Article details: 2015-0068	
Title	A methodologic framework to evaluate the number of cancers attributable to lifestyle and environment in Alberta, Canada
Authors	Anne Grundy PhD, Christine M. Friedenreich PhD, Abbey E. Poirier MSc, Farah Khandwala MSc, Darren R. Brenner PhD
Reviewer 1	Dr. Dena Schanzer
Institution	Public Health Agency of Canada, Ottawa, Ont.
General comments (author response in bold)	*This manuscript describes the protocol for a series of articles on the number of cancers attributable to various exposures. The objective of the manuscript is not clear. It is only in the conclusion section that the objective of providing an introduction to a series of articles is stated. The manuscript reads as a mix of content taken from the original project proposal and the methods used in completed studies. The discussion section discusses results of some of their studies, rather than the protocol itself. I have already reviewed CMAJOpen-2015-0070 on alcohol consumption, and this manuscript does not appear to add enough for a separate article. I did not see the supplementary tables, so I am not aware of the list of 24 proposed studies. As a manuscript with a focus on the protocol, I would expect to see much more effort put into validating the methods and identifying best practices. As a guidance document, I have a number of concerns.* Response: **The objective of the methods manuscript was to provide an overview of the overall principles used in our series of manuscripts concerning a project to estimate population attributable risks for modifiable lifestyle and environmental risk factors for cancer in Alberta. While we understand that some of this manuscript is similar to the methods described in our exposure-specific manuscripts, our intent with this manuscript was to outline the general principles used in our analysis across exposures, where exposure specific aspects of the methods will be described in more detail in exposure-specific manuscripts. Furthermore, the scope of this project did not include the development of new methods for estimation of population attributable risks, but rather for each exposure group to identify the most applicable existing method that could be used to estimate population attributable risks, given the data to which we had access. We have elaborated on the overall purpose of our population attributable risk project and the goals of this specific manuscript in at the end of the Background section on p. 3.** **We respectfully disagree with the reviewer that our discussion section discusses results of our studies rather than the methods themselves. The second paragraph of the discussion (Interpretation section) on p. 8 – 9, describes previous efforts to estimate population attributable risks both in Canada and in other countries and the extent to which we expect the results we will present in our series of manuscripts to be comparable to these previous efforts. In many cases, previous work on population attributable risk, particularly in Canada, has focused on individual exposures and as such will only be comparable to the corresponding exposure-specific manuscript in our series. We believe that the incorporation of methods from previous population attributable risk efforts into our project will facilitate interpretation and comparison of our results in the context of previous work. We have added a sentence to the discussion on p.10 to clarify this idea. Further, we believe the limitations discussed on p. 10 are generally applicable to the entire project, while more exposure-specific limitations will be discussed in exposure-specific manuscripts. We have added Supplementary Table 1 to the main body of this manuscript as Table 2 so that the contents are more readily accessible to readers and to clarify any potential confusion concerning the specific exposures that will be addressed in this series of manuscripts. The new Table 2 (formerly Supplementary Table 1) is included below for the reviewer's reference.**

Table 2: Exposure and Cancer Site Associations of Interest to be Included in this Project

Manuscript	Exposure	Cancer types consistently associated with exposure
1	Active Tobacco Exposure	Lung
		Oral cavity and pharynx
		Oesophagus
		Stomach
		Liver
		Pancreas
		Colorectum
		Larynx
		Cervix
		Ovarian (mucinous)

		Urinary bladder
		Kidney
		Acute myeloid leukemia
	Passive Tobacco Exposure	Lung
		Oral cavity and pharynx
		Oesophagus
		Larynx
2	High Alcohol Intake	Mouth
		Pharynx
		Larynx
		Liver
		Colorectum
		Breast (pre & post-menopause)
3	Overweight/Obesity	Breast (post-menopausal)
	(>25 kg/m^2)	Colorectum
		Oesophagus (adenocarcinoma)
		Kidney
		Endometrium
		Gall bladder
		Pancreas
4	Physical inactivity	Breast (post-menopausal)
		Colorectum
		Endometrium
		Lung
		Ovary
		Prostate
5	Low vegetable intake	Oral cavity and pharynx
	(non-starchy)	Oesophagus
		Stomach
		Larynx
	Low fruit intake	Oral cavity and pharynx
		Oesophagus
		Stomach
		Larynx
		Lung
6	High red meat intake	Colorectum
	High process meat intake	Colorectum
7	Low fibre intake	Colorectum
8	Low vitamin D High salt intake Low dietary calcium intake	Colorectum Breast Stomach Colorectum
9	Hormone therapies	
	Oral contraceptive use	Breast
		Endometrium
		Ovary
	Hormone Replacement Therapy	Breast
		Endometrium
		Ovary
10	Infectious agents	
	Human papillomavirus	Cervix
		Vagina
		Penis
		Anus
		Vulva
		Oropharynx
	Helicobacter Pylori	Stomach
	Epstein Barr Virus	Non-Hodgkin lymphoma
		Hodgkin lymphoma
		Burkitt's lymphoma
		Nasopharyngeal carcinoma
	Hepatitis B Virus	Liver
	Hepatitis C Virus	Liver
11	UV Exposure	Melanoma
12	Radon	Lung
13	Air pollution	
	PM 2.5	Lung

