1, 1976. The 1976 starting point was chosen due to the availability of administrative information in the Central Computerised Pay System (CCPS), an electronic DND dataset with information available beginning in 1976. The CCPS only includes Regular Force and Reserve Force "C" (Reservists on international operations and/or tours of duty) members. CCPS data have an accuracy and precision advantage over HRMS data when used to identify membership in the CF CAMS II cohort and other information associated with remuneration. CCPS data have a built-in feedback mechanism where both service members and the employer are motivated to rectify remuneration errors (e.g. not being paid the extra deployment-associated pay for the correct deployment dates) as soon as possible. This creates an informal data validation process that did not exist for CF CAMS I.

Reservists in CF CAMS II Cohort

CAF Reservists combine a full-time civilian career or higher education pursuit while devoting a portion of their time to military training and military service. Reservists are classified as "A" (serving on a part-time basis with durations of less than 12 consecutive days a year; usually one evening a week and one weekend a month), "B" (full-time contractual basis, usually within Canada), and "C" (full-time; either in a Regular Force position, or as part of contingency, routine, domestic or international operations) (29). Only Class C reservists are remunerated through the CCPS, similar to Regular Force personnel. Reserve A and B remuneration is managed through a different system (Reservist Pay System); this system is fraught with inaccuracies, in large part due to a high turnover of Class A reservists, as well as the lack of an established and agreed upon definition of what defines a reservist as "active," (30). This system has also been in place for a much shorter period of time than the CCPS. As such, use of the Reservist Pay System to identify additional reservist (A or B) cohort members and military service exposure would raise a number of concerns surrounding the completeness of such information. Therefore, the comparatively much lower exposure to military culture experienced by Reserve A and B personnel relative to that of Regular Force (and Reserve C) personnel indicates that quantifying this low exposure would be difficult, fraught with error and have little impact on the outcomes to be measured.

Given these factors, we elected to exclude reservists A and B. We acknowledge that this may under-estimate the effect of exposure to military culture on adverse health

outcomes, but accept that this is a compromise that must be made in order to maximize the quality and feasibility of this study.

2. Creating the Cohort File

Using data extracted from the DND CCPS, epidemiologists from the DFHP at DND created a cohort file (Figure 1). This cohort file was supplemented and validated using data from HRMS, thereby minimizing missing information and resolving imprecise or, apparently, anomalous information (e.g. unrealistic birth dates). The access and use of these data files was considered to be consistent with the purpose for which these data were collected and maintained (i.e. “consistent use”) under the Canadian Government InfoSource programme, and their authorized use was vetted by a Senior Privacy Officer and approved by an external, accredited Research Ethics Board (QUORUM Review IRB).

Data elements that were included in the cohort file for submission for linkage to STC included demographic information (name, sex, date of birth, social insurance number [SIN]), occupational data (rank, enrolment and release date[s], command [Regular or Reserve C], element [Army, Navy, Air Force]), and deployment and foreign posting data (including location and start and stop dates). Multiple enrolments and releases (if relevant) were also captured. Some of the aforementioned variables were included in the cohort file solely to facilitate the linkage process (e.g. name, SIN) and were stripped by STC prior to delivering the final cohort file for analysis.

3. Linking the Data

STC’s Social Data Linkage Environment (SDLE) is a highly secure linkage environment that facilitates the creation of linked population data files for social analysis. At the core of the SDLE is a Derived Record Depository (DRD), a national dynamic relational database containing only basic personal identifiers. The DRD is created by linking selected STC source index files for the purpose of producing a list of unique individuals. These files, which contain personal identifiers without analysis variables, are brought into the environment, processed and linked to the DRD. Updates to these data files are linked to the DRD on an ongoing basis.

All source index files are linked to the DRD either deterministically or probabilistically. Deterministic record linkage involves matching records based on unique identifiers shared by both files. Probabilistic record linkage works with non-unique identifiers (e.g. names, sex, date of birth, and postal code) and estimates the likelihood that records are referring to the same entity. The record linkage results, i.e., the association of the source index file identifiers and the DRD identifier referring to the same entity. The record linkage results, i.e., the association of the source index file identifiers and the DRD identifier referring t the same entity, are stored in a Key Registry.

Once a study requiring linked data has been defined and approved, the associated record identifiers from the Key Registry are used to find the individual records in the source data files. These data files contain analysis variables without personal identifiers.

