BMJ Open

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

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-036560
Article Type:	Protocol
Date Submitted by the Author:	19-Dec-2019
Complete List of Authors:	Majowicz, Shannon E. ; University of Waterloo, School of Public Health and Health Systems Panagiotoglou, Dimitra; McGill University, Department of Epidemiology, Biostatistics and Occupational Health Taylor, Marsha; BCCDC Gohari, Mahmood; University of Waterloo, School of Public Health and Health Systems Kaplan, Gilaad; University of Calgary, Departments of Medicine and Community Health Chaurasia, Ashok; University of Waterloo, School of Public Health and Health Systems Leatherdale, Scott T.; University of Waterloo, School of Public Health and Health Systems Cook, Richard; University of Waterloo, Department of Statistics and Actuarial Science Patrick, David; BCCDC; The University of British Columbia, School of Population and Public Health Ethelberg, Steen; Statens Serum Institut, Department of Infectious Disease Epidemiology and Prevention; University of Copenhagen Department of Public Health, Global Health Section Galanis, Eleni; BCCDC; The University of British Columbia, School of Population and Public Health
Keywords:	Epidemiology < INFECTIOUS DISEASES, Public health < INFECTIOUS DISEASES, Gastrointestinal infections < GASTROENTEROLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

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

Shannon E. Majowicz[1]*, Dimitra Panagiotoglou[2], Marsha Taylor[3], Mahmood R. Gohari[1], Gilaad G. Kaplan[4], Ashok Chaurasia[1], Scott T. Leatherdale[1], Richard J. Cook[5], David M. Patrick[3,6], Steen Ethelberg[7,8], Eleni Galanis[3,6]

[1] School of Public Health and Health Systems, University of Waterloo (Waterloo, Canada)

[2] Department of Epidemiology, Biostatistics and Occupational Health, McGill University (Montreal, Canada)

[3] British Columbia Centre for Disease Control (Vancouver, Canada)

[4] Departments of Medicine & Community Health Sciences, University of Calgary (Calgary, Canada)

[5] Department of Statistics and Actuarial Sciences, University of Waterloo (Waterloo, Canada)

[6] School of Population and Public Health, University of British Columbia (Vancouver, Canada)

[7] Department of Infectious Disease Epidemiology and Prevention, Statens Serum Institut (Copenhagen, Denmark)

[8] Global Health Section, Department of Public Health, University of Copenhagen (Copenhagen, Denmark)

*Corresponding Author:
200 University Avenue West
Waterloo, Ontario, Canada, N2L 3G1
Email: smajowicz@uwaterloo.ca
Telephone: +1 519 888 4567 extension 31790

KEY WORDS: foodborne infections, sequelae, cohort study, public health

ABSTRACT

Introduction: Over 1 in 8 Canadians is affected by a foodborne infection annually; however, the long-term consequences, including the risks and costs of sequelae, are unclear. We aim to estimate the health burden and direct costs of 14 infections commonly transmitted by food, considering the acute illness and subsequent sequelae and mortality, for the population of British Columbia (B.C.), Canada (~4.7 million).

Methods and analysis: We will conduct a population-based retrospective cohort study of the B.C. provincial population, over a 10-year study period (January 1, 2005 to December 31, 2014). Exposure is defined as a provincially-reported illness caused by: botulism, *Campylobacter*, *Cryptosporidium, Cyclospora, Giardia*, hepatitis A, *Listeria*, *Salmonella spp.* (non-typhoidal, Typhi, Paratyphi), Shiga toxin-producing *E. coli*, *Shigella*, *Vibrio parahaemolyticus*, or *Yersinia* (excluding *pestis*). We will link individual-level longitudinal data from eight province-wide administrative health and reportable disease databases that include physician visits, hospitalizations and day surgeries, deaths, stillbirths, prescription medications (except those to treat HIV), and reportable foodborne diseases. Using these linked databases we will investigate the likelihood of various sequelae and death. Hazard models will be used to estimate the risk of outcomes and their association with the type of foodborne infection. Epidemiologic analyses will be conducted to determine the progression of illness and the fraction of sequelae attributable to specific foodborne infections. Economic analyses will assess the consequent direct healthcare costs.

Ethics and dissemination: This study has been approved by a University of Waterloo Research Ethics Committee (#30645), the University of British Columbia Behavioral Research Ethics Board (#H16-00021), and McGill University's Institutional Review Board (#A03-M12-19A).

Results will be disseminated via presentations to academics, public health practitioners, and other knowledge users, and publication in peer-reviewed journals. Where such publications are not open access, manuscripts will also be available via the University of Waterloo's Institutional Repository (https://uwspace.uwaterloo.ca).

STRENGTHS AND LIMITATIONS

- This cohort is a near-complete set of individually-linked administrative health and reportable foodborne infection data, covering the ~4.7 million residents of British Columbia, Canada over 10 years (2005–2014).
- To the best of our knowledge, the study described in this protocol will be the most comprehensive assessment of the risk of sequelae following foodborne infections across multiple pathogens to-date.
- Because nearly all the British Columbia population is covered by a single provincial health insurance plan, movement of individuals within the province or between employers does not create loss to follow-up.
- Limitations include incomplete and lower quality (e.g., misclassification, use of non-specific codes) information associated with administrative health data, and under-ascertainment of foodborne infections typical to reportable disease data.

INTRODUCTION

Infections commonly transmitted via food, such as *Salmonella spp.* and Shiga toxin-producing *Escherichia coli* (STEC) are a global public health concern,[1] and in Canada they affect over 1 in 8 people annually.[2] Beyond the acute stage of illness (typically characterized by diarrhea and other gastrointestinal symptoms), these infections can cause severe and longer-term outcomes, such as spontaneous abortion, hemolytic uremic syndrome (HUS), Crohn's disease and ulcerative colitis (inflammatory bowel disease [IBD]), Guillain-Barré syndrome (GBS), and death.[3-11]

Estimates of the risk of sequelae following foodborne infection have come in part from prospective cohort studies conducted as follow-ups to outbreaks,[12-16] or to reports of sporadic cases (usually via laboratory testing and disease surveillance systems).[4, 17-23] While such prospective studies have the advantage of being able to tailor the data collection to address specific research questions, they have some important limitations. For example, outbreak follow-up studies are limited to specific strain(s) causing the outbreak and the specific population affected. Studies of sporadic cases often have short follow-up periods and rely on self-reported questionnaires to identify sequelae, both of which can lead to bias.

Retrospective, population-based cohort studies, in which administrative and registry data are analysed,[10-11, 24-32] offer several advantages that complement the prospective studies described above. They allow for a wider population to be covered, both sporadic and outbreak-associated infections (caused by the range of strains affecting the population) to be included, and the use of self-reports of event occurrence to be avoided. However, because they require population-wide, linked data on the exposures and outcomes of interest, they are less frequently conducted. To-date, such studies have not been conducted in Canada.

Although infections such as *Salmonella spp*. and STEC can be transmitted via several routes (e.g., person-to-person, water), their transmission via food and their presence throughout the food system (e.g., food animals as a reservoir for *Campylobacter spp.*,[33] food handlers shedding hepatitis A,[34] the ability of *Listeria monocytogenes* to persist in food production equipment [35]) mean that these infections are often termed "foodborne" although some fraction will not be transmitted via food directly. Here, we apply the term "foodborne infection" to 14 infections that can be transmitted via food (botulism, *Campylobacter*, *Cryptosporidium, Cyclospora, Giardia*, hepatitis A, *Listeria*, *Salmonella spp.* (non-typhoidal, Typhi, Paratyphi), STEC, *Shigella*, *Vibrio parahaemolyticus*, and *Yersinia* excluding *pestis*), recognizing that not all result from direct foodborne transmission.

The overall goal of this study is to estimate the health burden and costs of these 14 infections, considering the acute illness and subsequent sequelae and associated mortality, for the population of British Columbia (B.C.), Canada. Our specific objectives are to:

1. determine the risk of developing sequelae following infection;
2. describe the epidemiology and clinical progression across the range of outcomes, including acute illness, sequelae, and death;
3. quantify the direct healthcare costs due to these infections and their various outcomes; and
4. determine the risk of sequelae in the population attributable to these infections.

METHODS AND ANALYSIS

Study setting

B.C. is Canada's westernmost and third most populous province (~4.7 million circa 2014). The annual incidence of foodborne infections in B.C. is comparable to annual incidences in Canada, the United States, Australia, and Western Europe.[e.g., 2, 36-37]

All B.C. residents (defined as citizens or permanent residents of Canada who are physically present in B.C. for at least six months in a calendar year), their dependents, and certain other individuals (e.g., some holders of study or work permits) are covered by the province's health insurance program.[38] Enrolment is mandatory, and this program covers nearly all of the B.C. population (with the exception of members of the Canadian military, Royal Canadian Mounted Police, and some First Nations individuals covered by federal insurance programs). Physician visits, laboratory tests, hospitalizations, out-patient hospital services, and drugs for certain populations are among the publicly-funded benefits. The administrative datasets that contain these health care use data, along with vital statistics (e.g., births, deaths), demographic data, and other datasets, are held by the B.C. Ministry of Health and are accessible to researchers via Population Data B.C., a central repository and "multi-university, data and education resource" that "support[s] research access to individual-level, de-identified longitudinal data on British Columbia's 4.7 million residents".[39]

The B.C. Public Health Act mandates that reportable diseases, including several foodborne infections,[40] be reported by health professionals and laboratories to the local and provincial public health authorities, and these data are managed provincially by the B.C. Centre for Disease Control (BCCDC). These data are housed within Panorama, the provincial public health database of reportable diseases. In the Panorama database, as well as the administrative health and vital statistics databases, individuals are recorded by their unique Personal Health Number (PHN), allowing information from these data sources to be linked by individual.

Study design, population, and timeframe

This is a retrospective cohort study of the population of B.C., with additional descriptive, cost, and population attributable risk analyses. The study population is all individuals in B.C. registered with the provincial health insurance program at any point during the study period, i.e., all individuals with the following from 2005 to 2014 inclusive: one or more record in one or more of the Medical Services Plan (MSP), Discharge Abstracts Database (DAD), Vital Statistics Deaths, or PharmaNet; or record of coverage under the provincial insurance program within the Consolidation File database (see Table 1). The 10-year study period is January 1, 2005 to December 31, 2014, inclusive, with additional two-year wash-in (January 1, 2003 to December 31, 2004) and wash-out (January 1, 2015 to December 31, 2016) periods. During these periods we will identify occurrences of foodborne infections, sequelae, and death. The 10-year study period was selected to more than encompass timeframes for initial sequelae development and ensuing healthcare use currently reflected in the literature (i.e., days to years), although there is some evidence that sequelae can develop over longer timeframes (e.g., over decades).[41]

We assume that enrolment in the provincial health insurance program (i.e., entry into the study population) and reasons for exit from the cohort (e.g., moving away from B.C.) are not related to the exposures nor the outcomes of interest.

Data sources and linkage

The study will use individually-linked, longitudinal data from eight databases to investigate both acute and longer-term health outcomes following foodborne infection (Table 1). In totality, these data contain information on 14 reportable foodborne infections, physician and hospital visits, prescription medications, vital statistics, and various demographic descriptors, for

the B.C. population across the study period. All data will be stored and analyzed within Population Data B.C.'s virtual Secure Research Environment.

Table 1. Population-level administrative and reportable disease databases that will be used in this study (British Columbia [B.C.], Canada)

Database (Reference)	Database description and summary of variables included for this study	Date range
Health Care and Health Services Data		
Medical Services Plan Payment Information File [72]	Billing records for all medically necessary services provided by fee-for-service physicians. Includes service dates, up to five ICD*-9/ICD-10 diagnostic codes, MSP-specific fee-item codes, and physician specialties.	2003/01/01 to 2016/12/31
Discharge Abstracts Database (Hospital Separations) [73]	Data on discharges, transfers and deaths of in-patients and day surgery patients from acute care hospitals in B.C.; does not include emergency room visits. Includes admission and discharge dates, up to 25 ICD-10-CA diagnostic codes, service use and procedure codes, newborn and maternal data, discharge status, and province issuing health care number.	2003/01/01 to 2016/12/31
PharmaNet [74]	All prescriptions (for drugs and medical supplies) dispensed from community pharmacies, and from hospital outpatient pharmacies for patient use at home, in B.C. Includes date of dispensing, quantity, dose, costs/fees, the type of prescribing practitioner, and the Drug Information Number (or Product Information Number) assigned by Health Canada.	2003/01/01 to 2016/12/31
Population and Vital Statistics Data		
Vital Statistics Deaths [75]	All deaths registered in B.C. Includes time and place data, and ICD-10 codes for the nature and causes of death.	2005/01/01 to 2016/12/31
Vital Statistics Stillbirths [76]	All stillbirths registered in B.C. Includes the mother's unique study ID, time and place data, number of stillborn and live born children in the event, gestation period, and ICD-10 codes for the underlying cause of stillbirth.	2005/01/01 to 2016/12/31
Consolidation File [77]	Population Data B.C.'s central demographics file for research requests, containing all individuals who are eligible to receive services in B.C. Includes age, sex, neighbourhood income deciles, and local health authority area, as well as the date and number of days registered in the provincial health insurance program.	2003/01/01 to 2016/12/31

Statistics Canada Income Bands [78]	1000 income bands that contain information about the 6-digit postal code area in which the individual resides. Includes the average and median equivalised disposable income (derived from Statistics Canada tax-filer data, and available for the years 1992, 2002, and 2006), and the number of families, adults, and children in the area.	2002, 2006
Reportable Disease Data		
Panorama Public Health Information System	All cases of the following 14 reportable diseases reported in B.C.: botulism, *Campylobacter*, *Cryptosporidium, Cyclospora, Giardia*, hepatitis A, *Listeria*, *Salmonella spp.* (non-typhoidal, Typhi, Paratyphi), STEC, *Shigella*, *Vibrio parahaemolyticus*, and *Yersinia*. Includes onset date, reported date, health authority, and etiologic agent.	2005/01/01 to 2014/12/31

*ICD: International Classification of Diseases

Population Data B.C. will link the eight databases directly by individual using PHN, with additional identifiers (e.g., name, age, sex) used to validate linkages and link records probabilistically when PHNs are missing. Population Data B.C.'s detailed linkage process is fully described elsewhere.[42] Once the linkage is complete, each individual is then assigned a unique study identifier. All individually-linked, de-identified databases are provided by Population Data B.C. within their Secure Research Environment, a centralized online platform, accessible via virtual private network within Canada, for accessing and analyzing research data, with security standards that meet Data Steward requirements.