	1) Latency: The authors seem to be recommending a point estimate for latency. This may be acceptable in a steady state situation where levels of exposure are similar for all birth cohorts. When exposures differ significantly by birth cohort (perhaps due to public health interventions) the latency period of new cancers may be quite different from the average latency period. This is particularly true for hepatitis C and HCC, as persons born in the 1950s have the highest prevalence of chronic HCV. I would expect to see recommendations that a point estimate for latency is only appropriate when steady state exposure can be confirmed. This is an issue that should be addressed with each of the 24 planned manuscripts. This issue is mentioned as a limitation, however, as a guidance document, this exercise should be completed. **Response: The reviewer is correct that we are recommending a point estimate for estimates of latency. The rationale for this choice is that all estimates of population attributable risk presented in this series of manuscripts were conducted with the intention of applying the resulting population attributable risk values to cancers diagnosed in Alberta in 2012. Therefore, the intention of the latency aspect of our calculations was to address the specific issue that cancers diagnosed in 2012 would be attributable to past exposures and the selected latency period was intended to assist us in determining when the most appropriate past exposure time-window would be for estimates of exposure prevalence that would be associated with cancers diagnosed in 2012.** **We consider the latency period to be a biological phenomenon in the sense that the ability of the exposure to lead to cancer is determined by exposure to a given level of an individual exposure rather than by the population prevalence as the reviewer seems to be suggesting through the discussion concerning differing exposure prevalence values in different birth cohorts. As described in our response to the statistical reviewer, where available we utilized age- and sex-specific measures of exposure prevalence so if exposure prevalence is different in different age groups these differences would be captured. While this distinction is apparent in our exposure-specific manuscripts, we have added a sentence to the methods section on p.5 emphasizing that exposure prevalence measures were age and sex-specific where possible to improve the clarity of this aspect of our methods in this manuscript. We have also added a further discussion of the concept we refer to as latency to the methods section on p. 6 including the addition of a new figure (Figure 3) to clarify our terminology with regards to temporal issues of exposure, measurement and disease.** **Our decision to use the average latency period suggested by cohort studies (as described in the methods on p. 5 – 6) was influenced by the fact that previous work that has attempted to quantify appropriate latency periods had less formalized methods for determining the appropriate latency period (Parkin, 2011). We agree that, as described in the discussion on p. 10, our methods to assess latency remain somewhat imprecise but further sensitivity analyses examining the choice of latency period for individual exposures were often difficult to conduct due to limited availability of exposure prevalence data. For example, much of our prevalence data comes from the Canadian Community Health Survey which only includes data from the year 2000 onwards. As such, while we acknowledge that the lack of precision in our latency period estimates is a limitation of our analysis, our approach of establishing a standardized method through which latency periods for all exposure-specific analyses were assessed is an improvement over previous work, which has not included a standardized method. The authors are currently working on a larger population attributable risk project for all of Canada and as part of this work are will develop new methods for the assessment of concepts like latency periods. However, for the project described in this manuscript, such methods development was beyond the scope of our work.** *2) Details on how to estimate the 95% CI should be provided. This is a major challenge for all modelling studies, and one that should be built into the selection process of risk factors to study. If done properly, the authors may find that the precision of PAR estimates for many risk factors is poor, and may not warrant publication. To do a Monte Carlo simulation, the section on Methods/data sources should include methods to identify the precision of point estimates as well as the point estimate. As a rule of thumb, what level of precision is required to produce reasonable estimates of the PAR or number of cancers attributable to a modifiable risk factor?* **Response: Additional detail concerning the methods used to estimate our 95% confidence intervals, specifically related to the distributions used for these estimates, have been added to the methods on p. 8 – 9. The Monte Carlo simulations were chosen to estimate confidence intervals as these methods have been used in previous population attributable risk studies to estimate uncertainty associated with population attributable risk estimates (Renehan et al., 2010; Lee et al., 2012). Further as described in our response to the statistical reviewer, confidence intervals for impact estimates based on**