Both the CF CAMS II cohort file and the DRD contained the social insurance number (SIN), which is a unique identifier assigned by Employment and Social Development Canada (ESDC) to Canadian citizens, permanent residents and temporary residents as a requirement to work in Canada or to receive benefits and services from government programs, including the Canada Revenue Agency (CRA). All SDLE linkages preserve the uniqueness of SIN on the DRD so that no two people have the same SIN. Although rare, a person may have multiple SINs over their lifetime. A SIN is required for work in Canada, although people who have never worked, particularly young children, may not have been assigned a SIN. However, since the CF CAMS II cohort file included only members of the Canadian Forces, everyone in the population was expected to have been assigned a SIN.

Given the high proportion of cohort records with a SIN, the cohort file was to be linked to the DRD using a hierarchical deterministic linkage that relied on SIN, names, date of birth and sex.

4. Establishing the Denominators and a Comparator Population

Any individual whose initial enrolment in the CAF occurred in 1976 or later will generate person-years of experience within the cohort; this includes individuals whose first enrolment was in 1976 or later but released and subsequently re-enrolled. These individuals are assumed to be alive or cancer-free until they appear in either the CVSD or the CCR, respectively. Both databases are high quality sources of information for mortality and cancer events occurring in Canada; such events occurring outside Canada are less reliably captured. More details on the quality of the CVSD and CCR are provided below.

In order to generate standardized ratios (mortality [SMR] or incidence [Standardized Incidence Ratio (SIR)]), there is a requirement for an external population against which the study population can be compared. Applying the rate for a condition in the reference population, the observed rate in the study population (in this case the CF CAMS II cohort) can be compared to the expected rate, should the study population behave in the same manner as the reference population.

SMRs and/or SIRs are commonly found in the occupational epidemiology literature. Many of these studies remarked on the typically lower disease and mortality experience of employed populations as compared with the general population – the so-called “healthy worker effect”. The healthy worker effect refers to the confounding created when comparing health outcomes of occupational populations with those of general populations because those who gain and maintain employment will tend to be healthier (31). Therefore, an occupational cohort is expected to report lower SMRs or SIRs compared to the general population; ideally, the use of a more relevant external comparison population is a consideration. However, obtaining outcome data for an

employed population with the same decades of follow-up is a challenge. Of even greater concern is that, depending on the make-up of the population, it may have elevated risks due to occupational exposures that will impair detection of similarly elevated risks for diseases in the CAF population. Comparison with a military population of another country has the same problems with the additional concern that it will have many of the characteristics of the source population so that any differences observed may actually reflect differences in the national populations of the two countries.

Given these limitations, two approaches were adopted in comparing the CF CAMS II cohort:

1) Compare CF CAMS II cohort to the Canadian general population (CGP) for more general analyses (e.g. ICD chapter-level analysis of mortality burden), being cognizant that unless there is a massively elevated risk, SMRs or SIRs will still be below 1.0 and may mask truly elevated rates;

2) Use an internal reference population in which a group of workers defined by a common experience are compared with another group of workers without that experience. This approach is "typical of analyses that are primarily directed at exploring associations of disease rates with certain work areas, tasks, or exposure levels within the workplace," (32). For the CAF, the most obvious approach would be to compare individuals grouped by occupational classifications or those deployed versus non-deployed.

The latter approach reduces the healthy worker effect, but still needs to consider controlling for confounders such as employment status, time since first employment, length of follow-up, etc.

5. Complete Ascertainment of Mortality

CVSD Data Elements

The CVSD provides mortality information from 1950 onwards to the latest available year of data. Information from the CVSD is provided by the 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. They share their information (under the Statistics Act) with STC for national reporting, and to allow linkage for approved projects. Linkage for this study occurred up to the last year of available data (2012) for mortality outcomes.

As part of the linked file, the following mortality variables were requested:

- date of birth (month and year)
- sex
- province/country of birth
- province of residence
- full date of death

- province/country of death
- underlying cause of death (ICD code);
- all causes of death (for deaths 2000 onwards only, due to CVSD holdings)
- autopsy code

CVSD Data Quality

All causes of death are classified according to the World Health Organisation (WHO) International Statistical Classification of Diseases and Related Health Problems (ICD).

In addition to the quality control that occurs at the provincial/territorial level before the data are given to STC, STC also conducts quality control on the mortality data before they are publically released (33). Firstly, verification tables with univariate and bivariate frequencies, stratified by province/territory, are generated for the majority of the captured variables. These tables are then sent to the provincial/territorial registrars for review and approval. The data are also checked for internal consistencies, and are compared to previous years to detect unusual data patterns or changes.

In terms of data completeness, it was estimated that the data coverage is fairly complete, primarily because of the legislative requirement to report deaths. However, late or missing death registrations can occur, primarily in the event of unidentified bodies, deaths to Canadians that occur outside of Canada, and CAF deaths that occur on deployment or posting (33). The estimated magnitude of this problem was not provided, but it was assumed to be a fairly systematic underestimate across all years.