Measuring exposure and outcomes

The exposures of interest are infections with the 14 foodborne pathogens. Individuals will be considered exposed when and if they have a laboratory-confirmed and provincially-reported case of any of the following, recorded in the Panorama database during the study period: botulism, *Campylobacter*, *Cryptosporidium, Cyclospora, Giardia*, hepatitis A, *Listeria*, *Salmonella spp.* (non-typhoidal, Typhi, Paratyphi), STEC, *Shigella*, *Vibrio parahaemolyticus*,

and *Yersinia* (excluding *pestis*). Case definitions for each of these infections are specified by the BCCDC.[43]. Individuals without a reported foodborne infection, but who have International Classification of Disease (ICD) codes either for one of our infections (e.g., A02.0, '*Salmonella* enteritis') or for non-specific gastroenteritis (e.g., A08.4, 'viral intestinal infection, unspecified' [44]) within the MSP and DAD databases, will be considered potentially exposed. We will describe these individuals as a separate group in our descriptive, economic, and population attributable fraction analyses, but will remove them from analyses of sequelae risk.

It is possible for individuals to have more than one reported foodborne infection during the study period, either as a simultaneously occurring co-infection, or as two or more distinct events. For these individuals, we will treat this as a complex exposure problem; sequelae will be associated with the most plausible explanatory infection, considering biology and timing, and we will adjust for the presence of concurrent foodborne infections if applicable.[45-46]

Our primary outcomes of interest are those sequelae for which the link to a given foodborne infection is either established or is possible (Table 2). We selected sequelae (a) with evidence of an association with any of the 14 individual foodborne infections,[e.g., 47] and (b) that occur via direct effects of pathogens or their toxin, or via auto-immune or chronic inflammatory processes that can be triggered by the infection. We will classify individuals as having the sequelae via administrative case definitions, using International Classification of Disease (ICD) diagnostic codes within the MSP, DAD, and VS data, with the exception of stillbirths which will be determined using recorded events in the VS-Stillbirths database. The ICD codes in the MSP data are generally considered accurate to the third digit.[48] Although ICD codes in the VS and DAD data are ordered by most probable diagnosis, we will consider all codes, regardless of order.

Validation of the ICD codes is currently in progress, via a literature review to identify administrative case definitions that have been validated in the Canadian context, medical expert consultation, and, for those sequelae without a relevant validated definition, a targeted chart review in B.C. In addition to the ICD codes, we may also use PharmaNet data to improve sequelae classification if needed, by identifying pharmaceuticals given to patients (e.g., Tumor Necrosis Factor inhibitor use as an indication of reactive arthritis; intravenous immune globulin use for GBS).

Some of our sequelae of interest are lifelong (e.g., Graves' Disease), and some are transient in that complete recovery is possible (e.g., GBS, stillbirth). For lifelong sequelae, we will consider the individual as having the sequela on the earliest date they meet the administrative case definition for that sequela (with subsequent records considered as a continuation of the original event). For sequelae from which recovery and subsequent return to being at-risk is possible, we will consider the individual as first having the sequela on the earliest date they meet the administrative case definition for that sequela; we will then apply a post-sequela recovery time to determine the date on which the individual can be considered to be at-risk for a new, subsequent occurrence of that sequela.

Individuals may develop more than one sequela during the study period, either because they develop multiple different sequelae (e.g., HUS and GBS), or because they develop multiple occurrences of a single sequela from which complete recovery is possible (e.g., GBS, stillbirth). In all instances, the occurrence of multiple sequelae will be recorded and described. When individuals develop multiple different sequelae during the study period (e.g., HUS and GBS), we will treat these as distinct outcomes in our risk estimates. When individuals develop multiple

occurrences of the same sequela, we will treat these as distinct outcomes but account for recurrent events.[46]

Individuals with foodborne infections who develop a sequela listed in Table 2, but for which there is no current evidence of an established or possible link to the specific pathogen (e.g., *Campylobacter* and stillbirth), will be excluded from our estimates of sequelae risk (but included in sensitivity analyses).

For all 14 infections the secondary outcome of interest is death, which will be classified using recorded events in the VS-Mortality database. Finally, we are also interested in (a) the acute illnesses related to these infections (regardless of whether the individual develops sequelae or dies), and (b) additional outcomes following acute sequelae (e.g., end-stage kidney disease and kidney transplant following HUS); these will be include only in our descriptive and economic analyses.

Table 2. Established (E) and possible (P) sequelae of foodborne infections, that will be assessed in this study (British Columbia, Canada)

Foodborne Infection [1]	**Acute Kidney Injury**	**Celiac Disease**	**Erythema nodosum**	**Graves' Disease**	**Guillain-Barré syndrome** [2]	**Hemolytic Uremic Syndrome**	**Inflammatory Bowel Disease**	**Irritable Bowel Syndrome**	**Neonatal Listeriosis** [3]	**Stillbirth**	**Reactive Arthritis** [4]	**Thrombotic thrombocytopenic purpura** [5]
Campylobacter	E	P			E	P	P	P			E	
Cryptosporidium											P	
Cyclospora					P						P	
Giardia								P			P	
Hepatitis A	E				P							
Listeria monocytogenes									E [5]	E		
Salmonella (non-typhoidal)	E				P	P	P	P			E	
Salmonella Paratyphi	E				P	P	P	P			E	
Salmonella Typhi	E				P	P	P	P			E	
STEC	E					E					P	P
Shigella	E					E		P			E	P
Yersinia (excluding *pestis*)	E		E	P				P			E	

[1] botulism and *Vibrio parahemolyticus* do not have established or possible sequelae

[2] this includes GBS variants such as Miller Fischer syndrome; other neurological conditions such as chronic inflammatory demyelinating polyneuropathy will also be assessed

[3] considered here as a sequela of maternal *Listeria* infection

[4] this includes associated diagnoses such as anterior uveitis and ankylosing spondylitis

[5] shown not to be a sequela; retained to capture historical misdiagnosis of HUS

Measuring time at-risk

For all individuals, time-at-risk for sequelae (Figure 1) will be measured from the start of their entry into the study, which we define as the earliest registration date in the provincial health insurance program (recorded in the Consolidation File). Exposed individuals may contribute to both the exposed time-at-risk (during the post-infection 'at-risk' period, see below) and the unexposed time-at-risk (prior to, and after, the post-infection 'at-risk' period), while unexposed individuals will only contribute to the unexposed time-at-risk. Time-at-risk for a specific sequela will be measured in days, from the date of entry into the study, until: the development of that sequela, death, loss to follow-up, or the end of the study. We define loss to follow-up as the last date of coverage in the provincial health insurance plan, calculated using the start day registered in the most recent year plus the total days registered in that year.

During the unexposed time-at-risk, we will treat all individuals as having the potential to develop any of the sequelae (with the exception of neonatal listeriosis and stillbirth, for which only those who are pregnant are at risk). For those who develop a foodborne infection, unexposed time at-risk will end on the onset date of the infection. Infection onset date will be determined using the onset date reported in Panorama, and where this is missing, the date that the

infection was reported minus the number of days between onset to reporting (e.g., estimated using the Panorama data or from the literature).[49]

Exposed time at-risk will be measured starting from the infection onset date, plus any additional induction periods (specific to each sequela and currently being determined via literature review and medical expert consultation). The end of the exposed time at-risk period is currently being determined via literature review and medical expert consultation. During the exposed time-at-risk, individuals will be classified as having a sequela specific to their infection (Table 2) on the date within the 'at-risk' period on which they meet the administrative case definition for that sequela (e.g., the date of the physician visit or hospitalization). After the post-infection at-risk period ends, individuals will revert to contributing to the unexposed time-at-risk.

These data are subject to censoring and truncation. Individuals will be censored for the sequela in the event of: death, last date of coverage in the provincial health insurance plan (i.e., loss to follow up; calculated as above), or the end of the study period, whichever comes first.[50-51] In our descriptive and economic analyses, we will include all related health care use and prescription medication costs over the course of the infection and sequela(e), and in our estimates of mortality we will include any deaths recorded during the study period, following the sequela.

Analysis plan

Data will be analyzed and results reported following the STROBE and RECORD guidelines.[52-53]. Analyses will be performed using SAS 9.4 (SAS Institute Inc., Cary, NC, USA) and R.[54] The nature and extent of missing data will be described. If imputation is used to complete missing data, specific methods and assumptions will be reported. We will emphasize

estimation over tests of statistical significance by reporting relative measures of effect along with associated 95% confidence intervals.

The datasets in Table 1 contain the variable "sex" (identified via government records), that denotes whether individuals are 'male' or 'female', thereby capturing a composite of sex and gender. To reflect for potential sex- and gender-differences, we will report and interpret findings stratified by this variable, in addition to overall findings.

Objective 1: To determine the risk of developing sequelae following foodborne infection, we will estimate hazard ratios using Cox regression models,[55] that adjust for confounders and comorbidities (see below), along with the possible effect modifying role of age, sex, comorbidity, and medication use (e.g., antibiotics). We will include comorbidities as a composite score, using the revised Charlson comorbidity index and its associated coding algorithms.[56-59] Following foodborne infection, we will compare the cumulative risk of first diagnosis of each infection-specific sequela using life-table and Kaplan-Meier approaches.[55, 60] In the event an individual dies, we will use competing risk analysis.[61-62] For those who experience more than one foodborne infection across the study period, we will explore the impacts of having multiple infections on the risk of sequelae and mortality. We will also estimate the likelihood of dying following foodborne infection using the same methods described above.

Objective 2: To describe the epidemiology and clinical progression across the range of outcomes, including acute illness (typically diarrhea and other gastrointestinal symptoms), sequelae, and death, for each foodborne infection we will calculate incidence rates, demographic, geographic, and temporal distributions, timing and progression of outcomes, and case fatality rates, for both the acute stage, and sequelae associated to the foodborne infection.

Objective 3: To quantify the direct healthcare costs due to these infections and their various outcomes, we will determine health service use (i.e., patterns of use by type, frequency, timing of physician visits and hospitalizations), for both the acute foodborne infection and any sequelae. We will estimate direct healthcare costs of out-of-hospital physician visits, hospitalizations, and prescription medications. Out-of-hospital costs will be estimated using the MSP variables 'Fee Item' and 'Paid Service' and fee rates from the B.C. fee schedule.[63] Costs of in-patient and day-case hospitalizations will be calculated using established case-mix methodology (i.e., using the 'Resource Intensity Weight' of each hospitalization),[64] and the B.C. Ministry of Health unit costs for hospital stays.[65] Total prescription medication costs will be calculated using the drug cost claimed by the pharmacist which includes the ingredient cost, professional dispensing fee and other special service fees (if applicable). Because these costs are captured directly in the PharmaNet data, they will be tallied directly. We will also apply these methods to determine the direct costs per sequelae, regardless of exposure. Costs will be adjusted for inflation using the Canadian Consumer Price Index.[66] Results will be reported to allow comparability with other estimates (e.g., 2010/2011 Canadian and US dollars).

Objective 4: To determine the risk of sequelae in the population attributable to foodborne infections, we will calculate the proportion of cases of each sequela attributable to the specific foodborne infections. The total number of cases of each sequela occurring in B.C. during the study period will be the denominator (e.g., total number of cases of acute kidney injury), and the numerators will be the numbers of cases of each sequela occurring in those with specific foodborne infections (e.g., total number of cases of hepatitis A, and of STEC, with acute kidney injury). We will also describe the fraction of individuals with sequela who do not have a foodborne infection, but who do have an ICD code for prior gastroenteritis, and use this to

estimate the additional fractions of sequelae that may have an unidentified foodborne infection cause. We will calculate fractions for both established and possible sequelae, but clearly distinguish between the two when reporting findings.

