

BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

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:	<i>BMJ Open</i>
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

SCHOLARONE™
Manuscripts



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.

1
2
3 1 **Determining the long-term health burden and risk of sequelae for 14 foodborne infections,**
4
5 2 **in British Columbia, Canada: protocol for a retrospective population-based cohort study**
6
7
8 3

9
10 4 Shannon E. Majowicz^{1*}, Dimitra Panagiotoglou², Marsha Taylor³, Mahmood R. Gohari¹, Gilaad
11
12 5 G. Kaplan⁴, Ashok Chaurasia¹, Scott T. Leatherdale¹, Richard J. Cook⁵, David M. Patrick^{3,6},
13
14 6 Steen Ethelberg^{7,8}, Eleni Galanis^{3,6}
15
16

17 7 ¹ School of Public Health and Health Systems, University of Waterloo (Waterloo, Canada)
18

19 8 ² Department of Epidemiology, Biostatistics and Occupational Health, McGill University
20
21 (Montreal, Canada)
22
23

24 10 ³ British Columbia Centre for Disease Control (Vancouver, Canada)
25

26 11 ⁴ Departments of Medicine & Community Health Sciences, University of Calgary (Calgary,
27
28 Canada)
29

30
31 13 ⁵ Department of Statistics and Actuarial Sciences, University of Waterloo (Waterloo, Canada)
32

33 14 ⁶ School of Population and Public Health, University of British Columbia (Vancouver, Canada)
34

35 15 ⁷ Department of Infectious Disease Epidemiology and Prevention, Statens Serum Institut
36
37 (Copenhagen, Denmark)
38
39

40 17 ⁸ Global Health Section, Department of Public Health, University of Copenhagen (Copenhagen,
41
42 Denmark)
43
44

45 19
46
47 20 *Corresponding Author:
48 21 200 University Avenue West
49 22 Waterloo, Ontario, Canada, N2L 3G1
50 23 Email: smajowicz@uwaterloo.ca
51 24 Telephone: +1 519 888 4567 extension 31790
52
53

54 26 **KEY WORDS:** foodborne infections, sequelae, cohort study, public health
55
56
57
58
59
60

1
2
3 27 **ABSTRACT**
4

5 28 **Introduction:** Over 1 in 8 Canadians is affected by a foodborne infection annually; however, the
6
7 long-term consequences, including the risks and costs of sequelae, are unclear. We aim to
8 29
9 estimate the health burden and direct costs of 14 infections commonly transmitted by food,
10 30
11 considering the acute illness and subsequent sequelae and mortality, for the population of British
12 31
13 Columbia (B.C.), Canada (~4.7 million).
14 32

15
16
17 33 **Methods and analysis:** We will conduct a population-based retrospective cohort study of the
18
19 34 B.C. provincial population, over a 10-year study period (January 1, 2005 to December 31, 2014).
20
21 Exposure is defined as a provincially-reported illness caused by: botulism, *Campylobacter*,
22 35
23 *Cryptosporidium*, *Cyclospora*, *Giardia*, hepatitis A, *Listeria*, *Salmonella spp.* (non-typhoidal,
24 36
25 Typhi, Paratyphi), Shiga toxin-producing *E. coli*, *Shigella*, *Vibrio parahaemolyticus*, or *Yersinia*
26 37
27 (excluding *pestis*). We will link individual-level longitudinal data from eight province-wide
28 38
29 administrative health and reportable disease databases that include physician visits,
30 39
31 hospitalizations and day surgeries, deaths, stillbirths, prescription medications (except those to
32 40
33 treat HIV), and reportable foodborne diseases. Using these linked databases we will investigate
34 41
35 the likelihood of various sequelae and death. Hazard models will be used to estimate the risk of
36 42
37 outcomes and their association with the type of foodborne infection. Epidemiologic analyses will
38 43
39 be conducted to determine the progression of illness and the fraction of sequelae attributable to
40 44
41 specific foodborne infections. Economic analyses will assess the consequent direct healthcare
42 45
43 costs.
44 46
45
46

47 47 **Ethics and dissemination:** This study has been approved by a University of Waterloo Research
48 48
49 Ethics Committee (#30645), the University of British Columbia Behavioral Research Ethics
50
51 Board (#H16-00021), and McGill University's Institutional Review Board (#A03-M12-19A).
52 49
53
54
55
56
57
58
59
60

1
2
3 50 Results will be disseminated via presentations to academics, public health practitioners, and
4
5 51 other knowledge users, and publication in peer-reviewed journals. Where such publications are
6
7 52 not open access, manuscripts will also be available via the University of Waterloo's Institutional
8
9 53 Repository (<https://uwspace.uwaterloo.ca>).