CAF Out-of-Country Deaths

Among individuals with lengthy follow-up, there is greater potential for them to have left Canada after their release from the CAF. If they develop cancer or die outside Canada, that information may not make its way back into the STC databases. In such circumstances, the event does not get captured in the numerator of rate calculations, but that individual continues to contribute to the person-year denominator, thereby reducing the rate. The relative importance of this is dependent on the frequency with which CAF personnel relocate outside Canada post-release.

The CAF does not track movement of individuals post-release so no internal files will assist with this task. However, it was not perceived that released CAF personnel are any more likely to move out of the country than the general population. Fortunately, STC has the ability to assess the likelihood of someone being alive using an Alive Tax File, which contains information about the years individuals have filed a tax return, whether there was notification of a cause of death for the filer, or whether the filer moved out of the country. Information is available for individuals since 1984. In addition, a more comprehensive mortality database maintained by DFHP that includes chart reviews for all CAF Regular Force deaths from 2004 onwards. The findings from this database were shared with STC and were used to supplement the mortality data in the linked cohort data. Although this did not cover the complete cohort period, it did include the years of the Afghanistan involvement where excess mortality outside of Canada was noted.

6. Complete Ascertainment of Cancer

CCR Data Elements

The CCR contains cancer incident events from 1969 onwards to the latest available year of data (Canadian Cancer Registry Team, currently 2012). All new incident cases of cancer are recorded by Cancer Registries in the provinces and territories. Individuals can have more than one record in the CCR depending upon how many cancers have been reported or how many registrations have been submitted for that individual. Cancer incidence information in the CCR 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 STC for national reporting, and to allow linkage for approved projects. Linkage for this study will occur up to the last year of available data, expected to include dates of diagnosis up to December 31, 2013 by the time the cancer linkage is completed in 2017/18.

The following variables for all malignant tumours were requested for this study linkage:

- province of residence at date of diagnosis
- date of birth (month and year)
- sex
- province or country of birth
- province of diagnosis
- date of diagnosis (day, month, year)
- patient vital status
- date and province of death (if applicable)
- diagnostic information (ICD codes for morphology and topography, method of diagnosis, laterality)

A complete list of variables of interest is provided in Table 1.

CCR Data Quality

The responsibility to control the quality of the cancer incidence data befalls the provinces and territories that provide the data to STC for inclusion within the CCR. However, the data are provided in a standard format to STC, allowing for the use of validation edits to ensure that each captured field only includes valid entries. In addition, correlation edits are also run so as to ensure that relationships between patient and tumour records are coherent. Any recorded errors are sent back to the provincial/territorial registrar for verification and/or correction. Other data control methods used by the CCR are described in more detail elsewhere (34).

7. Changes in ICD-coding over time

The time period covered by the linked cohort data overlaps three discrete ICD coding periods:

- ICDA-8 used from 1969 to 1978

- ICD-9 used from 1979 to 1999
- ICD-10 used from 2000 to present (35).

For chapter-level analyses, ICD-8 and ICD-9 chapters will be recoded to their ICD-10 equivalent, using a number of Canadian (36) and international (37) sources. For cause-specific analyses, changes in definition and coding will be investigated and taken into consideration prior to any analysis. Specialists at STC will also be consulted should there be a lack of evidence to direct the decision-making process.

8. Military Service Variables

Several occupational variables are essential explanatory variables to the analysis of cancer and mortality events for military personnel. The following explanatory variables will be used and/or derived in the linked cohort file:

- Enrolment date(s)
- Release date(s)
- Reason for release
- Person-years of service (start and stop dates for Regular and Reserve Force Class C only)[1]
- Regular/Reserve "C" Force history – start-stop dates
- Occupational history – Military occupation codes (MOC), Military Occupation Structure Identification (MOSID) codes and dates
- Foreign posting history - locations, start-stop dates for each foreign posting
- Deployment history - location and start-stop dates of each deployment
- Element (Army, Navy, Air Force)
- Rank history - start and stop dates for each rank

For military occupations, the standardized coding (MOCs and/or MOSID) will be used to create similar occupational exposure categories or groupings that will be more meaningful in analyses. Recognizing that individuals can change their classification over time is critical (e.g. a medic might become a search and rescue technician). In theory, there are a variety of ways in which analyses on occupation could be conducted and the most relevant may depend on the study question: a) first job; b) last job; c) longest-held job; or d) most hazardous job. In discussing these options, Checkoway et al. (2004) favour item "c" and this is the most commonly used approach. Options "a" and "b" waste information and will ignore important histories of exposure. Option "d" requires investigator judgment regarding types and extent of exposures in different jobs. The mode of categorizing occupations for CF CAMS II will depend on the study question and will be adjusted accordingly for specific analyses. For example, analyses of suicide

[1] Data from CCPS provides the enlistment and release dates from the CAF Regular Force. For the purposes of CF CAMS, person-years of service will be compiled to capture individuals who are released and then subsequently re-enlist, which is a relatively common occurrence for military personnel.

mortality or deaths from motor vehicle accidents may require last known MOC while occupational investigations looking at cancer outcomes, may require longest held job, most hazardous job, or occupation of interest for a given occupational exposure.