Potential confounders and their adjustment

We will use propensity score matching, and inclusion of potential confounders as covariates in our analyses, as our primary methods to adjust for confounding.[67] The databases in Table 1 include direct measurements of important known, strong confounders (e.g., age, sex), as well as other potential factors (e.g., use of protein pump inhibitors and antibiotics, disease severity). We will consider the following variables as potential confounders: age, sex, local health area, income band/area income (as a proxy for socioeconomic status), month/year, seasonality, immune status (e.g., indication of immunosuppressant drugs, presence of conditions like cancer, pregnancy), use of medications like antibiotics, and Charlson comorbidity index. Because none of our data sources include ethnicity nor race, we are unable to adjust, or conduct sub-analyses, for these factors. Although ethnicity may impact health care seeking behaviours, we anticipate these impacts will apply equally regardless of exposure, and thus we expect any bias in our findings to be negligible. Nevertheless, to assess residual confounding, we will conduct sensitivity analyses,[68-69] and perform indirect adjustments.[70].

Planned sensitivity analyses and study limitations

We are planning several sensitivity analyses to explore assumptions, methodological decisions, limitations in the data, and robustness of results. We will explore the impact of propensity score matching on our sequelae and mortality risk estimates by also using (a) the

whole unexposed population, and (b) a random sample of unexposed individuals (matched on time), instead of propensity score-matched individuals. We may also explore additional matching and control strategies (e.g., matching on age and sex). We will also explore the impacts of including individuals with foodborne infections who develop a sequela for which there is no current evidence of an established or possible link to the specific pathogen (e.g., *Campylobacter* and stillbirth), in our estimates of sequelae risk. Our primary analyses will use a composite of all infections (i.e., a report of any of the 14 foodborne infections) and their various sequelae. We will also analyze and present results for each of the individual foodborne infections, and for each of the sequela.

A main recognized limitation of reportable disease data, such as the Panorama data in this study, is the under-ascertainment of foodborne infections. Here, this limitation means that individuals with foodborne infections who do not seek care nor get tested will be misclassified as unexposed. We will assess the impacts of such potential misclassification via sensitivity analyses that illustrate how our findings could be impacted by different misclassification rates, using estimates of misclassification from the literature,[e.g., 2, 71] and from our data (e.g., individuals with non-specific gastroenteritis). An additional limitation is that if sequelae develop over longer timeframes than our 10-year study (e.g., over decades),[41] our study cannot assess this scenario.

Patient and public involvement statement

Patients were not involved in the development of this protocol, nor were members of the public.

ETHICS AND DISSEMINATION

This study has received approval by a University of Waterloo Research Ethics Committee (#30645), the University of British Columbia Behavioral Research Ethics Board (#H16-00021), and McGill University's Institutional Review Board (#A03-M12-19A). In addition to conference presentations and dissemination to public health practitioners and other knowledge users, results will be published in peer-reviewed journals, and where such publications are not open access, they will also be stored on UWSpace, the University of Waterloo's Institutional Repository (https://uwspace.uwaterloo.ca).

ACKNOWLEDGEMENTS

We thank Bei Yuan (Ethan) Zhang (University of British Columbia) and Dr. Jenna Dixon (University of Waterloo) for helping draft early versions of this paper, and Dr. Yuen Hung (University of Waterloo) for compiling relevant background literature.

Study Team members Drs. Jonathan Chan (University of British Columbia), Kristine Chapman (Vancouver General Hospital), Susan Horton (University of Waterloo), Douglas Matsell (B.C. Children's Hospital Research Institute), and Sara Pires (Technical University of Denmark), served as expert advisors and critically reviewed the study plan.

AUTHORS' CONTRIBUTIONS

Majowicz and Galanis are the co-Principal Investigators on this study. Majowicz, Galanis, and Taylor conceived the study. The overall design was first developed by Majowicz, Galanis, Taylor, and Panagiotoglou, with critical revisions from Cook, Ethelberg, Leatherdale, Kaplan, and Patrick. All coauthors developed the analysis plan, with specific statistical expertise provided by Cook, Chaurasia, and Gohari, and economic expertise by Panagiotoglou.

Majowicz drafted the manuscript; all authors provided feedback on manuscript drafts and approved the final version to be published. All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

FUNDING

This work is supported by the Canadian Institutes of Health Research (CIHR), Project Grant program, grant number 156385 (PI: Shannon Majowicz), and also received seed funding

from the University of Waterloo's Research Incentive Fund (PI: Shannon Majowicz) and the British Columbia Centre for Disease Control Foundation for Public Health's Open Award Program (PI: Eleni Galanis).

DISCLAIMER

All inferences, opinions, and conclusions drawn in this research protocol are those of the authors, and do not reflect the opinions or policies of the Data Steward(s).

COMPETING INTERESTS

Drs. Majowicz and Galanis report funding for this study as per the funding statement. Dr. Majowicz reports other relationships; she is an Associate Editor at Epidemiology and Infection (for which she receives a small honorarium); she has served as a paid Expert on behalf of the Attorney General of Canada in legal proceedings, providing evidence on the public health risks and benefits of unpasteurized milk; and she is an expert on the Joint FAO/WHO Expert Meetings on Microbiological Risk Assessment (JEMRA) Roster of Experts. Dr. Kaplan reports honoraria for speaking or consultancy from Abbvie, Janssen, Pfizer, and Takeda. He has received research support from Ferring, Janssen, Abbvie, GlaxoSmith Kline, Merck, and Shire. He shares ownership of a patent: TREATMENT OF INFLAMMATORY DISORDERS, AUTOIMMUNE DISEASE, AND PBC. UTI Limited Partnership, assignee. Patent WO2019046959A1. PCT/CA2018/051098. 7 Sept. 2018. Dr. Galanis' spouse works for QHR Technologies, a Canadian medical records company; these records were not used in this study. All other authors have nothing to disclose.

DATA SHARING

Open access for these data is not permitted by the data stewards; further details on the legislation and agreements can be found at: https://www.popdata.bc.ca/dataaccess/rdaf/history and https://www.popdata.bc.ca/dataaccess/rdaf/expectations. To access the data used for this study, researchers must submit a Data Access Request through Population Data B.C. (for all databases except Panorama) and the Panorama Data Governance Committee (Panorama data), who have record of the files and fields provided (cite project: Majowicz Galanis 15-180). We will make the programming code used to clean and analyse the data available (on request, or via publications where possible).

REFERENCES

1. Havelaar AH, Kirk MD, Torgerson PR, et al. World Health Organization global estimates and regional comparisons of the burden of foodborne disease in 2010. *PLoS Med* 2015;12(12):e1001923.
2. Thomas MK, Murray R, Flockhart L, et al. Estimates of the burden of foodborne illness in Canada for 30 specified pathogens and unspecified agents, circa 2006. *Foodborne Pathog Dis.* 2013;10(7):639-48.
3. Awofisayo A, Amar C, Ruggles R, et al. Pregnancy-associated listeriosis in England and Wales. *Epidemiol Infect*. 2015;143(2):249-56.
4. Charlier C, Perrodeau É, Leclercq A, et al. Clinical features and prognostic factors of listeriosis: the MONALISA national prospective cohort study. *Lancet Infect Dis*. 2017;17(5):510-9.
5. Karmali MA, Petric M, Lim C, et al. The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing *Escherichia coli*. *J Infect Dis*. 1985;151(5):775-82.
6. Gould LH, Demma L, Jones TF, et al. Hemolytic uremic syndrome and death in persons with *Escherichia coli* O157: H7 infection, Foodborne Diseases Active Surveillance Network sites, 2000–2006. *Clin Infect Dis*. 2009;49(10):1480-5.
7. Rodríguez LA, Ruigómez A, Panés J. Acute gastroenteritis is followed by an increased risk of inflammatory bowel disease. *Gastroenterol*. 2006;130(6):1588-94.
8. Small CL, Xing L, McPhee J, et al. Acute infectious gastroenteritis potentiates a Crohn's disease pathobiont to fuel ongoing inflammation in the post-infectious period. *PLoS Pathog*. 2016. http://dx.doi.org/10.1371/journal.ppat.1005907.

9. Ruzante JM, Majowicz SE, Fazil A, et al. Hospitalization and deaths for select enteric illnesses and associated sequelae in Canada, 2001-2004. *Epidemiol Infect*. 2011;139(6):937-45.
10. Tam CC, Rodrigues LC, Petersen I, et al. Incidence of Guillain-Barré syndrome among patients with *Campylobacter* infection: a general practice research database study. *J Infect Dis*. 2006;194(1):95-7.
11. Helms M, Vastrup P, Gerner-Smidt P, et al. Short and long term mortality associated with foodborne bacterial infections: registry-based study. *Br Med J*. 2003;326:1-5.
12. Bremell T, Bjelle A, Svedhem A. Rheumatic symptoms following an outbreak of campylobacter enteritis: a five year follow up. *Ann Rheum Dis*. 1991;50(12):934-8.
13. Marshall JK, Thabane M, Garg AX, et al. Incidence and epidemiology of irritable bowel syndrome after a large waterborne outbreak of bacterial dysentery. *Gastroenterol.* 2006;131(2):445-50.
14. Garg AX, Pope JE, Thiessen-Philbrook H, et al. Arthritis risk after acute bacterial gastroenteritis. *Rheumatol*. 2008;47(2):200-4.
15. Hanevik K, Wensaas KA, Rortveit G, et al. Irritable bowel syndrome and chronic fatigue 6 years after giardia infection: a controlled prospective cohort study. *Clin Infect Dis*. 2014;59(10):1394-400.
16. Rehn M, Wallensten A, Widerström M, et al. Post-infection symptoms following two large waterborne outbreaks of *Cryptosporidium hominis* in Northern Sweden, 2010–2011. *BMC Public Health*. 2015;15(1):529.
17. Hannu T, Mattila L, Rautelin H, et al. Campylobacter-triggered reactive arthritis: a population-based study. *Rheumatol*. 2002;41(3):312-8.

18. Hunter PR, Hughes S, Woodhouse S, et al J. Health sequelae of human cryptosporidiosis in immunocompetent patients. *Clin Infect Dis*. 2004;39(4):504-10.

19. Wang LH, Fang XC, Pan GZ. Bacillary dysentery as a causative factor of irritable bowel syndrome and its pathogenesis. *Gut*. 2004;53(8):1096-101.

20. Hannu T, Mattila L, Siitonen A, et al. Reactive arthritis attributable to *Shigella* infection: a clinical and epidemiological nationwide study. *Ann Rheum Dis*. 2005;64(4):594-8.

21. Saps M, Pensabene L, Di Martino L, et al. Post-infectious functional gastrointestinal disorders in children. *J Pediatr*. 2008;152(6):812-6.

22. Rosner BM, Werber D, Höhle M, et al. Clinical aspects and self-reported symptoms of sequelae of *Yersinia enterocolitica* infections in a population-based study, Germany 2009–2010. *BMC Infect Dis*. 2013;13(1):236.

23. Litleskare S, Rortveit G, Eide GE, et al. Prevalence of irritable bowel syndrome and chronic fatigue 10 years after *Giardia* infection. *Clin Gastroenterol Hepatol.* 2018;16(7):1064-72.

24. Helms M, Simonsen J, Mølbak K. Foodborne bacterial infection and hospitalization: a registry-based study. *Clin Infect Dis*. 2006;42(4):498-506.

25. Rodríguez LA, Ruigómez A, Panés J. Acute gastroenteritis is followed by an increased risk of inflammatory bowel disease. *Gastroenterol*. 2006;130(6):1588-94.

26. Ternhag A, Törner A, Svensson Å, et al. Short-and long-term effects of bacterial gastrointestinal infections. *Emerg Infect Dis*. 2008;14(1):143.

27. Gradel KO, Nielsen HL, Schønheyder HC, et al. Increased short-and long-term risk of inflammatory bowel disease after *Salmonella* or *Campylobacter* gastroenteritis. *Gastroenterol*. 2009;137(2):495-501.

28. Jess T, Simonsen J, Nielsen NM, et al. Enteric *Salmonella* or *Campylobacter* infections and the risk of inflammatory bowel disease. *Gut*. 2011;60(3):318-24.

29. Porter CK, Choi D, Cash B, et al. Pathogen-specific risk of chronic gastrointestinal disorders following bacterial causes of foodborne illness. *BMC Gastroenterol*. 2013;13(1):46.

30. Riddle MS, Murray JA, Cash BD, et al. Pathogen-specific risk of celiac disease following bacterial causes of foodborne illness: a retrospective cohort study. *Digestive Diseases and Sciences*. 2013;58(11):3242-5.

31. Nakao JH, Collier SA, Gargano JW. Giardiasis and subsequent irritable bowel syndrome: a longitudinal cohort study using health insurance data. *J Infect Dis*. 2017;215(5):798-805.