10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

54 STRENGTHS AND LIMITATIONS

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

67 INTRODUCTION

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

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

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

1
2
3 90 Although infections such as *Salmonella spp.* and STEC can be transmitted via several
4
5 91 routes (e.g., person-to-person, water), their transmission via food and their presence throughout
6
7 92 the food system (e.g., food animals as a reservoir for *Campylobacter spp.*,[33] food handlers
8
9 93 shedding hepatitis A,[34] the ability of *Listeria monocytogenes* to persist in food production
10
11 94 equipment [35]) mean that these infections are often termed “foodborne” although some fraction
12
13 95 will not be transmitted via food directly. Here, we apply the term “foodborne infection” to 14
14
15 96 infections that can be transmitted via food (botulism, *Campylobacter*, *Cryptosporidium*,
16
17 97 *Cyclospora*, *Giardia*, hepatitis A, *Listeria*, *Salmonella spp.* (non-typhoidal, Typhi, Paratyphi),
18
19 98 STEC, *Shigella*, *Vibrio parahaemolyticus*, and *Yersinia* excluding *pestis*), recognizing that not
20
21 99 all result from direct foodborne transmission.
22
23
24
25

26 100 The overall goal of this study is to estimate the health burden and costs of these 14
27
28 101 infections, considering the acute illness and subsequent sequelae and associated mortality, for the
29
30 102 population of British Columbia (B.C.), Canada. Our specific objectives are to:
31
32

- 33 103 1. determine the risk of developing sequelae following infection;
- 34
35 104 2. describe the epidemiology and clinical progression across the range of outcomes, including
36
37 105 acute illness, sequelae, and death;
- 38
39 106 3. quantify the direct healthcare costs due to these infections and their various outcomes; and
40
41 107 4. determine the risk of sequelae in the population attributable to these infections.
42
43
44

45 108

46 109 **METHODS AND ANALYSIS**

47 110 **Study setting**

48
49 111 B.C. is Canada’s westernmost and third most populous province (~4.7 million circa
50
51 112 2014). The annual incidence of foodborne infections in B.C. is comparable to annual incidences
52
53 113 in Canada, the United States, Australia, and Western Europe.[e.g., 2, 36-37]
54
55
56
57
58
59
60

1
2
3 114 All B.C. residents (defined as citizens or permanent residents of Canada who are
4
5 115 physically present in B.C. for at least six months in a calendar year), their dependents, and
6
7 116 certain other individuals (e.g., some holders of study or work permits) are covered by the
8
9 117 province's health insurance program.[38] Enrolment is mandatory, and this program covers
10
11 118 nearly all of the B.C. population (with the exception of members of the Canadian military, Royal
12
13 119 Canadian Mounted Police, and some First Nations individuals covered by federal insurance
14
15 120 programs). Physician visits, laboratory tests, hospitalizations, out-patient hospital services, and
16
17 121 drugs for certain populations are among the publicly-funded benefits. The administrative datasets
18
19 122 that contain these health care use data, along with vital statistics (e.g., births, deaths),
20
21 123 demographic data, and other datasets, are held by the B.C. Ministry of Health and are accessible
22
23 124 to researchers via Population Data B.C., a central repository and "multi-university, data and
24
25 125 education resource" that "support[s] research access to individual-level, de-identified
26
27 126 longitudinal data on British Columbia's 4.7 million residents".[39]
28
29
30
31
32

33 127 The B.C. Public Health Act mandates that reportable diseases, including several
34
35 128 foodborne infections,[40] be reported by health professionals and laboratories to the local and
36
37 129 provincial public health authorities, and these data are managed provincially by the B.C. Centre
38
39 130 for Disease Control (BCCDC). These data are housed within Panorama, the provincial public
40
41 131 health database of reportable diseases. In the Panorama database, as well as the administrative
42
43 132 health and vital statistics databases, individuals are recorded by their unique Personal Health
44
45 133 Number (PHN), allowing information from these data sources to be linked by individual.
46
47
48
49
50