Deployment as an explanatory variable is arguably one of the more difficult ones to define. Deployment is a term used to broadly refer to heterogeneous movements of military personnel. Deployments include different mission types (peacekeeping [e.g. Rwanda, Kosovo], active combat [e.g. Afghanistan], disaster response missions [e.g. Tsunami in Southeast Asia, earthquake in the Philippines], Navy deployments), of different lengths, with different roles and different risks (both at the mission and at the individual levels). Lack of statistical power has historically prevented DND from investigating deployment beyond a dichotomous yes/no exposure, which ignores the heterogeneity of exposures related to deployment (30). It is our hope that part of the broader work on CF CAMS II, the analyses will include evaluations on a number of different approaches to defining deployment that will take direction from the existing literature. Alternate approaches in the literature to defining deployment have included looking at the total number of deployments (38-40), length of first deployment (39), total duration of all deployments (39, 40), specific location(s) of deployment (39), isolation level of deployment(s) (38), occupation during deployment (38), and the categorization of combat exposure (38, 41).

9. Ethics and Privacy

This project was reviewed and was approved by an external Institutional Review Board (QUORUM Review IRB, Reference#QR#31460CDN/1). A consent waiver was requested and approved by the external IRB, allowing the study to be conducted without the individual consent of all participants. This is in accordance with Article 3.7 of the second edition of the Canadian Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (42).

Furthermore, all record linkages conducted by STC must be vetted through a formal application process with the Chief Statistical Officer of STC, in accordance with the rules and regulations of the Canadian Federal Statistics Act. Access to the linked data required further application to and approval by the Federal Research Data Centre (FRDC) at STC.

The risk to participating subjects will be low to negligible. They will be exposed to no physical risk. The record linkage requires some degree of loss of privacy that is considered minimal.

No microdata files will be externally released by STC. Analyses will take place with STC's FRDC. All data will be anonymized by STC prior to providing access to DND and VAC. STC will only release aggregated data, once vetted using strict rules specific to this study. It should also be noted that epidemiological staff from DND and VAC will complete statistical analyses as deemed employees of STC and under specifications outlined in the Statistics Act (1970).

In an effort to ensure that the needs and privacy of the subjects are maintained, all CF CAMS II consensus decisions are made by a Steering Committee whose membership includes the CF CAMS II researchers (DND and VAC), other DND researchers from other DND departments, STC representatives, and VAC and CAF Senior Leadership. In addition, representatives from both the VAC and the CAF Ombudsman Offices are also members of the Steering Committee, allowing them to advocate for still serving and released personnel, as well as allowing them to relay pertinent information back to their respective stakeholders.

DISCUSSION

As an employer of a high-risk workforce, the CAF has a requirement to monitor the health of its personnel to identify adverse occupational exposures so as to mitigate their effect on the long-term health of its personnel. The CAF has implemented a health surveillance strategy over the nearly last two decades that includes surveillance programs (e.g. injury, suicide, mortality, disease and injury on deployment), cross-sectional surveys (e.g. Health and Lifestyle Information Survey), cluster investigations, and record linkage studies (e.g. Gulf War Veterans Study and CF CAMS I). At VAC, CF CAMS contributes to the health surveillance of Veterans, after release from the military, in conjunction with the Life After Service Studies (LASS) (43, 44). This is of particular importance as it is estimated that VAC provides services to approximately 8% of the total living Veteran population.

However, each of these approaches has a number of methodological or statistical limitations/problems that have limited the applicability of these findings to evidence-based care and policy.

It is our strong belief that the proposed methodology in this protocol mitigates the following eight challenges that we have historically identified in our efforts to conduct occupational cancer and mortality surveillance:

1. Defining the cohort
2. Creating a cohort file
3. Data completeness in linked file
4. Establishing denominators and comparator populations
5. Achieving the complete ascertainment of mortality and/or cancer morbidity
6. Changes in ICD coding over time
7. Selecting key explanatory variables
8. Adhering to ethics and privacy responsibilities

By opting to use pay data, the CF CAMS II has identified the most complete and accurate data source for developing a full cohort over an appreciable amount of time. The confidence in the remuneration data has implications not only on the quality and completeness of the linked data, but also in the generation of denominators and of any comparisons made with reference populations. The remuneration data also provides broad access to possible occupationally-related independent variables (e.g. rank, deployments, service start and stop dates) that allow for advanced statistical analyses

(e.g. Cox regression models) to determine whether any, and to what extent, these independent variables are linked to adverse health outcomes. High quality data also means that we will be able to more optimally derive important explanatory variables, including deployment, and estimate their importance in affecting adverse health outcomes.