32. Painter JE, Collier SA, Gargano JW. Association between *Giardia* and arthritis or joint pain in a large health insurance cohort: could it be reactive arthritis? *Epidemiol Infect.* 2017;145(3):471-7.

33. Kaakoush NO, Castaño-Rodríguez N, Mitchell HM, et al. Global epidemiology of *Campylobacter* infection. *Clin Microbiol Rev*. 2015;28(3):687-720.

34. Tricco AC, Duval B, De Serres G, et al. A review of interventions triggered by hepatitis A infected food-handlers in Canada. *BMC Health Serv Res*. 2006;6(1):157.

35. Carpentier B, Cerf O. Persistence of *Listeria monocytogenes* in food industry equipment and premises. *Int J Food Microbiol*. 2011;145(1):1-8.

36. British Columbia Centre for Disease Control. British Columbia Annual Summary of Reportable Diseases, 2017. http://www.bccdc.ca/resource-gallery/Documents/Statistics%20and%20Research/Statistics%20and%20Reports/Epid/Annual%20Reports/2017CDAnnualReportFinal.pdf (accessed 12 November 2019).

37. Scallan E, Majowicz SE, Hall G, et al. Prevalence of diarrhoea in the community in Australia, Canada, Ireland, and the United States. *Int J Epidemiol*. 2005;34(2):454-60.

38. Government of British Columbia. https://www2.gov.bc.ca/gov/content/health/health-drug-coverage/msp/bc-residents/eligibility-and-enrolment/are-you-eligible (accessed 12 November 2019).

39. Population Data BC. About PopData. https://www.popdata.bc.ca/aboutus (accessed 12 November 2019).

40. Government of British Columbia. Public Health Act [SBC 2008] Chapter 28. http://www.bclaws.ca/civix/document/id/complete/statreg/08028_01 (accessed 12 November 2019).

41. Moorin RE, Heyworth JS, Forbes GM, et al. Long-term health risks for children and young adults after infective gastroenteritis. *Emerg Infect Dis*. 2010;16(9):1440.

42. Population Data BC. Data Linkage. https://www.popdata.bc.ca/datalinkage (accessed 12 November 2019).

43. British Columbia Centre for Disease Control. Case Definitions. http://www.bccdc.ca/health-professionals/clinical-resources/case-definitions (accessed 12 November 2019).

44. World Health Organization. ICD-10 Version:2016. https://icd.who.int/browse10/2016/en (accessed 12 November 2019).

45. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. Springer Science & Business Media 2013.

46. Cook RJ, Lawless JF. The Statistical Analysis of Recurrent Events. New York, NY: Springer-Verlag, 2007.

47. Batz MB, Henke E, Kowalcyk B. Long-term consequences of foodborne infections. *Infect Dis Clinics N Am*. 2013;27(3):599-616.

48. Hu W. Diagnostic Codes in MSP Claim Data, Summary Report. Victoria: Medical Services Plan; 1996.

49. Galanis E, Taylor M, Romanowski K, et al. Evaluating the timeliness of enteric disease surveillance in British Columbia, Canada, 2012-13. *Can J Infect Dis Med Microbiol.* 2017 Article ID 9854103, 7 pages, 2017.

50. Klein JP, Moeschberger ML. Survival Analysis: Techniques for Censored and Truncated Data, 2nd ed. New York, NY: Springer-Verlag New York Inc., 2003.

51. Qin J, Shen Y. Statistical methods for analyzing right-censored length-biased data under Cox model. *Biometrics*. 2010;66(2):382–392.

52. von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ*. 2007;335(7624):806-8.

53. Benchimol EI, Smeeth L, Guttman A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Med.* 2015;12(10): e1001885. https://doi.org/10.1371/journal.pmed.1001885

54. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

55. Cox DRJ. Regression models and life-tables (with discussion). *J R Stat Soc, B*. 1972;34:187-220.

56. Charlson ME, Pompeo P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chron Dis*. 1987;40(5):373-83.

57. Charlson M, Szatrowski TP, Peterson J, et al. Validation of a combined comorbidity index. *J Clin Epidemiol*. 1884;47(11):1245-51.

58. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care.* 2005;43(11):1130-9.

59. Hude Quan, Bing Li, Chantal M, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *Am J Epidemiol*. 2011;173(6):676–82.

60. Kaplan EL, Meier PJ. Nonparametric estimation for incomplete observations. *J Am Stat Assoc*. 1958;53:457-81.

61. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc.* 1999;94(446):496-509.

62. Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. *Am J Epidemiol.* 2009;170(2):244-256.

63. Government of British Columbia. Medical Services Commission Payment Schedule. December 1, 2015. https://www2.gov.bc.ca/assets/gov/health/practitioner-pro/medical-services-plan/msc_payment_schedule.pdf (accessed 12 November 2019).

64. Poole B, Robinson S, MacKinnon M. Resource Intensity Weights™ and Canadian hospital costs: some preliminary data. *Healthc Manage Forum*. 1998:11(1):22-26.

65. Canadian Institute for Health Information. Patient Cost Estimator: Methodological Notes and Glossary. Ottawa, ON: CIHI; 2016. https://www.cihi.ca/en/pce_methodology_notes_en.pdf (accessed 26 November 2019).

66. Statistics Canada, Consumer Price Index (CPI). http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=2301 (accessed 12 November 2019).

67. Holmes WM. Using Propensity Scores in Quasi-Experimental Designs. Los Angeles, Calif.: Sage Publishing, 2014.

68. Stürmer T, Wyss R, Glynn RJ, et al. Propensity scores for confounder adjustment when assessing the effects of medical interventions using nonexperimental study designs. *J Int Med*. 2014;275(6):570-80.

69. Brookhart MA, Schneeweiss S, Rothman KJ, et al. Variable selection for propensity score models. *Am J Epidemiol*. 2006;163(12):1149-56.

70. Kuenzig ME, Barnabe C, Seow CH, et al. Asthma is associated with subsequent development of inflammatory bowel disease: a population-based case–control study. *Clin Gastroenterol Hepatol*. 2017;15(9):1405-12.

71. MacDougall L, Majowicz S, Doré K, et al. Under-reporting of infectious gastrointestinal illness in British Columbia, Canada: who is counted in provincial communicable disease statistics? *Epidemiol Infect*. 2008;136(2):248-56.

72. BC Ministry of Health [creator] (2018): Medical Services Plan (MSP) Payment Information File. Population Data BC [publisher]. Data Extract. MOH (2018). http://www.popdata.bc.ca/data

73. Canadian Institute for Health Information [creator] (2019): Discharge Abstract Database (Hospital Separations). Population Data BC [publisher]. Data Extract. MOH (2018). http://www.popdata.bc.ca/data

74. BC Ministry of Health [creator] (2019): PharmaNet. BC Ministry of Health [publisher]. Data Extract. Data Stewardship Committee (2018). http://www.popdata.bc.ca/data

75. BC Vital Statistics Agency [creator] (2019): Vital Statistics Deaths. Population Data BC [publisher]. Data Extract BC Vital Statistics Agency (2018). http://www.popdata.bc.ca/data

76. BC Vital Statistics Agency [creator] (2019): Vital Statistics Stillbirths. Population Data BC [publisher]. Data Extract. BC Vital Statistics Agency (2018). http://www.popdata.bc.ca/data

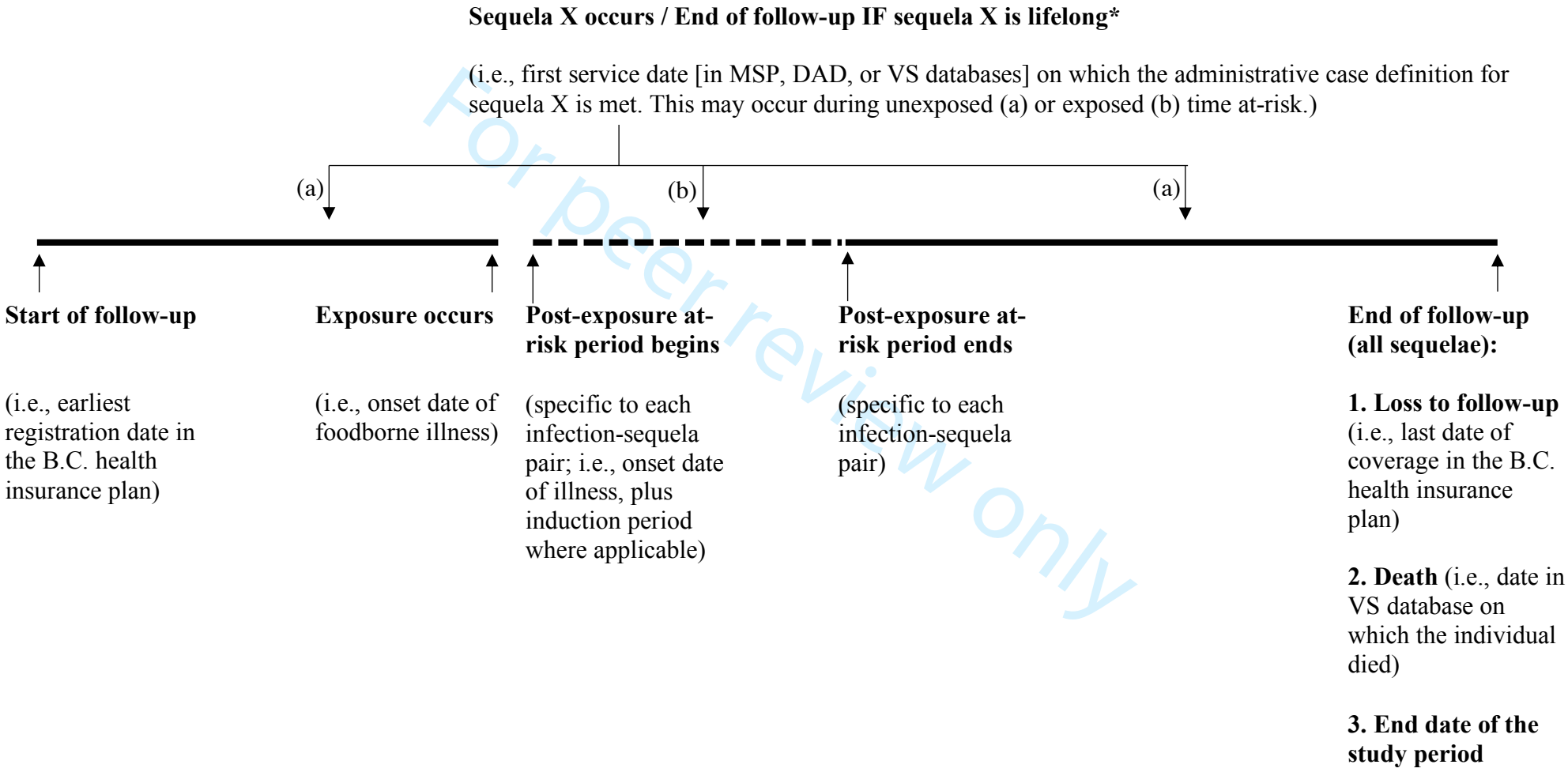
77. British Columbia Ministry of Health [creator] (2019): Consolidation File (MSP Registration & Premium Billing). Population Data BC [publisher]. Data Extract. MOH (2018). http://www.popdata.bc.ca/data

78. Statistics Canada [creator] (2009): Statistics Canada Income Band Data. Catalogue Number: 13C0016. Population Data BC [publisher]. Data Extract. Population Data BC (2018). http://www.popdata.bc.ca/data

FIGURE LEGEND

Figure 1. **Study follow-up period and time-at-risk for development of Sequela X (solid lines: unexposed time at-risk; dashed lines: exposed time at-risk; B.C., British Columbia, Canada; MSP: Medical Services Plan; DAD: Discharge Abstracts Database; VS: Vital Statistics)**

FIGURE 1.



* For sequelae where recovery and return to being at-risk is possible: this date + recovery time for sequela X = date of return to being at risk for a subsequent occurrence of sequela X.