	computer intensive methods have been shown to outperform confidence intervals based on asymptotic estimates (Lehnert-Batar et al., 2006). As some previous population attributable risk work has not included confidence intervals at all (Parkin, 2011), our objective with this project was to identify the most appropriate existing methods (Renehan et al., 2010; Lee et al., 2012) with which to incorporate measures of uncertainty into our work. *3) Quality of estimates. The quality of a study is not sufficient to assess the suitability of the estimate for the particular use. For example, a cohort study may focus on a sub-population that is not generalizable to the whole population (for example, household surveys do not include homeless, and will under estimate HCV prevalence in the population). The GRADE system (The Grading of Recommendations, Assessment, Development and Evaluation) is commonly used in evaluating evidence for a public health decision. A series of articles on GRADE published in the Journal of Clinical Epidemiology (http://www.gradeworkinggroup.org/publications/JCE_series.htm) provides a comprehensive introduction. It is also an example of an introductory article for a series of articles on a similar topic.* **Response: We consider the risk estimate quantifying the association between a given exposure and cancer site to be a biological phenomenon and attempted to obtain the risk estimates that would be the best representation of this relationship in the general Alberta population. As shown in Figure 1, we ranked potential risk estimate sources in order: 1) from international collaborative panels; 2) from recent meta-analyses; 3) from recent pooled analyses; and, 4) conducting our own meta-analysis if one was not already available. This hierarchy allowed us to use risk estimates that combine the results from multiple case-control and cohort studies from around the world to obtain the most representative risk estimate wherever possible. While we agree that these risk estimates may not be appropriate when considering specific population subgroups, the objective of our project was to provide estimates of population attributable risk for the general Alberta population and as such, we feel our chosen method of selecting risk estimates is appropriate. Additional text clarifying the target population for this study and its influence on the selection of risk estimates has been added to the Methods on p. 4. For most risk estimates we were also able to obtain measures of uncertainty (95% confidence intervals) that were used in our Monte Carlo simulations to estimate the uncertainty around our estimated population attributable risks.** **Further, while we thank the reviewer for the recommendation of the GRADE system, we have reviewed the GRADE material and following this review, are unsure whether this system is appropriate for our project. Specifically, in the introductory article to GRADE (Guyatt *et al.*, 2011), the authors note that the GRADE system is designed to assess the quality of evidence related to alternative management strategies, interventions or policies. Specifically, the proposed evaluation system automatically rates observational studies as lower quality evidence. While we understand that this approach may be appropriate for the evaluation of interventions, in the epidemiological context utilized in our population attributable risk project, many of the associations for which we were attempting to obtain risk estimates were only evaluated in observational studies. Furthermore, as alternative study designs such as randomized clinical trials may be infeasible and/or unethical.**
Reviewer 2	Lawrence Paszat
Institution	Institute for Clinical Evaluative Sciences
General comments (author response in bold)	*The protocol manuscript is thoughtful and considers peer reviewed scientific literature pertinent to the methods along with work done by Cancer Care Ontario and other organizations not published in peer review publications. It is written as a protocol rather than a systematic review. It may have some utility as a working document for a public health department or department of health. However, it does not evaluate the protocol against any test data or current data.* **Response: We thank the reviewer for the constructive feedback concerning our manuscript. The purpose of this manuscript was to serve as an introduction to a series of exposure-specific manuscripts that would include (but not be limited to) the tobacco and alcohol manuscripts submitted in conjunction with this methods manuscript and thus was intended to read as a methodological outline rather than a systematic review. The objective of the methods manuscript is to provide a general overview of the methodological principles used in the exposure-specific manuscripts, similar to a comparable manuscript published by Parkin (2011) as an introduction to a similar project conducted in the United Kingdom in 2010. As shown in the alcohol and tobacco manuscripts the exposure-specific methods are described in more detail and current data are used to estimate population attributable risks associated with individual exposures.**
Reviewer 3	Dr. Agnihotram Ramana-Kumar
Institution	McGill University, Epidemiology and Biostatistics. Montréal, Que.