Strengths and Limitations of the Proposed Study

The main strengths of this proposed study are the length of the follow-up period (nearly 40 years) and the completeness of the cohort for this time period.

Because there are direct remuneration implications in having incorrect information captured in the CCPS, we are confident that the data that were used to build the CF CAMS II cohort are both more complete and accurate. Furthermore, the inclusion of Reserve Class C service within the CF CAMS II cohort file will allow for a better accounting of time served which will, in turn, allow for more accurate ascertainment of any causal links between military service and adverse health outcomes.

The additional strength of this study is that it is a living study; as the CAF military cohort continues to grow over time, the addition of new records for new enrolees as well as updates to existing study participants to the existing cohort file within Statistics Canada is planned. In turn, these new and updated records can be relinked to the CVSD and CRDB to continue to monitor the health of the CAF over time. This long-term study will be particularly interesting from a lag-time perspective; the longer the follow-up period, the broader the age representation within the cohort. In turn, the longer the follow-up period, the longer the person-time contributing the study, and the more outcome data are expected. It will also allow to monitor more recent military events (e.g. Afghanistan) for which there is currently insufficient lag-time and/or events to identify possible premature mortality or cancer morbidity.

To the best of our knowledge, there are no other studies that have documented mortality and cancer morbidity outcomes in a full military cohort, nor over such a long period of time. The nearly 40 years of follow-up time may also provide sufficient lead up time for conditions with delayed expression following exposure (e.g. certain cancers). The large sample size (approximately 230,000 individuals contributing a total of more than 5 million person-years) may also provide sufficient statistical power to investigate less common outcomes.

The main limitation of this approach is the exclusion of any Reservist A or B time served, thereby underestimating total time served as well as the possible relationship between time served and excess mortality or cancer morbidity. However, given that the military “exposure” time for A and B Reservists is fairly small, and proportionally much smaller than the time spent in a civilian capacity, it is expected that this underestimate will not appreciably change the results reported as part of this study. This assumption is supported by research into Canadian Reservists (43).

It is expected that additional limitations may be identified upon using the linked data; these will be disclosed and attempts to mitigate them will be described.

How Will the Findings Emanating from this Proposed Study Be Used?

Results from the CF CAMS II will enhance the understanding of risk factors for mortality and cancer in still serving and released military populations. Specifically, this study has the potential to answer many questions, such as:

1. What are the causes of death among persons who served in the military?
2. What are the leading preventable causes of death?
3. Are there any causes of death or occurrences of cancer that are elevated in CAF personnel during or after employment?
4. Do certain military occupational groups have a higher risk of death or cancer compared to other personnel?
5. Are there certain causes of death that are higher than expected compared to the Canadian population of the same age and sex?
6. Are there certain types of cancer that are higher than expected compared to the Canadian population of the same age and sex?

This study has the potential to provide novel and sound evidence on the risks and protective factors of military life to an extent not yet seen in the literature. The length of the follow-up period/cohort (nearly 40 years), the complete population coverage, and the availability of sound occupational risk factor data make this record-linkage study a potentially groundbreaking study on the relationship 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 sound evidence that may also benefit our allied militaries.

It is expected that the results from this study will be disseminated both within the involved organisations (DND/CAF, VAC, STC) as well as to the general public, through their publication and dissemination in peer-reviewed journals.

ABBREVIATIONS

ACM	All-cause mortality
AJCC	American Joint Committee on Cancer
CAF	Canadian Armed Forces
CCR	Canadian Cancer Registry
CCPS	Central Computerized Pay System
CF CAMS	Canadian Forces Cancer and Mortality Study
CGP	Canadian General Population
CIHI	Canadian Institutes for Health Information
COD	Cause of death
CRA	Canada Revenue Agency
CRO	Cancer registration organisation
CS	Collaborative stage
CVSD	Canadian Vital Statistics Database
DCSM	Directorate Casualty Support Management
DFHP	Directorate of Force Health Protection
DND	Department of National Defence
DRD	Derived Record Depository
Dx	Diagnosis
ESDC	Employment and Social Development Canada
FRDC	Federal Research Data Centre
HRMS	Human Resources Management System
ICD	International Classification of Disease
ICD-O	International Classification of Disease for oncology
IRB	Institutional Review Board
LASS	Life After Service Study
Mets	Metastases
MOC	Military Occupation Code
MOSID	Military Occupational System Identification
NYSIIS	New York State Identification and Intelligence System
SDLE	Social Data Linkage Environment
SIN	Social insurance number
SIR	Standardized Incidence Ratio
SMR	Standardized Mortality Ratio
STC	Statistics Canada
TNM	Tumor, node, metastatic classification of malignant tumours
VAC	Veterans Affairs Canada