BMJ Open

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

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-036560.R1
Article Type:	Protocol
Date Submitted by the Author:	19-May-2020
Complete List of Authors:	Majowicz, Shannon E. ; University of Waterloo, School of Public Health and Health Systems Panagiotoglou, Dimitra; McGill University, Department of Epidemiology, Biostatistics and Occupational Health Taylor, Marsha; BCCDC Gohari, Mahmood; University of Waterloo, School of Public Health and Health Systems Kaplan, Gilaad; University of Calgary, Departments of Medicine and Community Health Chaurasia, Ashok; University of Waterloo, School of Public Health and Health Systems Leatherdale, Scott T.; University of Waterloo, School of Public Health and Health Systems Cook, Richard; University of Waterloo, Department of Statistics and Actuarial Science Patrick, David; BCCDC; The University of British Columbia, School of Population and Public Health Ethelberg, Steen; Statens Serum Institut, Department of Infectious Disease Epidemiology and Prevention; University of Copenhagen Department of Public Health, Global Health Section Galanis, Eleni; BCCDC; The University of British Columbia, School of Population and Public Health
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Infectious diseases, Public health
Keywords:	Epidemiology < INFECTIOUS DISEASES, Public health < INFECTIOUS DISEASES, Gastrointestinal infections < GASTROENTEROLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

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

Shannon E. Majowicz[1]*, Dimitra Panagiotoglou[2], Marsha Taylor[3], Mahmood R. Gohari[1], Gilaad G. Kaplan[4], Ashok Chaurasia[1], Scott T. Leatherdale[1], Richard J. Cook[5], David M. Patrick[3,6], Steen Ethelberg[7,8], Eleni Galanis[3,6]

[1] School of Public Health and Health Systems, University of Waterloo (Waterloo, Canada)

[2] Department of Epidemiology, Biostatistics and Occupational Health, McGill University (Montreal, Canada)

[3] British Columbia Centre for Disease Control (Vancouver, Canada)

[4] Departments of Medicine & Community Health Sciences, University of Calgary (Calgary, Canada)

[5] Department of Statistics and Actuarial Sciences, University of Waterloo (Waterloo, Canada)

[6] School of Population and Public Health, University of British Columbia (Vancouver, Canada)

[7] Department of Infectious Disease Epidemiology and Prevention, Statens Serum Institut (Copenhagen, Denmark)

[8] Global Health Section, Department of Public Health, University of Copenhagen (Copenhagen, Denmark)

*Corresponding Author:
200 University Avenue West
Waterloo, Ontario, Canada, N2L 3G1
Email: smajowicz@uwaterloo.ca
Telephone: +1 519 888 4567 extension 31790

KEY WORDS: foodborne infections, sequelae, cohort study, public health

ABSTRACT

Introduction: Over 1 in 8 Canadians is affected by a foodborne infection annually; however, the long-term consequences, including the risks and costs of sequelae, are unclear. We aim to estimate the health burden and direct costs of 14 infections commonly transmitted by food, considering the acute illness and subsequent sequelae and mortality, for the population of British Columbia (B.C.), Canada (~4.7 million).

Methods and analysis: We will conduct a population-based retrospective cohort study of the B.C. provincial population, over a 10-year study period (January 1, 2005 to December 31, 2014). Exposure is defined as a provincially-reported illness caused by: *Clostridium botulinum*, *Campylobacter*, *Cryptosporidium, Cyclospora, Giardia*, hepatitis A virus, *Listeria*, non-typhoidal *Salmonella* spp., *Salmonella* Typhi, *Salmonella* Paratyphi, Shiga toxin-producing *E. coli*, *Shigella*, *Vibrio parahaemolyticus*, or *Yersinia* (excluding *pestis*). We will link individual-level longitudinal data from eight province-wide administrative health and reportable disease databases that include physician visits, hospitalizations and day surgeries, deaths, stillbirths, prescription medications (except those to treat HIV), and reportable foodborne diseases. Using these linked databases we will investigate the likelihood of various sequelae and death. Hazard models will be used to estimate the risk of outcomes and their association with the type of foodborne infection. Epidemiologic analyses will be conducted to determine the progression of illness and the fraction of sequelae attributable to specific foodborne infections. Economic analyses will assess the consequent direct healthcare costs.

Ethics and dissemination: This study has been approved by a University of Waterloo Research Ethics Committee (#30645), the University of British Columbia Behavioral Research Ethics Board (#H16-00021), and McGill University's Institutional Review Board (#A03-M12-19A).

Results will be disseminated via presentations to academics, public health practitioners, and knowledge users, and publication in peer-reviewed journals. Where such publications are not open access, manuscripts will also be available via the University of Waterloo's Institutional Repository (https://uwspace.uwaterloo.ca).

STRENGTHS AND LIMITATIONS

- This cohort is a near-complete set of individually-linked administrative health and reportable foodborne infection data, covering the ~4.7 million residents of British Columbia, Canada over 10 years (2005–2014).
- To the best of our knowledge, the study described in this protocol will be the most comprehensive assessment of the risk of sequelae following foodborne infections across multiple pathogens to-date.
- Because all residents of British Columbia population are covered by a mandatory, single provincial health insurance plan (with only a few exceptions, e.g., members of the military), movement of individuals within the province or between employers does not create loss to follow-up.
- Limitations include incomplete and lower quality (e.g., misclassification, use of non-specific codes) information associated with administrative health data, and under-ascertainment of foodborne infections typical to reportable disease data.

INTRODUCTION

Infections commonly transmitted via food, such as *Salmonella* spp. and Shiga toxin-producing *Escherichia coli* (STEC) are a global public health concern,[1] and in Canada they affect over 1 in 8 people annually.[2] Beyond the acute stage of illness (typically characterized by diarrhea and other gastrointestinal symptoms), these infections can cause severe and longer-term outcomes, such as spontaneous abortion, hemolytic uremic syndrome (HUS), Crohn's disease and ulcerative colitis (inflammatory bowel disease [IBD]), Guillain-Barré syndrome (GBS), and death.[3-11]

Estimates of the risk of sequelae following foodborne infection have come in part from prospective cohort studies conducted as follow-ups to outbreaks,[12-16] or to reports of sporadic cases (usually via laboratory testing and disease surveillance systems).[4, 17-23] While such prospective studies have the advantage of being able to tailor the data collection to address specific research questions, they have some important limitations. For example, outbreak follow-up studies are limited to specific strain(s) causing the outbreak and the specific population affected. Studies of sporadic cases often have short follow-up periods and rely on self-reported questionnaires to identify sequelae, both of which can lead to bias.

Retrospective, population-based cohort studies, in which administrative and registry data are analysed,[10-11, 24-32] offer several advantages that complement the prospective studies described above. They allow for a wider population to be covered, both sporadic and outbreak-associated infections (caused by the range of strains affecting the population) to be included, and the use of self-reports of event occurrence to be avoided. However, because they require population-wide, linked data on the exposures and outcomes of interest, they are less frequently conducted. To-date, such studies have not been conducted in Canada.

Although infections such as *Salmonella* spp. and STEC can be transmitted via several routes (e.g., person-to-person, water), their transmission via food and their presence throughout the food system (e.g., food animals as a reservoir for *Campylobacter* spp.,[33] food handlers shedding hepatitis A virus,[34] the ability of *Listeria monocytogenes* to persist in food production equipment [35]) mean that these infections are often termed "foodborne" although some fraction will not be transmitted via food directly. Here, we apply the term "foodborne infection" to 14 infections that can be transmitted via food (*Clostridium botulinum*, *Campylobacter*, *Cryptosporidium, Cyclospora, Giardia*, hepatitis A virus, *Listeria*, non-typhoidal *Salmonella* spp., *Salmonella* Typhi, *Salmonella* Paratyphi, STEC, *Shigella*, *Vibrio parahaemolyticus*, and *Yersinia* excluding *pestis*), recognizing that not all result from direct foodborne transmission.

The overall goal of this study is to estimate the health burden and costs of these 14 infections, considering the acute illness and subsequent sequelae and associated mortality, for the population of British Columbia (B.C.), Canada. Our specific objectives are to:

1. determine the risk of developing sequelae following infection;
2. describe the epidemiology and clinical progression across the range of outcomes, including acute illness, sequelae, and death;
3. quantify the direct healthcare costs due to these infections and their various outcomes; and
4. determine the risk of sequelae in the population attributable to these infections.

METHODS AND ANALYSIS

Study setting

B.C. is Canada's westernmost and third most populous province (~4.7 million circa 2014). The annual incidence of foodborne infections in B.C. is comparable to annual incidences in Canada, the United States, Australia, and Western Europe.[e.g., 2, 36-37]

All B.C. residents (defined as citizens or permanent residents of Canada who are physically present in B.C. for at least six months in a calendar year), their dependents, and certain other individuals (e.g., some holders of study or work permits) are covered by the province's health insurance program.[38] Enrolment is mandatory, and this program covers nearly all of the B.C. population (with the exception of members of the Canadian military, Royal Canadian Mounted Police, and some First Nations individuals covered by federal insurance programs). Physician visits, laboratory tests, hospitalizations, out-patient hospital services, and drugs for certain populations are among the publicly-funded benefits. The administrative datasets that contain these health care use data, along with vital statistics (e.g., births, deaths), demographic data, and other datasets, are held by the B.C. Ministry of Health and are accessible to researchers via Population Data B.C.,[39] a central repository and "multi-university, data and education resource" that "support[s] research access to individual-level, de-identified longitudinal data on British Columbia's 4.7 million residents".[40]

The B.C. Public Health Act mandates that reportable diseases, including several foodborne infections,[41] be reported by health professionals and laboratories to the local and provincial public health authorities, and these data are managed provincially by the B.C. Centre for Disease Control (BCCDC). These data are housed within Panorama, the provincial public health database of reportable diseases. In the Panorama database, as well as the administrative health and vital statistics databases, individuals are recorded by their unique Personal Health Number (PHN), allowing information from these data sources to be linked by individual.

Study design, population, and timeframe

This is a retrospective cohort study of the population of B.C., with additional descriptive, cost, and population attributable risk analyses. Because this is a dynamic population in which exposure status of individuals changes over time, our study design assesses the risk and effect of exposure in terms of person-time. Thus, rather than using fixed cohorts of exposed versus unexposed individuals, we will track individuals over time and assign their person-time at risk to either "unexposed person-time" (e.g., prior to foodborne infection) or "exposed person-time" (e.g., after foodborne infection), as described further below.

The study population is all individuals in B.C. registered with the provincial health insurance program at any point during the study period, i.e., all individuals with the following from 2005 to 2014 inclusive: one or more record in one or more of the Medical Services Plan (MSP), Discharge Abstracts Database (DAD), Vital Statistics Deaths, or PharmaNet; or record of coverage under the provincial insurance program within the Consolidation File database (see Table 1). The 10-year study period is January 1, 2005 to December 31, 2014, inclusive, with additional two-year wash-in (January 1, 2003 to December 31, 2004) and wash-out (January 1, 2015 to December 31, 2016) periods. During these periods we will identify occurrences of foodborne infections, sequelae, and death. The 10-year study period was selected to more than encompass timeframes for initial sequelae development and ensuing healthcare use currently reflected in the literature (i.e., days to years), although there is some evidence that sequelae can develop over longer timeframes (e.g., over decades).[42]

We assume that enrolment in the provincial health insurance program (i.e., entry into the study population) and reasons for exit from the cohort (e.g., moving away from B.C.) are not related to the exposures nor the outcomes of interest.

Data sources and linkage

The study will use individually-linked, longitudinal data from eight databases to investigate both acute and longer-term health outcomes following foodborne infection (Table 1). In totality, these data contain information on 14 reportable foodborne infections, physician and hospital visits, prescription medications, vital statistics, and various demographic descriptors, for the B.C. population across the study period. All data will be stored and analyzed within Population Data B.C.'s virtual Secure Research Environment.

Table 1. Population-level administrative and reportable disease databases that will be used in this study (British Columbia [B.C.], Canada)

Database (Reference)	Database description and summary of variables included for this study	Date range
Health Care and Health Services Data		
Medical Services Plan Payment Information File [43]	Billing records for all medically necessary services provided by fee-for-service physicians. Includes PHN, service dates, up to five ICD*-9/ICD-10 diagnostic codes, MSP-specific fee-item codes, and physician specialties.	2003/01/01 to 2016/12/31
Discharge Abstracts Database (Hospital Separations) [44]	Data on discharges, transfers and deaths of in-patients and day surgery patients from acute care hospitals in B.C.; does not include emergency room visits. Includes PHN, admission and discharge dates, up to 25 ICD-10-CA diagnostic codes, service use and procedure codes, newborn and maternal data, discharge status, and province issuing health care number.	2003/01/01 to 2016/12/31
PharmaNet [45]	All prescriptions (for drugs and medical supplies) dispensed from community pharmacies, and from hospital outpatient	2003/01/01 to 2016/12/31

	pharmacies for patient use at home, in B.C. Includes PHN, date of dispensing, quantity, dose, costs/fees, the type of prescribing practitioner, and the Drug Information Number (or Product Information Number) assigned by Health Canada.	
Population and Vital Statistics Data		
Vital Statistics Deaths [46]	All deaths registered in B.C. Includes PHN, time and place data, and ICD-10 codes for the nature and causes of death.	2005/01/01 to 2016/12/31
Vital Statistics Stillbirths [47]	All stillbirths registered in B.C. Includes the mother's PHN, time and place data, number of stillborn and live born children in the event, gestation period, and ICD-10 codes for the underlying cause of stillbirth.	2005/01/01 to 2016/12/31
Consolidation File [48]	Population Data B.C.'s central demographics file for research requests, containing all individuals who are eligible to receive services in B.C. Includes PHN, age, sex, neighbourhood income deciles, and local health authority area, as well as the date and number of days registered in the provincial health insurance program.	2003/01/01 to 2016/12/31
Statistics Canada Income Bands [49]	1000 income bands that contain information about the 6-digit postal code area in which the individual resides. Includes the average and median equivalised disposable income (derived from Statistics Canada tax-filer data, and available for the years 1992, 2002, and 2006), and the number of families, adults, and children in the area.	2002, 2006
Reportable Disease Data		
Panorama Public Health Information System	All cases of the following 14 reportable diseases reported in B.C.: *Clostridium botulinum*, *Campylobacter*, *Cryptosporidium, Cyclospora, Giardia*, hepatitis A virus, *Listeria*, non-typhoidal *Salmonella* spp., *Salmonella* Typhi, *Salmonella* Paratyphi, STEC, *Shigella*, *Vibrio parahaemolyticus*, and *Yersinia*. Includes PHN, onset date, reported date, health authority, and etiologic agent.	2005/01/01 to 2014/12/31

*ICD: International Classification of Diseases

Population Data B.C. will link the eight databases directly by individual using PHN, with additional identifiers (e.g., name, age, sex) used to validate linkages and link records probabilistically when PHNs are missing. Population Data B.C.'s detailed linkage process is fully described elsewhere.[50] Note that because the Statistics Canada Income Bands database contains area-level data (whereas the other seven databases contain individual-level data), these

data are first linked to individuals (and their PHNs) using their 6-digit postal code. Once the linkage is complete, each individual is then assigned a unique study identifier. All individually-linked, de-identified databases are provided by Population Data B.C. within their Secure Research Environment, a centralized online platform, accessible via virtual private network within Canada, for accessing and analyzing research data, with security standards that meet Data Steward requirements.

