PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Determining the long-term health burden and risk of sequelae for 14 foodborne infections, in British Columbia, Canada: protocol for a retrospective population-based cohort study
AUTHORS	Majowicz, Shannon E.; Panagiotoglou, Dimitra; Taylor, Marsha; Gohari, Mahmood; Kaplan, Gilaad; Chaurasia, Ashok; Leatherdale, Scott T.; Cook, Richard; Patrick, David; Ethelberg, Steen; Galanis, Eleni

VERSION 1 – REVIEW

REVIEWER	John Harris University of Liverpool, UK
REVIEW RETURNED	17-Feb-2020

GENERAL COMMENTS	Comprehensive study protocol with well-defined objectives. I have only a couple of minor points that might be considered by the authors for the study design and or analysis, or for clarity on the protocol. The use of databases for diagnosis of illness will give the number officially diagnosed but does underestimate the true burden of disease. This is accounted for in terms of patients who do not report symptoms and is a recognised limitation. However, there are two forms of under-ascertainment, firstly those who do not consult healthcare services as a result of their illness and secondly those who do but for whom no diagnosis is made (where laboratory tests do not find a positive result for any pathogen). The second study of Infectious Intestinal Disease in the UK found a considerable diagnostic gap; for 60% of patients tested no pathogen was identified. How will this be considered in their sensitivity analysis when dealing with misclassification? The authors state that the majority of the population are covered by the MSP system, but they have not given a proportion of the population covered. This could (albeit marginally) affect their estimates of effect of deprivation if, for example, those of a lower socio-economic status are not likely to be covered by the system.

REVIEWER	Brecht Devleesschauwer Sciensano, Belgium I have co-authored publications with the manuscript's first author.
REVIEW RETURNED	01-Mar-2020

GENERAL COMMENTS	The manuscript presents the protocol for a much needed

	assessment of the risk of developing sequelae following foodborne disease. The manuscript is well written and provides a high level of detail. 1. My main comment refers to the apparent lack of a control group in the main analysis. The authors do briefly mention this as a sensitivity analysis, so I would guess that there would be no practical or ethical issues restricting the authors from including a control group? 2. Linked to the above comment, the definition of the numerator and denominator under Objective 4 does not seem to match the definition of a population attributable fraction. The numerator should be the difference in the risk of disease (incidence) for the total population compared to the unexposed, while the denominator should be the risk in the total population. 3. What was the rationale for selecting the 14 included infections? Why for instance exclude Toxoplasma gondii, which is also associated with an important burden? 4. L184: I believe it could be interesting to estimate sequela risks for the "unknown" group, if the data would be at hand. Some studies make estimates for this unknown group, and then such estimates would come in handy. Some minor comments: 5. The authors state they include 14 infections, but list only 12 pathogens -- so I guess that Salmonella counts for three infections? Maybe this could be revised or at least made explicit? 6. For consistency: maybe use "Clostridium botulinum" instead of botulism? 7. Likewise, maybe use "hepatitis A virus" instead of "hepatitis A"? 8. The "spp." in "Salmonella spp." and "Campylobacter spp." should not be in italics. 9. L165: do all 8 databases have this PHN? 10. L277, L310: how will "related" health care use and prescription medication be defined? Do your databases on health care use and prescribed medications include the diagnoses linked to the use or prescriptions?

VERSION 1 – AUTHOR RESPONSE

VIEWER COMMENT	**AUTHOR RESPONSE**
Reviewer: 1, John Harris (University of Liverpool, UK)	
Comprehensive study protocol with well-defined objectives. I have only a couple of minor points that	Thank you for your review and assessment of our protocol. Thank you for your question about under-