REFERENCES

1. Li C, Sung F. A review of the healthy worker effect in occupational epidemiology. Occup Med. 1999;49(4):225-9.
2. Attfield MD, Kuempel ED. Mortality among U.S. underground coal miners: A 23-year follow-up. Am J Indust Med. 2008;51(4):231-45.
3. Miller BH, MacCalman L. Cause-specific mortality in British coal workers and exposure to respirable dust and quartz. Occup Environ Med. 2010;67:270-6.
4. Hendryx M, O'Donnell K, Hom K. Lung cancer mortality is elevated in coal-mining areas of Appalachia. Lung Cancer. 2008;62(1):1-7.
5. Band PR, Le ND, Fang R, Astrakianakis G, Bert J, Keefe A. Cohort cancer incidence among pulp and paper mill workers in British Columbia. Scand J Work Environ Health. 2001;27(2):113-9.
6. Mills PK, Kwong S. Cancer incidence in the United Farmworkers of America (UFW), 1987-1997. Am J Ind Med. 2001;40(5):596-603.
7. Mills PK, Yang R. Prostate cancer risk in California farm workers. J Occup Environ Med. 2003;45(3):249-58.
8. Ma F, Fleming LE, Lee DJ, Trapido E, Gerace TA, Lai H, et al. Mortality in Florida professional firefighters, 1972-1999. Am J Indust Med. 2005;47(6):509-17.
9. Kahn SA, Woods JFF, Rae L. Line of Duty Firefighter Fatalities: An Evolving Trend Over Time. J Burn Care Res. 2015;36(1):218-24.
10. Sung TI, Chen PC, Lee KJH, Lin YP, Hsieh GY, Wang JD. Increased standardized incidence ratio of breast cancer in female electronics workers. BMC Public Health. 2006;7:102.
11. Yassi A, Tate RB, Routledge M. Cancer incidence and mortality in workers employed at a transformer manufacturing plant: update to a cohort study. Am J Ind Med. 2003;44(1):58-62.
12. Akhtar FZ, Garabrant DH, Ketchum NS, Michalek JE. Cancer in US Air Force Veterans of the Vietnam War. J Occup Environ Med. 2004;46(2):123-36.
13. Yi S-W. Cancer Incidence in Korean Vietnam Veterans During 1992-2003: The Korean Veterans Health Study. J Prev Med Public Health. 2013.
14. Yi S-W, Ohrr H. Agent Orange Exposure and Cancer Incidence in Korean Vietnam Veterans: A Prospective Cohort Study. Cancer. 2014;120:3699-706.
15. Yi S-W, Ryu S-Y, Ohrr H, Hong J-S. Agent Orange exposure and risk of death in Korean Vietnam veterans: Korean Veterans Health Study. Int J Epidemiol. 2014:1825-34.
16. National Academis of Sciences. Veterans and Agent Orange: Update 2014. Washington, DC: National Academies Press; 2016.
17. Storm HH, Jorgensen HO, Kejs AMT, Engholm G. Depleted uranium and cancer in Danish Balkan veterans deployed 1992-2001. Eur J Cancer. 2006;42:2355-8.
18. Veterans Affairs Canada. Depleted Uranium and Canadian Veterans: A Review of Potential Exposures and Health Effects. Charlottetown, PEI: VAC; 2013.
19. Bullman TA, Mahan CM, Kang HK, Page WF. Mortality in the US Army Gulf War Veterans Exposed to 1991 Khamisiyah Chemical Munitions Destruction. Am J Public Health. 2005;95(8):1382-8.