General comments (author response in bold)	This paper is a very informative and useful for cancer research community. *1. When you say that the risk association measures were selected according to a hierarchy, does this imply a weighting scheme in the estimation?* **Response: No. The hierarchy was used to select risk estimates in the sense that risk estimates were chosen from highest available rank on the hierarchy for use in our analysis. Once a risk estimate was selected no weighting system was applied. A sentence clarifying this issue has been added to the methods on p. 4.** *2. My understanding is that a single measure was derived from a combination of RR, OR, HR, and IRR. How was this done?* **Response: As described in the response to the previous comment directly above, the single risk estimate (i.e. RR, OR, HR or IRR) from the highest available rank on our hierarchy of sources was chosen. For example, if risk estimates were available from both international collaborative panels and recent meta-analyses, the estimate from an international collaborative panel was used as it corresponded to a higher rank on the hierarchy presented in Figure 1. This clarification has been added to the methods section on p. 4.** *3. Similarly, were prevalence estimates weighted based on the hierarchy as well?* **Response: No. As described for the risk estimates, the prevalence estimate from the source with the highest rank on our hierarchy shown in Figure 2 was utilized. This information has been added to the methods section on p. 5.** *4. Supplementary Tables 1 and 2 were not included in the manuscript.* **Response: Supplementary Table 1 has been added as a main table in the manuscript and is now Table 2. Supplementary Table 2 is now referred to as Supplementary Table 1. We will ask the journal to include this in future revision cycles.** I also have minor comments to improve the text: *1. Page 4 of 17, line 36: "risk factors" appears twice consecutively. This is most likely a typo. Please correct.* **Response: This item has been corrected as suggested.** *2. Page 6 of 17, lines 32-36: For BMI, it is not clear how we determined the risk, were only higher levels of exposure considered to lead to higher risk, or were deviations from the level associated with the lowest risk considered to lead to higher risk?* **Response: Only levels of exposure that were considered deviations from 'normal' were used to evaluate population attributable risk. For example, with BMI individuals in the 'normal' range (i.e. BMI < 25) were considered unexposed, while those in the overweight (BMI = 25 – 29) or obese (BMI ≥ 30) range were considered exposed. We have attempted to clarify the text on p. 5 of the methods section to clarify this issue.** 3. *Page 6 of 17, lines 38-42: The latent period is defined as the time from disease initiation and detection though. So, the period being measured is the sum of the induction period (relative to the exposure of interest) and the latent period.* *A illustrative example is needed.* **Response: We thank the reviewer for raising this issue and have attempted to clarify our justification of the approach. We have revised the text of the manuscript to include a figure (Figure 3) illustrating what we refer to as the latent period. Specifically, we consider there to be a theoretical latent period which incorporates both the biologically relevant exposure period (the time between initiation of exposure and the initiation of carcinogenesis) and the induction period (the time between initiation of carcinogenesis and cancer diagnosis), while the measured latent period is the time between exposure measurement and cancer diagnosis. In our work we attempted to quantify the measured latent period, which has subsequently been referred to as simply the "latent period" for simplicity in the remainder of our manuscript and in subsequent manuscripts in this series. Additional language clarifying these considerations has been added to the methods section on p. 6.** *4. Page 8 of 17, equation 3: "Pc" or "pc"? (i.e., lower case or upper case?) Lower case was not defined. Use the uniform cases.* **Response: We have fixed the discrepancy here to use the lower case exclusively.** *5. Page 12 of 17: From Table 1, it is not clear how the exposures will be grouped into distinct manuscripts. Please consider clarifying.* **Response: We have added Supplementary Table 1 as a main table (Table 2) in the manuscript. We hope that this will clarify any confusion around the issue of**

	how exposures will be grouped. References: Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J et al. GRADE guidelines: 1. Introduction - GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011; 64(4):383-394. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. The Lancet 2012; 380(9838):219-299. Lehnert-Batar A, Pfahlberg A, Gefeller O. Comparison of confidence intervals for adjusted attributable risk estimates under multinomial sampling. Biom J 2006; 48(5):805-819. Parkin DM. I. The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010 – Introduction. Br J Cancer 2010; 105:S2-S5. Renehan AG, Soerjomataram I, Tyson M, Egger M, Zwahlen M, Coebergh JW et al. Incident cancer burden attributable to excess body mass index in 30 European countries. Int J Cancer 2010; 126(3):692-702.