Measuring exposure and outcomes

The exposures of interest are infections with the 14 foodborne pathogens. These 14 infections were selected because they are (a) considered a priority in terms of prevention potential and health impacts, and (b) capture nearly all reportable foodborne infections in B.C. Note that brucellosis and paralytic shellfish poisoning were also reportable foodborne infections in B.C. during the study period. However, since brucellosis is very rare and nearly always travel-related, and paralytic shellfish poisoning (also rare) is syndromic and diagnosis is uncertain, these two were not considered for inclusion in this study.

Individuals will be considered exposed when and if they have a laboratory-confirmed and provincially-reported case of any of the following, recorded in the Panorama database during the study period: *Clostridium botulinum*, *Campylobacter*, *Cryptosporidium, Cyclospora, Giardia*, hepatitis A virus, *Listeria*, *Salmonella* spp. (non-typhoidal, Typhi, Paratyphi), STEC, *Shigella*, *Vibrio parahaemolyticus*, and *Yersinia* (excluding *pestis*). Case definitions for each of these infections are specified by the BCCDC.[51]. Individuals without a reported foodborne infection, but who have International Classification of Disease (ICD) codes either for one of our infections (e.g., A02.0, '*Salmonella* enteritis') or for non-specific gastroenteritis (e.g., A08.4, 'viral

intestinal infection, unspecified' [52]) within the MSP and DAD databases, will be considered potentially exposed. We will describe these individuals as a separate group in our descriptive, economic, and population attributable fraction analyses, but will remove them from the main analyses of sequelae risk. We will, however, estimate the risk of sequelae among those who are potentially exposed as a secondary analysis.

It is possible for individuals to have more than one reported foodborne infection during the study period, either as a simultaneously occurring co-infection, or as two or more distinct events. For these individuals, we will treat this as a complex exposure problem; sequelae will be associated with the most plausible explanatory infection, considering biology and timing, and we will adjust for the presence of concurrent foodborne infections if applicable.[53-54]

Our primary outcomes of interest are those sequelae for which the link to a given foodborne infection is either established or is possible (Table 2). We selected sequelae (a) with evidence of an association with any of the 14 individual foodborne infections,[e.g., 55] and (b) that occur via direct effects of pathogens or their toxin, or via auto-immune or chronic inflammatory processes that can be triggered by the infection. We will classify individuals as having the sequelae via administrative case definitions, using International Classification of Disease (ICD) diagnostic codes within the MSP, DAD, and VS data, with the exception of stillbirths which will be determined using recorded events in the VS-Stillbirths database. The ICD codes in the MSP data are generally considered accurate to the third digit.[56] Although ICD codes in the VS and DAD data are ordered by most probable diagnosis, we will consider all codes, regardless of order.

Validation of the ICD codes is currently in progress, via a literature review to identify administrative case definitions that have been validated in the Canadian context, medical expert

consultation, and, for those sequelae without a relevant validated definition, a targeted chart review in B.C. In addition to the ICD codes, we may also use PharmaNet data to improve sequelae classification if needed, by identifying pharmaceuticals given to patients (e.g., Tumor Necrosis Factor inhibitor use as an indication of reactive arthritis; intravenous immune globulin use for GBS).

Some of our sequelae of interest are lifelong (e.g., Graves' Disease), and some are transient in that complete recovery is possible (e.g., GBS, stillbirth). For lifelong sequelae, we will consider the individual as having the sequela on the earliest date they meet the administrative case definition for that sequela (with subsequent records considered as a continuation of the original event). For sequelae from which recovery and subsequent return to being at-risk is possible, we will consider the individual as first having the sequela on the earliest date they meet the administrative case definition for that sequela; we will then apply a post-sequela recovery time to determine the date on which the individual can be considered to be at-risk for a new, subsequent occurrence of that sequela.

Individuals may develop more than one sequela during the study period, either because they develop multiple different sequelae (e.g., HUS and GBS), or because they develop multiple occurrences of a single sequela from which complete recovery is possible (e.g., GBS, stillbirth). In all instances, the occurrence of multiple sequelae will be recorded and described. When individuals develop multiple different sequelae during the study period (e.g., HUS and GBS), we will treat these as distinct outcomes in our risk estimates. When individuals develop multiple occurrences of the same sequela, we will treat these as distinct outcomes but account for recurrent events.[54]

Individuals with foodborne infections who develop a sequela listed in Table 2, but for which there is no current evidence of an established or possible link to the specific pathogen (e.g., *Campylobacter* and stillbirth), will be excluded from our estimates of sequelae risk (but included in sensitivity analyses).

For all 14 infections the secondary outcome of interest is death, which will be classified using recorded events in the VS-Mortality database. Finally, we are also interested in (a) the acute illnesses related to these infections (regardless of whether the individual develops sequelae or dies), and (b) additional outcomes following acute sequelae (e.g., end-stage kidney disease and kidney transplant following HUS); these will be include only in our descriptive and economic analyses.

Table 2. Established (E) and possible (P) sequelae of foodborne infections, that will be assessed in this study (British Columbia, Canada)

Foodborne Infection [1]	**Acute Kidney Injury**	**Celiac Disease**	**Erythema nodosum**	**Graves' Disease**	**Guillain-Barré syndrome** [2]	**Hemolytic Uremic Syndrome**	**Inflammatory Bowel Disease**	**Irritable Bowel Syndrome**	**Neonatal Listeriosis** [3]	**Stillbirth**	**Reactive Arthritis** [4]	**Thrombotic thrombocytopenic purpura** [5]
Campylobacter	E	P			E	P	P	P			E	

Cryptosporidium											P	
Cyclospora					P						P	
Giardia								P			P	
Hepatitis A virus	E				P							
Listeria monocytogenes									E [5]	E		
Salmonella (non-typhoidal)	E				P	P	P	P			E	
Salmonella Paratyphi	E				P	P	P	P			E	
Salmonella Typhi	E				P	P	P	P			E	
STEC	E					E					P	P
Shigella	E					E		P			E	P
Yersinia (excluding *pestis*)	E		E	P				P			E	

[1] *Clostridium botulinum* and *Vibrio parahemolyticus* do not have established or possible sequelae

[2] this includes GBS variants such as Miller Fischer syndrome; other neurological conditions such as chronic inflammatory demyelinating polyneuropathy will also be assessed

[3] considered here as a sequela of maternal *Listeria* infection

[4] this includes associated diagnoses such as anterior uveitis and ankylosing spondylitis

[5] shown not to be a sequela; retained to capture historical misdiagnosis of HUS

Measuring time at-risk

For all individuals, time-at-risk for sequelae (Figure 1) will be measured from the start of their entry into the study, which we define as the earliest registration date in the provincial health insurance program (recorded in the Consolidation File). Individuals with foodborne infections

may contribute to both the exposed time-at-risk (during the post-infection 'at-risk' period, see below) and the unexposed time-at-risk (prior to, and after, the post-infection 'at-risk' period), while individuals without foodborne infections will only contribute to the unexposed time-at-risk. Time-at-risk for a specific sequela will be measured in days, from the date of entry into the study, until: the development of that sequela, death, loss to follow-up, or the end of the study. We define loss to follow-up as the last date of coverage in the provincial health insurance plan, calculated using the start day registered in the most recent year plus the total days registered in that year.

During the unexposed time-at-risk, we will treat all individuals as having the potential to develop any of the sequelae (with the exception of neonatal listeriosis and stillbirth, for which only those who are pregnant are at risk). For those who develop a foodborne infection, unexposed time at-risk will end on the onset date of the infection. Infection onset date will be determined using the onset date reported in Panorama, and where this is missing, the date that the infection was reported minus the number of days between onset to reporting (e.g., estimated using the Panorama data or from the literature).[57]

Exposed time at-risk will be measured starting from the infection onset date, plus any additional induction periods (specific to each sequela and currently being determined via literature review and medical expert consultation). The end of the exposed time at-risk period is currently being determined via literature review and medical expert consultation. During the exposed time-at-risk, individuals will be classified as having a sequela specific to their infection (Table 2) on the date within the 'at-risk' period on which they meet the administrative case definition for that sequela (e.g., the date of the physician visit or hospitalization). After the post-infection at-risk period ends, individuals will revert to contributing to the unexposed time-at-risk.

These data are subject to censoring and truncation. Individuals will be censored for the sequela in the event of: death, last date of coverage in the provincial health insurance plan (i.e., loss to follow up; calculated as above), or the end of the study period, whichever comes first.[58-59] In our descriptive and economic analyses, we will include all related health care use and prescription medication costs over the course of the infection and sequela(e), and in our estimates of mortality we will include any deaths recorded during the study period, following the sequela. 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.

Analysis plan

Data will be analyzed and results reported following the STROBE and RECORD guidelines.[60-61]. Analyses will be performed using SAS 9.4 (SAS Institute Inc., Cary, NC, USA) and R.[62] The nature and extent of missing data will be described. If imputation is used to complete missing data, specific methods and assumptions will be reported. We will emphasize estimation over tests of statistical significance by reporting relative measures of effect along with associated 95% confidence intervals.

The datasets in Table 1 contain the variable "sex" (identified via government records), that denotes whether individuals are 'male' or 'female', thereby capturing a composite of sex and gender. To reflect for potential sex- and gender-differences, we will report and interpret findings stratified by this variable, in addition to overall findings.

Objective 1: To determine the risk of developing sequelae following foodborne infection, we will estimate hazard ratios using Cox regression models,[63] that adjust for confounders and

comorbidities (see below), along with the possible effect modifying role of age, sex, comorbidity, and medication use (e.g., antibiotics). We will include comorbidities as a composite score, using the revised Charlson comorbidity index and its associated coding algorithms.[64-67] Following foodborne infection, we will compare the cumulative risk of first diagnosis of each infection-specific sequela using life-table and Kaplan-Meier approaches.[63, 68] In the event an individual dies, we will use competing risk analysis.[69-70] For those who experience more than one foodborne infection across the study period, we will explore the impacts of having multiple infections on the risk of sequelae and mortality. We will also estimate the likelihood of dying following foodborne infection using the same methods described above.

Objective 2: To describe the epidemiology and clinical progression across the range of outcomes, including acute illness (typically diarrhea and other gastrointestinal symptoms), sequelae, and death, for each foodborne infection we will calculate incidence rates, demographic, geographic, and temporal distributions, timing and progression of outcomes, and case fatality rates, for both the acute stage, and sequelae associated to the foodborne infection.

Objective 3: To quantify the direct healthcare costs due to these infections and their various outcomes, we will determine health service use (i.e., patterns of use by type, frequency, timing of physician visits and hospitalizations), for both the acute foodborne infection and any sequelae. We will estimate direct healthcare costs of out-of-hospital physician visits, hospitalizations, and prescription medications. Out-of-hospital costs will be estimated using the MSP variables 'Fee Item' and 'Paid Service' and fee rates from the B.C. fee schedule.[71] Costs of in-patient and day-case hospitalizations will be calculated using established case-mix methodology (i.e., using the 'Resource Intensity Weight' of each hospitalization),[72] and the B.C. Ministry of Health unit costs for hospital stays.[73] Total prescription medication costs will

be calculated using the drug cost claimed by the pharmacist which includes the ingredient cost, professional dispensing fee and other special service fees (if applicable). Because these costs are captured directly in the PharmaNet data, they will be tallied directly. We will also apply these methods to determine the direct costs per sequelae, regardless of exposure. Costs will be adjusted for inflation using the Canadian Consumer Price Index.[74] Results will be reported to allow comparability with other estimates (e.g., 2010/2011 Canadian and US dollars).