might be considered by the authors for the study design and or analysis, or for clarity on the protocol. The use of databases for diagnosis of illness will give the number officially diagnosed but does underestimate the true burden of disease. This is accounted for in terms of patients who do not report symptoms and is a recognised limitation. However, there are two forms of under-ascertainment, firstly those who do not consult healthcare services as a result of their illness and secondly those who do but for whom no diagnosis is made (where laboratory tests do not find a positive result for any pathogen). The second study of Infectious Intestinal Disease in the UK found a considerable diagnostic gap; for 60% of patients tested no pathogen was identified. How will this be considered in their sensitivity analysis when dealing with misclassification?	ascertainment, particularly how we will handle the two types that you describe. In Canada, our under-ascertainment estimates have been calculated with a slightly different division: by (a) "under-diagnosis", calculated as the under-reporting that occurs between the total cases in the community, to those with positive laboratory tests, and (b) "under-reporting", calculated as the proportion of positive lab tests reported to notifiable disease systems. Please see Thomas et al [manuscript reference #2] (doi: 10.1089/fpd.2012.1389) for full details. In our study, we will use overall estimates that account for all under-ascertainment (i.e., from community cases to positive tests, and from positive tests to reports to the notifiable disease system). We apologize that our original text did not make this clear, and we have clarified this on lines 391–392.
The authors state that the majority of the population are covered by the MSP system, but they have not given a proportion of the population covered. This could (albeit marginally) affect their estimates of effect of deprivation if, for example, those of a lower socio-economic status are not likely to be covered by the system.	All residents of B.C. are required to register with the MSP by law (with the very few exceptions noted in the paper). Thus, while there are no official estimates of the proportion of the BC population covered, because the MSP is the sole, universal health care plan for the province, it is considered to cover all residents (except for those noted in the paper). There is no reason to expect this coverage would differ by socio-economic status, as everyone regardless of SES is covered. We have not made any revisions to the main text on this point, but have clarified this in the 3rd bullet of the strengths/limitations (lines 61–64).
Reviewer: 2, Brecht Devleesschauwer (Sciensano, Belgium)	
The manuscript presents the protocol for a much needed assessment of the risk of developing sequelae following foodborne disease. The manuscript is well written and provides a high level of detail.	Thank you for your review and assessment of our protocol.
1. My main comment refers to the apparent lack of a control group in the main analysis. The authors do briefly mention this as a sensitivity analysis, so I would guess that there would be no practical or ethical issues restricting the authors from including a control group?	In this research, we are studying a dynamic population in which the exposure status of individuals changes over time. This allows us to assess the risk and effect of exposure in terms of person-time, rather than using fixed (e.g., from baseline) cohorts of exposed versus unexposed individuals. In other words, instead of conceptualizing individuals in the population as belonging to one of two distinct "exposed" and "unexposed" groups, we instead track individuals over time and assign their person-time at risk to either an "unexposed person-time" period (e.g., prior to foodborne infection) or "exposed person-time" period (e.g., after foodborne infection). We have explained this on lines 139–144,

	and clarified lines 271 and 274. This is a conventional approach routinely used in life history analyses of disease processes studying the effects of time-varying covariates, and we hope you find this reasonable.
2. Linked to the above comment, the definition of the numerator and denominator under Objective 4 does not seem to match the definition of a population attributable fraction. The numerator should be the difference in the risk of disease (incidence) for the total population compared to the unexposed, while the denominator should be the risk in the total population.	Thank you for noting this. We have corrected this text to clearly differentiate between our PAF calculations (that will be done per standard formulae as you suggest), from our other calculation of the proportion of sequelae who have different reported foodborne infections, or who are possibly exposed (lines 348–358).
3. What was the rationale for selecting the 14 included infections? Why for instance exclude Toxoplasma gondii, which is also associated with an important burden?	These 14 infections were selected by our provincial public health partner (i.e., British Columbia Centre for Disease Control) as the most important foodborne infections in the province, in terms of health impact and prevention potential. Note we were limited to those foodborne infections that were reportable during the study period. There are two additional reportable infections that we could have included (brucellosis, paralytic shellfish poisoning). In B.C., brucellosis is very rare and nearly always travel-related, and paralytic shellfish poisoning (also rare) is syndromic and diagnosis is uncertain, so these two were not considered a priority for this study. Unfortunately, in B.C., only congenital toxoplasmosis is reportable, and thus we excluded it as well. We have explained this on lines 186–192.
4. L184: I believe it could be interesting to estimate sequela risks for the "unknown" group, if the data would be at hand. Some studies make estimates for this unknown group, and then such estimates would come in handy.	Thank you for this suggestion; we will make these estimates as suggested, and have noted this on lines 204–205.
5. The authors state they include 14 infections, but list only 12 pathogens -- so I guess that Salmonella counts for three infections? Maybe this could be revised or at least made explicit?	Clarified as suggested, lines 37, 99 and also in Table 1.
6. For consistency: maybe use "Clostridium botulinum" instead of botulism? 7. Likewise, maybe use "hepatitis A virus" instead of "hepatitis A"? 8. The "spp." in "Salmonella spp." and "Campylobacter spp." should not be in italics.	For all these points, we have corrected as suggested, throughout the paper.
9. L165: do all 8 databases have this PHN?	Thank you for this question. To be precise, seven of the eight databases (the ones with individual-level data) have PHN. The eighth database, the Statistics

	Canada Income Band data, is area-level information at the geographic level of the 6-digit postal code. For this database, Population Data BC links individuals (and their PHN) to their income band using their 6-digit postal code, resulting in individuals' PHNs associated to income bands. The PHNs are then used to link individuals across databases. We have clarified this in the text on lines 176–178, and in Table 1.
10. L277, L310: how will "related" health care use and prescription medication be defined? Do your databases on health care use and prescribed medications include the diagnoses linked to the use or prescriptions?	Thank you for this question. We will determine whether health care use is related to infection and sequelae using ICD diagnosis codes, and we will determine whether prescription medication use is related via medical expert consultation. We have noted this on lines 301–303.

VERSION 2 – REVIEW

REVIEWER	John Harris Public Health England, United Kingdom
REVIEW RETURNED	16-Jun-2020

GENERAL COMMENTS	I have no further comments to make.

REVIEWER	Brecht Devleesschauwer Sciensano, Belgium I have co-authored publications with some of the authors in the past 3 years.
REVIEW RETURNED	21-May-2020

GENERAL COMMENTS	Thank you for addressing my comments.