51
52
53
54
55
56
57
58
59
60

135 **Study design, population, and timeframe**

1
2
3 136 This is a retrospective cohort study of the population of B.C., with additional descriptive,
4
5 137 cost, and population attributable risk analyses. The study population is all individuals in B.C.
6
7 138 registered with the provincial health insurance program at any point during the study period, i.e.,
8
9
10 139 all individuals with the following from 2005 to 2014 inclusive: one or more record in one or
11
12 140 more of the Medical Services Plan (MSP), Discharge Abstracts Database (DAD), Vital Statistics
13
14 141 Deaths, or PharmaNet; or record of coverage under the provincial insurance program within the
15
16 142 Consolidation File database (see Table 1). The 10-year study period is January 1, 2005 to
17
18 143 December 31, 2014, inclusive, with additional two-year wash-in (January 1, 2003 to December
19
20 144 31, 2004) and wash-out (January 1, 2015 to December 31, 2016) periods. During these periods
21
22 145 we will identify occurrences of foodborne infections, sequelae, and death. The 10-year study
23
24 146 period was selected to more than encompass timeframes for initial sequelae development and
25
26 147 ensuing healthcare use currently reflected in the literature (i.e., days to years), although there is
27
28 148 some evidence that sequelae can develop over longer timeframes (e.g., over decades).[41]
29
30
31
32

33 149 We assume that enrolment in the provincial health insurance program (i.e., entry into the
34
35 150 study population) and reasons for exit from the cohort (e.g., moving away from B.C.) are not
36
37 151 related to the exposures nor the outcomes of interest.
38
39
40
41

42 153 **Data sources and linkage**

43
44 154 The study will use individually-linked, longitudinal data from eight databases to
45
46 155 investigate both acute and longer-term health outcomes following foodborne infection (Table 1).
47
48 156 In totality, these data contain information on 14 reportable foodborne infections, physician and
49
50 157 hospital visits, prescription medications, vital statistics, and various demographic descriptors, for
51
52
53
54
55
56
57
58
59
60

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

160

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

Database (Reference)	Database description and summary of variables included for this study	Date range
<i>Health Care and Health Services Data</i>		
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
<i>Population and Vital Statistics Data</i>		
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

1 2 3 4 5 6 7 8 9	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
10	Reportable Disease Data		
11 12 13 14 15 16 17	Panorama Public Health Information System	All cases of the following 14 reportable diseases reported in B.C.: botulism, <i>Campylobacter</i> , <i>Cryptosporidium</i> , <i>Cyclospora</i> , <i>Giardia</i> , hepatitis A, <i>Listeria</i> , <i>Salmonella spp.</i> (non-typhoidal, Typhi, Paratyphi), STEC, <i>Shigella</i> , <i>Vibrio parahaemolyticus</i> , and <i>Yersinia</i> . Includes onset date, reported date, health authority, and etiologic agent.	2005/01/01 to 2014/12/31

18 *ICD: International Classification of Diseases
19
20
21
22

163
164
165 Population Data B.C. will link the eight databases directly by individual using PHN, with
166 additional identifiers (e.g., name, age, sex) used to validate linkages and link records
167 probabilistically when PHNs are missing. Population Data B.C.'s detailed linkage process is
168 fully described elsewhere.[42] Once the linkage is complete, each individual is then assigned a
169 unique study identifier. All individually-linked, de-identified databases are provided by
170 Population Data B.C. within their Secure Research Environment, a centralized online platform,
171 accessible via virtual private network within Canada, for accessing and analyzing research data,
172 with security standards that meet Data Steward requirements.
173

174 **Measuring exposure and outcomes**

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

1
2
3 180 and *Yersinia* (excluding *pestis*). Case definitions for each of these infections are specified by the
4
5 181 BCCDC.[43]. Individuals without a reported foodborne infection, but who have International
6
7 182 Classification of Disease (ICD) codes either for one of our infections (e.g., A02.0, ‘*Salmonella*
8
9 183 enteritis’) or for non-specific gastroenteritis (e.g., A08.4, ‘viral intestinal infection, unspecified’
10
11 184 [44]) within the MSP and DAD databases, will be considered potentially exposed. We will
12
13 185 describe these individuals as a separate group in our descriptive, economic, and population
14
15 186 attributable fraction analyses, but will remove them from analyses of sequelae risk.