Objective 4: To determine the risk of sequelae in the population attributable to foodborne infections, we will calculate population attributable fractions using standard formulae.[75] We will also describe the proportion of cases of each sequela with specific foodborne infections. Here, the total number of cases of each sequela occurring in B.C. during the study period will be the denominator (e.g., total number of cases of acute kidney injury), and the numerators will be the numbers of cases of each sequela occurring in those with specific foodborne infections (e.g., total number of cases of hepatitis A virus, and of STEC, with acute kidney injury). We will also describe the proportion of individuals with sequela who do not have a foodborne infection, but who do have an ICD code for prior gastroenteritis, and use this to estimate the additional proportions of sequelae that may have an unidentified foodborne infection. We will calculate population attributable fractions and proportions for both established and possible sequelae, but clearly distinguish between the two when reporting findings.

Potential confounders and their adjustment

We will use propensity score matching, and inclusion of potential confounders as covariates in our analyses, as our primary methods to adjust for confounding.[76] The databases in Table 1 include direct measurements of important known, strong confounders (e.g., age, sex),

as well as other potential factors (e.g., use of protein pump inhibitors and antibiotics, disease severity). We will consider the following variables as potential confounders: age, sex, local health area, income band/area income (as a proxy for socioeconomic status), month/year, seasonality, immune status (e.g., indication of immunosuppressant drugs, presence of conditions like cancer, pregnancy), use of medications like antibiotics, and Charlson comorbidity index. Because none of our data sources include ethnicity nor race, we are unable to adjust, or conduct sub-analyses, for these factors. Although ethnicity may impact health care seeking behaviours, we anticipate these impacts will apply equally regardless of exposure, and thus we expect any bias in our findings to be negligible. Nevertheless, to assess residual confounding, we will conduct sensitivity analyses,[77-78] and perform indirect adjustments.[79].

Planned sensitivity analyses and study limitations

We are planning several sensitivity analyses to explore assumptions, methodological decisions, limitations in the data, and robustness of results. We will explore the impact of propensity score matching on our sequelae and mortality risk estimates by also using (a) the whole unexposed population, and (b) a random sample of unexposed individuals (matched on time), instead of propensity score-matched individuals. We may also explore additional matching and control strategies (e.g., matching on age and sex). We will also explore the impacts of including individuals with foodborne infections who develop a sequela for which there is no current evidence of an established or possible link to the specific pathogen (e.g., *Campylobacter* and stillbirth), in our estimates of sequelae risk. Our primary analyses will use a composite of all infections (i.e., a report of any of the 14 foodborne infections) and their various sequelae. We

will also analyze and present results for each of the individual foodborne infections, and for each of the sequela.

A main recognized limitation of reportable disease data, such as the Panorama data in this study, is the under-ascertainment of foodborne infections. Here, this limitation means that individuals with foodborne infections who do not seek care, do not get tested, or who test negative, will be misclassified as unexposed. We will assess the impacts of such potential misclassification via sensitivity analyses that illustrate how our findings could be impacted by different misclassification rates, using estimates of misclassification from the literature,[e.g., 2, 80] and from our data (e.g., individuals with non-specific gastroenteritis). An additional limitation is that if sequelae develop over longer timeframes than our 10-year study (e.g., over decades),[42] our study cannot assess this scenario.

Patient and public involvement statement

Patients were not involved in the development of this protocol, nor were members of the public.

ETHICS AND DISSEMINATION

This study has received approval by a University of Waterloo Research Ethics Committee (#30645), the University of British Columbia Behavioral Research Ethics Board (#H16-00021), and McGill University's Institutional Review Board (#A03-M12-19A). In addition to conference presentations and dissemination to public health practitioners and other knowledge users, results will be published in peer-reviewed journals, and where such publications are not open access,

they will also be stored on UWSpace, the University of Waterloo's Institutional Repository (https://uwspace.uwaterloo.ca).

ACKNOWLEDGEMENTS

We thank Bei Yuan (Ethan) Zhang (University of British Columbia) and Dr. Jenna Dixon (University of Waterloo) for helping draft early versions of this paper, and Dr. Yuen Hung (University of Waterloo) for compiling relevant background literature.

Study Team members Drs. Jonathan Chan (University of British Columbia), Kristine Chapman (Vancouver General Hospital), Susan Horton (University of Waterloo), Douglas Matsell (B.C. Children's Hospital Research Institute), and Sara Pires (Technical University of Denmark), served as expert advisors and critically reviewed the study plan.

AUTHORS' CONTRIBUTIONS

Majowicz and Galanis are the co-Principal Investigators on this study. Majowicz, Galanis, and Taylor conceived the study. The overall design was first developed by Majowicz, Galanis, Taylor, and Panagiotoglou, with critical revisions from Cook, Ethelberg, Leatherdale, Kaplan, and Patrick. All coauthors developed the analysis plan, with specific statistical expertise provided by Cook, Chaurasia, and Gohari, and economic expertise by Panagiotoglou.

Majowicz drafted the manuscript; all authors provided feedback on manuscript drafts and approved the final version to be published. All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

FUNDING

This work is supported by the Canadian Institutes of Health Research (CIHR), Project Grant program, grant number 156385 (PI: Shannon Majowicz), and also received seed funding

from the University of Waterloo's Research Incentive Fund (PI: Shannon Majowicz) and the British Columbia Centre for Disease Control Foundation for Public Health's Open Award Program (PI: Eleni Galanis).

DISCLAIMER

All inferences, opinions, and conclusions drawn in this research protocol are those of the authors, and do not reflect the opinions or policies of the Data Steward(s).

COMPETING INTERESTS

Drs. Majowicz and Galanis report funding for this study as per the funding statement. Dr. Majowicz reports other relationships; she is an Associate Editor at Epidemiology and Infection (for which she receives a small honorarium); she has served as a paid Expert on behalf of the Attorney General of Canada in legal proceedings, providing evidence on the public health risks and benefits of unpasteurized milk; and she is an expert on the Joint FAO/WHO Expert Meetings on Microbiological Risk Assessment (JEMRA) Roster of Experts. Dr. Kaplan reports honoraria for speaking or consultancy from Abbvie, Janssen, Pfizer, and Takeda. He has received research support from Ferring, Janssen, Abbvie, GlaxoSmith Kline, Merck, and Shire. He shares ownership of a patent: TREATMENT OF INFLAMMATORY DISORDERS, AUTOIMMUNE DISEASE, AND PBC. UTI Limited Partnership, assignee. Patent WO2019046959A1. PCT/CA2018/051098. 7 Sept. 2018. Dr. Galanis' spouse works for QHR Technologies, a Canadian medical records company; these records were not used in this study. All other authors have nothing to disclose.

DATA SHARING

Open access for these data is not permitted by the data stewards; further details on the legislation and agreements can be found at: https://www.popdata.bc.ca/dataaccess/rdaf/history and https://www.popdata.bc.ca/dataaccess/rdaf/expectations. To access the data used for this study, researchers must submit a Data Access Request through Population Data B.C. (for all databases except Panorama) and the Panorama Data Governance Committee (Panorama data), who have record of the files and fields provided (cite project: Majowicz Galanis 15-180). We will make the programming code used to clean and analyse the data available (on request, or via publications where possible).

REFERENCES

1. Havelaar AH, Kirk MD, Torgerson PR, et al. World Health Organization global estimates and regional comparisons of the burden of foodborne disease in 2010. *PLoS Med* 2015;12(12):e1001923.
2. Thomas MK, Murray R, Flockhart L, et al. Estimates of the burden of foodborne illness in Canada for 30 specified pathogens and unspecified agents, circa 2006. *Foodborne Pathog Dis.* 2013;10(7):639-48.
3. Awofisayo A, Amar C, Ruggles R, et al. Pregnancy-associated listeriosis in England and Wales. *Epidemiol Infect*. 2015;143(2):249-56.
4. Charlier C, Perrodeau É, Leclercq A, et al. Clinical features and prognostic factors of listeriosis: the MONALISA national prospective cohort study. *Lancet Infect Dis*. 2017;17(5):510-9.
5. Karmali MA, Petric M, Lim C, et al. The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing *Escherichia coli*. *J Infect Dis*. 1985;151(5):775-82.
6. Gould LH, Demma L, Jones TF, et al. Hemolytic uremic syndrome and death in persons with *Escherichia coli* O157: H7 infection, Foodborne Diseases Active Surveillance Network sites, 2000–2006. *Clin Infect Dis*. 2009;49(10):1480-5.
7. Rodríguez LA, Ruigómez A, Panés J. Acute gastroenteritis is followed by an increased risk of inflammatory bowel disease. *Gastroenterol*. 2006;130(6):1588-94.
8. Small CL, Xing L, McPhee J, et al. Acute infectious gastroenteritis potentiates a Crohn's disease pathobiont to fuel ongoing inflammation in the post-infectious period. *PLoS Pathog*. 2016. http://dx.doi.org/10.1371/journal.ppat.1005907.

9. Ruzante JM, Majowicz SE, Fazil A, et al. Hospitalization and deaths for select enteric illnesses and associated sequelae in Canada, 2001-2004. *Epidemiol Infect*. 2011;139(6):937-45.
10. Tam CC, Rodrigues LC, Petersen I, et al. Incidence of Guillain-Barré syndrome among patients with *Campylobacter* infection: a general practice research database study. *J Infect Dis*. 2006;194(1):95-7.
11. Helms M, Vastrup P, Gerner-Smidt P, et al. Short and long term mortality associated with foodborne bacterial infections: registry-based study. *Br Med J*. 2003;326:1-5.
12. Bremell T, Bjelle A, Svedhem A. Rheumatic symptoms following an outbreak of campylobacter enteritis: a five year follow up. *Ann Rheum Dis*. 1991;50(12):934-8.
13. Marshall JK, Thabane M, Garg AX, et al. Incidence and epidemiology of irritable bowel syndrome after a large waterborne outbreak of bacterial dysentery. *Gastroenterol.* 2006;131(2):445-50.
14. Garg AX, Pope JE, Thiessen-Philbrook H, et al. Arthritis risk after acute bacterial gastroenteritis. *Rheumatol*. 2008;47(2):200-4.
15. Hanevik K, Wensaas KA, Rortveit G, et al. Irritable bowel syndrome and chronic fatigue 6 years after giardia infection: a controlled prospective cohort study. *Clin Infect Dis*. 2014;59(10):1394-400.
16. Rehn M, Wallensten A, Widerström M, et al. Post-infection symptoms following two large waterborne outbreaks of *Cryptosporidium hominis* in Northern Sweden, 2010–2011. *BMC Public Health*. 2015;15(1):529.
17. Hannu T, Mattila L, Rautelin H, et al. Campylobacter-triggered reactive arthritis: a population-based study. *Rheumatol*. 2002;41(3):312-8.

18. Hunter PR, Hughes S, Woodhouse S, et al J. Health sequelae of human cryptosporidiosis in immunocompetent patients. *Clin Infect Dis*. 2004;39(4):504-10.

19. Wang LH, Fang XC, Pan GZ. Bacillary dysentery as a causative factor of irritable bowel syndrome and its pathogenesis. *Gut*. 2004;53(8):1096-101.

20. Hannu T, Mattila L, Siitonen A, et al. Reactive arthritis attributable to *Shigella* infection: a clinical and epidemiological nationwide study. *Ann Rheum Dis*. 2005;64(4):594-8.

21. Saps M, Pensabene L, Di Martino L, et al. Post-infectious functional gastrointestinal disorders in children. *J Pediatr*. 2008;152(6):812-6.

22. Rosner BM, Werber D, Höhle M, et al. Clinical aspects and self-reported symptoms of sequelae of *Yersinia enterocolitica* infections in a population-based study, Germany 2009–2010. *BMC Infect Dis*. 2013;13(1):236.

23. Litleskare S, Rortveit G, Eide GE, et al. Prevalence of irritable bowel syndrome and chronic fatigue 10 years after *Giardia* infection. *Clin Gastroenterol Hepatol.* 2018;16(7):1064-72.

24. Helms M, Simonsen J, Mølbak K. Foodborne bacterial infection and hospitalization: a registry-based study. *Clin Infect Dis*. 2006;42(4):498-506.

25. Rodríguez LA, Ruigómez A, Panés J. Acute gastroenteritis is followed by an increased risk of inflammatory bowel disease. *Gastroenterol*. 2006;130(6):1588-94.

26. Ternhag A, Törner A, Svensson Å, et al. Short-and long-term effects of bacterial gastrointestinal infections. *Emerg Infect Dis*. 2008;14(1):143.

27. Gradel KO, Nielsen HL, Schønheyder HC, et al. Increased short-and long-term risk of inflammatory bowel disease after *Salmonella* or *Campylobacter* gastroenteritis. *Gastroenterol*. 2009;137(2):495-501.