20. Kang HK, Bullman TA. Mortality Among U.S. Veterans of the Persian Gulf War. New England Journal of Medicine. 1996;335:1498-504.
21. Kang HK, Cypel Y, Kilbourne AM, Magruder KM, Serpi T, Collins JF, et al. HealthViEWS: Mortality Study of Female US Vietnam Era Veterans, 1965-2010. Am J Public Health. 2014.
22. Young HA, Maillard JD, Levine PH, Simmens SJ, Mahan CM, Kang HK. Investigating the Risk of Cancer in 1990-1991 US Gulf War Veterans With the Use of State Cancer Registry Data. Ann Epidemiol. 2010;20:265-72.
23. Knoke JD, Gray GC, Garland FC. Testicular Cancer and Persian Gulf War Service. Epidemiol. 1998;9(6):648-83.
24. McBride D, Cox B, Broughton J, Tong D. The mortality and cancer experience of New Zealand Vietnam war veterans: a cohort study. BMJ Open. 2013;3:e003379.
25. Strand LA, Martinsen JI, Borud EK. Cancer risk and all-cause mortality among Norwegian military United Nations peacekeepers deployed to Kosovo between 1999 and 2011. Cancer Epidemiol. 2014.
26. Rolland-Harris E, Maher M. Surveillance of Mortality and Cancer Morbidity in the Canadian Armed Forces: A Life-Course Approach. STO-HFM-239 Military Health Surveillance; 2015; Paris, France: NATO STO.
27. Statistics Canada. The Canadian Persian Gulf cohort study: summary report. Ottawa: Statistics Canada; 2005.
28. Birkett N, Brodsky L. Gulf War Veterans Study Protocol. 2001.
29. Park J. A profile of the Canadian Forces. Perspectives on Labour and Income. 2008;9(7).
30. Rolland-Harris E, Whitehead J, Matheson H, Zamorski M. 2015 Report on Suicide Mortality in the Canadian Armed Forces (1995 to 2014). In: Program SGHR, editor. Ottawa: Department of National Defence; 2015.
31. 4th ed. Toronto: OUP; 2001. A Dictionary of Epidemiology.
32. Checkoway H, Pearce N, Kriebel D. Research methods in occupational epidemiology. Second ed. New York: Oxford University Press; 2004.
33. Statistics Canada. Vital Statistics - Death Database (CVSD) Ottawa: Minister of Industry; 2017 [Available from: <http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=347902>.
34. Statistics Canada. Canadian Cancer Registry (CCR) Ottawa: Minister of Industry; 2017 [Available from: <http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=329031>
35. Statistics Canada. Vital Statistics - Death Database (CVSD) Ottawa: Minister of Industry; 2017 [Available from: http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3233.
36. Wang F-L, Morrison K. ICD-9 to ICD-10 Coding with Reference to Causes of Death Grouping in Alberta. In: Wellness AHa, editor. Edmonton, AB: Alberta Health and Wellness; 2006.
37. National Cancer Institute (USA). ICD-9-CM to ICD-10-CM Conversion Program Washington, DC2016 [Available from: https://seer.cancer.gov/tools/conversion/.
38. Garber B, Zamorski M, Jetly R. Mental Health of Canadian Forces Members While on Deployment to Afghanistan. Can J Psych. 2012;57(12):736-44.

39. Boulos D, Zamorski M. Deployment-related mental disorders among Canadian Forces personnel deployed in support of the mission in Afghanistan, 2001-2008. Can Med Assoc J. 2013;185(11):E545-52.
40. LeardMann C, Powell T, Smith T, Bell M, Smith B, Boyko E, et al. Risk Factors Associated With Suicide in Current and Former US Military Personnel. JAMA. 2013;310(5):496-506.
41. Wells T, Miller S, Adler A, Engel C, Smith T, Fairbank J. Mental health impact of the Iraq and Adghanistan conflicts: a review of US research, service provision, and programmatic responses. Int Rev Psychiatry. 2011;23:144-52.
42. Canadian Institutes for Health Research, Natural Sciences and Engineering Research Council of Canada, Social Sciences and Humanities Research Council of Canada. Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans. In: Interagency Secretariat on Research Ethics, editor. Ottawa: Interagency Secretariat on Research Ethics; 2010.
43. VanTil L, MacLean MB, Poirier A, McKinnon K, Keough J, Pedlar D, et al. Veterans of the Reserve Force: Life After Service Study 2013. In: Veterans Affairs Canada, editor. Charlottetown: Veterans Affairs Canada; 2016.
44. MacLean MB, Van Til L, Thompson JM, Pedlar D, Poirier A, Adams J, et al. Life After Service Study: Data Collection Methodology for Income Linkage and Transition to Civilian Life Survey. In: Research Directorate, editor. Charlottetown, PEI: Veterans Affairs Canada; 2010.