16
17
18
19 187 It is possible for individuals to have more than one reported foodborne infection during
20
21 188 the study period, either as a simultaneously occurring co-infection, or as two or more distinct
22
23 189 events. For these individuals, we will treat this as a complex exposure problem; sequelae will be
24
25 190 associated with the most plausible explanatory infection, considering biology and timing, and we
26
27 191 will adjust for the presence of concurrent foodborne infections if applicable.[45-46]

28
29
30
31 192 Our primary outcomes of interest are those sequelae for which the link to a given
32
33 193 foodborne infection is either established or is possible (Table 2). We selected sequelae (a) with
34
35 194 evidence of an association with any of the 14 individual foodborne infections,[e.g., 47] and (b)
36
37 195 that occur via direct effects of pathogens or their toxin, or via auto-immune or chronic
38
39 196 inflammatory processes that can be triggered by the infection. We will classify individuals as
40
41 197 having the sequelae via administrative case definitions, using International Classification of
42
43 198 Disease (ICD) diagnostic codes within the MSP, DAD, and VS data, with the exception of
44
45 199 stillbirths which will be determined using recorded events in the VS-Stillbirths database. The
46
47 200 ICD codes in the MSP data are generally considered accurate to the third digit.[48] Although
48
49 201 ICD codes in the VS and DAD data are ordered by most probable diagnosis, we will consider all
50
51 202 codes, regardless of order.
52
53
54
55
56
57
58
59
60

1
2
3 203 Validation of the ICD codes is currently in progress, via a literature review to identify
4
5 204 administrative case definitions that have been validated in the Canadian context, medical expert
6
7
8 205 consultation, and, for those sequelae without a relevant validated definition, a targeted chart
9
10 206 review in B.C. In addition to the ICD codes, we may also use PharmaNet data to improve
11
12 207 sequelae classification if needed, by identifying pharmaceuticals given to patients (e.g., Tumor
13
14 208 Necrosis Factor inhibitor use as an indication of reactive arthritis; intravenous immune globulin
15
16
17 209 use for GBS).

18
19 210 Some of our sequelae of interest are lifelong (e.g., Graves' Disease), and some are
20
21 211 transient in that complete recovery is possible (e.g., GBS, stillbirth). For lifelong sequelae, we
22
23 212 will consider the individual as having the sequela on the earliest date they meet the
24
25 213 administrative case definition for that sequela (with subsequent records considered as a
26
27 214 continuation of the original event). For sequelae from which recovery and subsequent return to
28
29 215 being at-risk is possible, we will consider the individual as first having the sequela on the earliest
30
31 216 date they meet the administrative case definition for that sequela; we will then apply a post-
32
33 217 sequela recovery time to determine the date on which the individual can be considered to be at-
34
35 218 risk for a new, subsequent occurrence of that sequela.

36
37
38 219 Individuals may develop more than one sequela during the study period, either because
39
40 220 they develop multiple different sequelae (e.g., HUS and GBS), or because they develop multiple
41
42 221 occurrences of a single sequela from which complete recovery is possible (e.g., GBS, stillbirth).
43
44 222 In all instances, the occurrence of multiple sequelae will be recorded and described. When
45
46 223 individuals develop multiple different sequelae during the study period (e.g., HUS and GBS), we
47
48 224 will treat these as distinct outcomes in our risk estimates. When individuals develop multiple
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 225 occurrences of the same sequela, we will treat these as distinct outcomes but account for
4
5 226 recurrent events.[46]
6

7
8 227 Individuals with foodborne infections who develop a sequela listed in Table 2, but for
9
10 228 which there is no current evidence of an established or possible link to the specific pathogen
11
12 229 (e.g., *Campylobacter* and stillbirth), will be excluded from our estimates of sequelae risk (but
13
14 230 included in sensitivity analyses).
15

16
17 231 For all 14 infections the secondary outcome of interest is death, which will be classified
18
19 232 using recorded events in the VS-Mortality database. Finally, we are also interested in (a) the
20
21 233 acute illnesses related to these infections (regardless of whether the individual develops sequelae
22
23 234 or dies), and (b) additional outcomes following acute sequelae (e.g., end-stage kidney disease
24
25 235 and kidney transplant following HUS); these will be include only in our descriptive and
26
27 236 economic analyses.
28
29

30
31 237

32
33 238 **Table 2. Established (E) and possible (P) sequelae of foodborne infections, that will be**
34
35 239 **assessed in this study (British Columbia, Canada)**
36

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
<i>Campylobacter</i>	E	P			E	P	P	P			E	
<i>Cryptosporidium</i>											P	
<i>Cyclospora</i>					P						P	
<i>Giardia</i>								P			P	
Hepatitis A	E				P							
<i>Listeria monocytogenes</i>									E ⁵	E		
<i>Salmonella</i> (non-typhoidal)	E				P	P	P	P			E	
<i>Salmonella</i> Paratyphi	E				P	P	P	P			E	
<i>Salmonella</i> Typhi	E				P	P	P	P			E	
STEC	E					E					P	P
<i>Shigella</i>	E					E		P			E	P
<i>Yersinia</i> (excluding <i>pestis</i>)	E		E	P				P			E	