28. Jess T, Simonsen J, Nielsen NM, et al. Enteric *Salmonella* or *Campylobacter* infections and the risk of inflammatory bowel disease. *Gut*. 2011;60(3):318-24.
29. Porter CK, Choi D, Cash B, et al. Pathogen-specific risk of chronic gastrointestinal disorders following bacterial causes of foodborne illness. *BMC Gastroenterol*. 2013;13(1):46.
30. Riddle MS, Murray JA, Cash BD, et al. Pathogen-specific risk of celiac disease following bacterial causes of foodborne illness: a retrospective cohort study. *Digestive Diseases and Sciences*. 2013;58(11):3242-5.
31. Nakao JH, Collier SA, Gargano JW. Giardiasis and subsequent irritable bowel syndrome: a longitudinal cohort study using health insurance data. *J Infect Dis*. 2017;215(5):798-805.
32. Painter JE, Collier SA, Gargano JW. Association between *Giardia* and arthritis or joint pain in a large health insurance cohort: could it be reactive arthritis? *Epidemiol Infect.* 2017;145(3):471-7.
33. Kaakoush NO, Castaño-Rodríguez N, Mitchell HM, et al. Global epidemiology of *Campylobacter* infection. *Clin Microbiol Rev*. 2015;28(3):687-720.
34. Tricco AC, Duval B, De Serres G, et al. A review of interventions triggered by hepatitis A infected food-handlers in Canada. *BMC Health Serv Res*. 2006;6(1):157.
35. Carpentier B, Cerf O. Persistence of *Listeria monocytogenes* in food industry equipment and premises. *Int J Food Microbiol*. 2011;145(1):1-8.
36. British Columbia Centre for Disease Control. British Columbia Annual Summary of Reportable Diseases, 2017. http://www.bccdc.ca/resource-gallery/Documents/Statistics%20and%20Research/Statistics%20and%20Reports/Epid/Annual%20Reports/2017CDAnnualReportFinal.pdf (accessed 12 November 2019).

37. Scallan E, Majowicz SE, Hall G, et al. Prevalence of diarrhoea in the community in Australia, Canada, Ireland, and the United States. *Int J Epidemiol*. 2005;34(2):454-60.

38. Government of British Columbia. https://www2.gov.bc.ca/gov/content/health/health-drug-coverage/msp/bc-residents/eligibility-and-enrolment/are-you-eligible (accessed 12 November 2019).

39. Ark TK, Kesselring S, Hills B, et al. Population Data BC: Supporting population data science in British Columbia. *Int J Pop Data Sci.* 2020;4(2):14.

40. Population Data BC. About PopData. https://www.popdata.bc.ca/aboutus (accessed 12 November 2019).

41. Government of British Columbia. Public Health Act [SBC 2008] Chapter 28. http://www.bclaws.ca/civix/document/id/complete/statreg/08028_01 (accessed 12 November 2019).

42. Moorin RE, Heyworth JS, Forbes GM, et al. Long-term health risks for children and young adults after infective gastroenteritis. *Emerg Infect Dis*. 2010;16(9):1440.

43. BC Ministry of Health [creator] (2018): Medical Services Plan (MSP) Payment Information File. Population Data BC [publisher]. Data Extract. MOH (2018). http://www.popdata.bc.ca/data

44. Canadian Institute for Health Information [creator] (2019): Discharge Abstract Database (Hospital Separations). Population Data BC [publisher]. Data Extract. MOH (2018). http://www.popdata.bc.ca/data

45. BC Ministry of Health [creator] (2019): PharmaNet. BC Ministry of Health [publisher]. Data Extract. Data Stewardship Committee (2018). http://www.popdata.bc.ca/data

46. BC Vital Statistics Agency [creator] (2019): Vital Statistics Deaths. Population Data BC [publisher]. Data Extract BC Vital Statistics Agency (2018). http://www.popdata.bc.ca/data

47. BC Vital Statistics Agency [creator] (2019): Vital Statistics Stillbirths. Population Data BC [publisher]. Data Extract. BC Vital Statistics Agency (2018). http://www.popdata.bc.ca/data

48. British Columbia Ministry of Health [creator] (2019): Consolidation File (MSP Registration & Premium Billing). Population Data BC [publisher]. Data Extract. MOH (2018). http://www.popdata.bc.ca/data

49. Statistics Canada [creator] (2009): Statistics Canada Income Band Data. Catalogue Number: 13C0016. Population Data BC [publisher]. Data Extract. Population Data BC (2018). http://www.popdata.bc.ca/data

50. Population Data BC. Data Linkage. https://www.popdata.bc.ca/datalinkage (accessed 12 November 2019).

51. British Columbia Centre for Disease Control. Case Definitions. http://www.bccdc.ca/health-professionals/clinical-resources/case-definitions (accessed 12 November 2019).

52. World Health Organization. ICD-10 Version:2016. https://icd.who.int/browse10/2016/en (accessed 12 November 2019).

53. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. Springer Science & Business Media 2013.

54. Cook RJ, Lawless JF. The Statistical Analysis of Recurrent Events. New York, NY: Springer-Verlag, 2007.

55. Batz MB, Henke E, Kowalcyk B. Long-term consequences of foodborne infections. *Infect Dis Clinics N Am*. 2013;27(3):599-616.

56. Hu W. Diagnostic Codes in MSP Claim Data, Summary Report. Victoria: Medical Services Plan; 1996.

57. Galanis E, Taylor M, Romanowski K, et al. Evaluating the timeliness of enteric disease surveillance in British Columbia, Canada, 2012-13. *Can J Infect Dis Med Microbiol.* 2017 Article ID 9854103, 7 pages, 2017.

58. Klein JP, Moeschberger ML. Survival Analysis: Techniques for Censored and Truncated Data, 2nd ed. New York, NY: Springer-Verlag New York Inc., 2003.

59. Qin J, Shen Y. Statistical methods for analyzing right-censored length-biased data under Cox model. *Biometrics*. 2010;66(2):382–392.

60. von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ*. 2007;335(7624):806-8.

61. Benchimol EI, Smeeth L, Guttman A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Med.* 2015;12(10): e1001885. https://doi.org/10.1371/journal.pmed.1001885

62. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

63. Cox DRJ. Regression models and life-tables (with discussion). *J R Stat Soc, B*. 1972;34:187-220.

64. Charlson ME, Pompeo P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chron Dis*. 1987;40(5):373-83.

65. Charlson M, Szatrowski TP, Peterson J, et al. Validation of a combined comorbidity index. *J Clin Epidemiol*. 1884;47(11):1245-51.

66. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care.* 2005;43(11):1130-9.

67. Hude Quan, Bing Li, Chantal M, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *Am J Epidemiol*. 2011;173(6):676–82.

68. Kaplan EL, Meier PJ. Nonparametric estimation for incomplete observations. *J Am Stat Assoc*. 1958;53:457-81.

69. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc.* 1999;94(446):496-509.

70. Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. *Am J Epidemiol.* 2009;170(2):244-256.

71. Government of British Columbia. Medical Services Commission Payment Schedule. December 1, 2015. https://www2.gov.bc.ca/assets/gov/health/practitioner-pro/medical-services-plan/msc_payment_schedule.pdf (accessed 12 November 2019).

72. Poole B, Robinson S, MacKinnon M. Resource Intensity Weights™ and Canadian hospital costs: some preliminary data. *Healthc Manage Forum*. 1998:11(1):22-26.

73. Canadian Institute for Health Information. Patient Cost Estimator: Methodological Notes and Glossary. Ottawa, ON: CIHI; 2016. https://www.cihi.ca/en/pce_methodology_notes_en.pdf (accessed 26 November 2019).

74. Statistics Canada, Consumer Price Index (CPI). http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=2301 (accessed 12 November 2019).

75. Mansournia MA, Altman DG. Population attributable fraction. *BMJ*. 2018;360:k757.

76. Holmes WM. Using Propensity Scores in Quasi-Experimental Designs. Los Angeles, Calif.: Sage Publishing, 2014.

77. Stürmer T, Wyss R, Glynn RJ, et al. Propensity scores for confounder adjustment when assessing the effects of medical interventions using nonexperimental study designs. *J Int Med*. 2014;275(6):570-80.

78. Brookhart MA, Schneeweiss S, Rothman KJ, et al. Variable selection for propensity score models. *Am J Epidemiol*. 2006;163(12):1149-56.

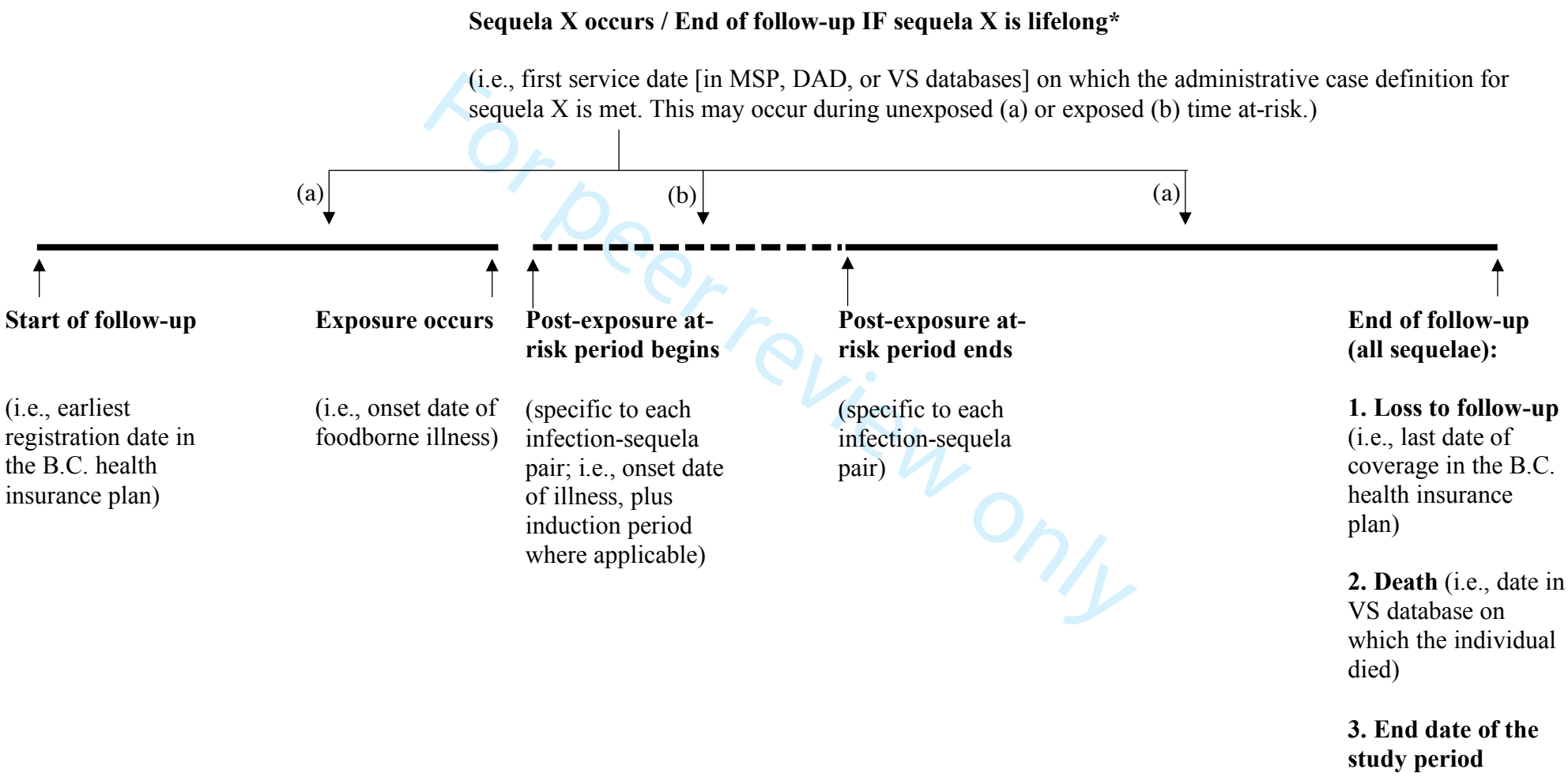
79. Kuenzig ME, Barnabe C, Seow CH, et al. Asthma is associated with subsequent development of inflammatory bowel disease: a population-based case–control study. *Clin Gastroenterol Hepatol*. 2017;15(9):1405-12.

80. MacDougall L, Majowicz S, Doré K, et al. Under-reporting of infectious gastrointestinal illness in British Columbia, Canada: who is counted in provincial communicable disease statistics? *Epidemiol Infect*. 2008;136(2):248-56.

FIGURE LEGEND

Figure 1. **Study follow-up period and time-at-risk for development of Sequela X (solid lines: unexposed time at-risk; dashed lines: exposed time at-risk; B.C., British Columbia, Canada; MSP: Medical Services Plan; DAD: Discharge Abstracts Database; VS: Vital Statistics)**

FIGURE 1.



* For sequelae where recovery and return to being at-risk is possible: this date + recovery time for sequela X = date of return to being at risk for a subsequent occurrence of sequela X.