Figure 1. Flow Chart Describing the CF CAMS II Cohort Building Process

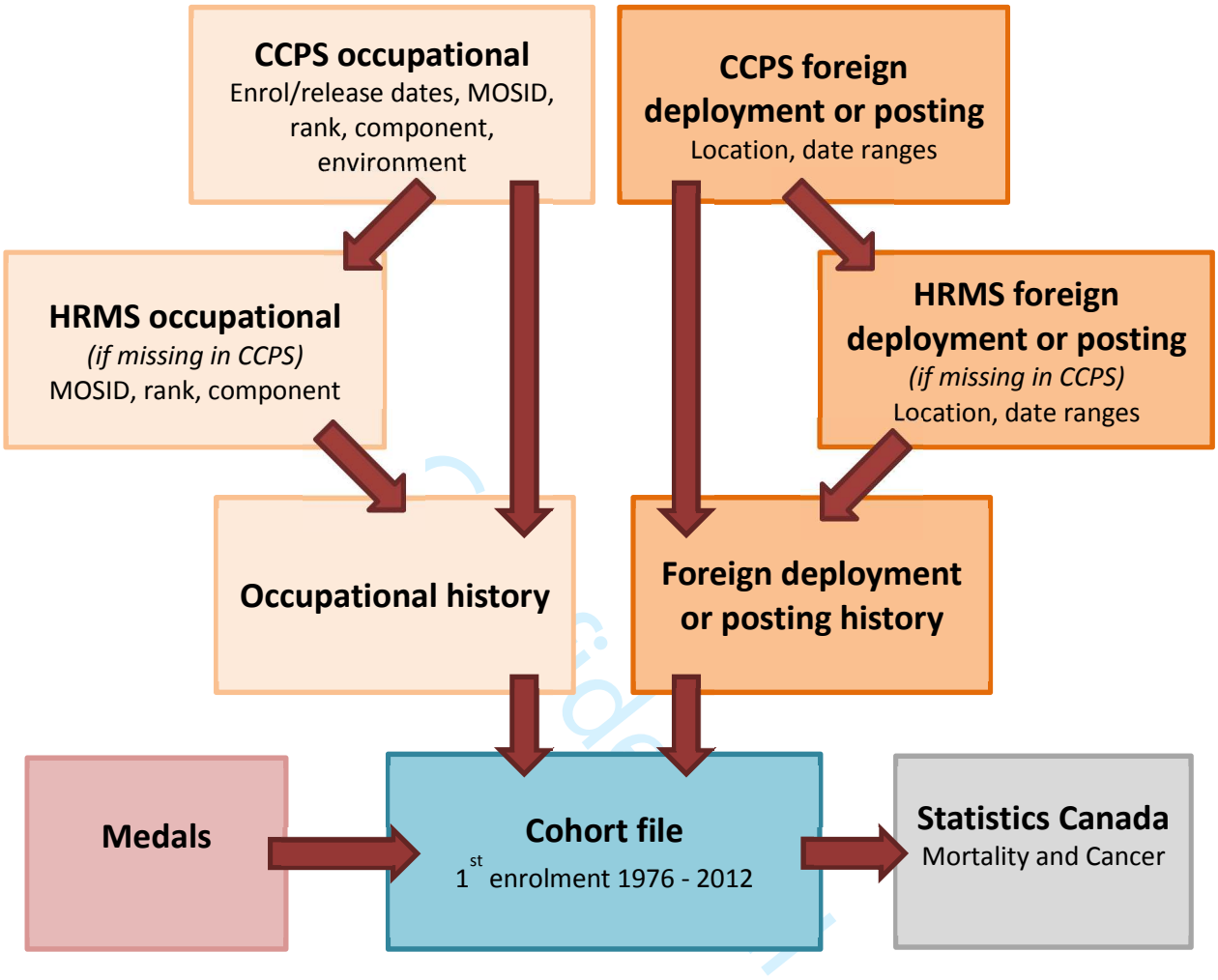

Table 1. Requested Canadian Cancer Registry Data Elements

Variable	Description
Patient variables	
Patient record type	Type of record (new, update, delete)
Type of current surname*	Code describing type of surname currently used by patient in "Current surname" variable
Current surname*	
First given name*	
Second given name*	
Sex	
Date of birth	
Birth surname*	Legal surname at birth
Date of death	
Underlying cause of death	
Autopsy confirming cause of death	Code indicating whether COD from official death certificate takes into account autopsy findings
Derived patient variables	
Vital status	
Number of tumours	Number of tumour records belonging to patient record
Tumour variables	
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 localisation 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/diametre of the primary tumour (mm)
CS extension	Primary tumour growth within the organ of origin or its direct extension into neighbouring organs
CS tumour size/ext eval	Code indicating how the “CS tumour size” and “CS extension” were determined (based on diagnostic method employed)
CD 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 (based on diagnostic methods employed)
Regional nodes examined	Total number of regional lymph nodes that were removed/examined by pathologist
Regional nodes positive	Exact number 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 (based on diagnostic methods employed)
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

* Fields to be used for linkage only – not available for analysis