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

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

243 ³ considered here as a sequela of maternal *Listeria* infection

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

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

246

247

248 **Measuring time at-risk**

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

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

1
2
3 264 infection was reported minus the number of days between onset to reporting (e.g., estimated
4
5 265 using the Panorama data or from the literature).[49]
6

7
8 266 Exposed time at-risk will be measured starting from the infection onset date, plus any
9
10 267 additional induction periods (specific to each sequela and currently being determined via
11
12 268 literature review and medical expert consultation). The end of the exposed time at-risk period is
13
14 269 currently being determined via literature review and medical expert consultation. During the
15
16 270 exposed time-at-risk, individuals will be classified as having a sequela specific to their infection
17
18 271 (Table 2) on the date within the 'at-risk' period on which they meet the administrative case
19
20 272 definition for that sequela (e.g., the date of the physician visit or hospitalization). After the post-
21
22 273 infection at-risk period ends, individuals will revert to contributing to the unexposed time-at-risk.
23
24

25
26 274 These data are subject to censoring and truncation. Individuals will be censored for the
27
28 275 sequela in the event of: death, last date of coverage in the provincial health insurance plan (i.e.,
29
30 276 loss to follow up; calculated as above), or the end of the study period, whichever comes first.[50-
31
32 277 51] In our descriptive and economic analyses, we will include all related health care use and
33
34 278 prescription medication costs over the course of the infection and sequela(e), and in our
35
36 279 estimates of mortality we will include any deaths recorded during the study period, following the
37
38 280 sequela.
39
40

41
42 281
43

44 282 **Analysis plan**

45
46
47 283 Data will be analyzed and results reported following the STROBE and RECORD
48
49 284 guidelines.[52-53]. Analyses will be performed using SAS 9.4 (SAS Institute Inc., Cary, NC,
50
51 285 USA) and R.[54] The nature and extent of missing data will be described. If imputation is used
52
53 286 to complete missing data, specific methods and assumptions will be reported. We will emphasize
54
55
56
57
58
59

1
2
3 287 estimation over tests of statistical significance by reporting relative measures of effect along with
4
5 288 associated 95% confidence intervals.
6

7
8 289 The datasets in Table 1 contain the variable “sex” (identified via government records),
9
10 290 that denotes whether individuals are ‘male’ or ‘female’, thereby capturing a composite of sex and
11
12 291 gender. To reflect for potential sex- and gender-differences, we will report and interpret findings
13
14 292 stratified by this variable, in addition to overall findings.
15

16
17 293 *Objective 1:* To determine the risk of developing sequelae following foodborne infection,
18
19 294 we will estimate hazard ratios using Cox regression models,[55] that adjust for confounders and
20
21 295 comorbidities (see below), along with the possible effect modifying role of age, sex,
22
23 296 comorbidity, and medication use (e.g., antibiotics). We will include comorbidities as a composite
24
25 297 score, using the revised Charlson comorbidity index and its associated coding algorithms.[56-59]
26
27 298 Following foodborne infection, we will compare the cumulative risk of first diagnosis of each
28
29 299 infection-specific sequela using life-table and Kaplan-Meier approaches.[55, 60] In the event an
30
31 300 individual dies, we will use competing risk analysis.[61-62] For those who experience more than
32
33 301 one foodborne infection across the study period, we will explore the impacts of having multiple
34
35 302 infections on the risk of sequelae and mortality. We will also estimate the likelihood of dying
36
37 303 following foodborne infection using the same methods described above.
38
39
40
41

42 304 *Objective 2:* To describe the epidemiology and clinical progression across the range of
43
44 305 outcomes, including acute illness (typically diarrhea and other gastrointestinal symptoms),
45
46 306 sequelae, and death, for each foodborne infection we will calculate incidence rates, demographic,
47
48 307 geographic, and temporal distributions, timing and progression of outcomes, and case fatality
49
50 308 rates, for both the acute stage, and sequelae associated to the foodborne infection.
51
52
53
54
55
56
57
58
59
60

1
2
3 309 *Objective 3:* To quantify the direct healthcare costs due to these infections and their
4
5 310 various outcomes, we will determine health service use (i.e., patterns of use by type, frequency,
6
7 311 timing of physician visits and hospitalizations), for both the acute foodborne infection and any
8
9 312 sequelae. We will estimate direct healthcare costs of out-of-hospital physician visits,
10
11 313 hospitalizations, and prescription medications. Out-of-hospital costs will be estimated using the
12
13 314 MSP variables ‘Fee Item’ and ‘Paid Service’ and fee rates from the B.C. fee schedule.[63] Costs
14
15 315 of in-patient and day-case hospitalizations will be calculated using established case-mix
16
17 316 methodology (i.e., using the ‘Resource Intensity Weight’ of each hospitalization),[64] and the
18
19 317 B.C. Ministry of Health unit costs for hospital stays.[65] Total prescription medication costs will
20
21 318 be calculated using the drug cost claimed by the pharmacist which includes the ingredient cost,
22
23 319 professional dispensing fee and other special service fees (if applicable). Because these costs are
24
25 320 captured directly in the PharmaNet data, they will be tallied directly. We will also apply these
26
27 321 methods to determine the direct costs per sequelae, regardless of exposure. Costs will be adjusted
28
29 322 for inflation using the Canadian Consumer Price Index.[66] Results will be reported to allow
30
31 323 comparability with other estimates (e.g., 2010/2011 Canadian and US dollars).

32 324 *Objective 4:* To determine the risk of sequelae in the population attributable to foodborne
33
34 325 infections, we will calculate the proportion of cases of each sequela attributable to the specific
35
36 326 foodborne infections. The total number of cases of each sequela occurring in B.C. during the
37
38 327 study period will be the denominator (e.g., total number of cases of acute kidney injury), and the
39
40 328 numerators will be the numbers of cases of each sequela occurring in those with specific
41
42 329 foodborne infections (e.g., total number of cases of hepatitis A, and of STEC, with acute kidney
43
44 330 injury). We will also describe the fraction of individuals with sequela who do not have a
45
46 331 foodborne infection, but who do have an ICD code for prior gastroenteritis, and use this to
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 332 estimate the additional fractions of sequelae that may have an unidentified foodborne infection
4
5 333 cause. We will calculate fractions for both established and possible sequelae, but clearly
6
7
8 334 distinguish between the two when reporting findings.
9

10 335

11 12 336 **Potential confounders and their adjustment**

13
14
15 337 We will use propensity score matching, and inclusion of potential confounders as
16
17 338 covariates in our analyses, as our primary methods to adjust for confounding.[67] The databases
18
19 339 in Table 1 include direct measurements of important known, strong confounders (e.g., age, sex),
20
21 340 as well as other potential factors (e.g., use of protein pump inhibitors and antibiotics, disease
22
23 341 severity). We will consider the following variables as potential confounders: age, sex, local
24
25 342 health area, income band/area income (as a proxy for socioeconomic status), month/year,
26
27 343 seasonality, immune status (e.g., indication of immunosuppressant drugs, presence of conditions
28
29 344 like cancer, pregnancy), use of medications like antibiotics, and Charlson comorbidity index.
30
31 345 Because none of our data sources include ethnicity nor race, we are unable to adjust, or conduct
32
33 346 sub-analyses, for these factors. Although ethnicity may impact health care seeking behaviours,
34
35 347 we anticipate these impacts will apply equally regardless of exposure, and thus we expect any
36
37 348 bias in our findings to be negligible. Nevertheless, to assess residual confounding, we will
38
39 349 conduct sensitivity analyses,[68-69] and perform indirect adjustments.[70].
40
41
42
43
44

45 350

46 47 351 **Planned sensitivity analyses and study limitations**

48
49 352 We are planning several sensitivity analyses to explore assumptions, methodological
50
51 353 decisions, limitations in the data, and robustness of results. We will explore the impact of
52
53 354 propensity score matching on our sequelae and mortality risk estimates by also using (a) the
54
55
56
57
58
59

1
2
3 355 whole unexposed population, and (b) a random sample of unexposed individuals (matched on
4
5 356 time), instead of propensity score-matched individuals. We may also explore additional matching
6
7
8 357 and control strategies (e.g., matching on age and sex). We will also explore the impacts of
9
10 358 including individuals with foodborne infections who develop a sequela for which there is no
11
12 359 current evidence of an established or possible link to the specific pathogen (e.g., *Campylobacter*
13
14 360 and stillbirth), in our estimates of sequelae risk. Our primary analyses will use a composite of all
15
16 361 infections (i.e., a report of any of the 14 foodborne infections) and their various sequelae. We
17
18 362 will also analyze and present results for each of the individual foodborne infections, and for each
19
20 363 of the sequela.
21
22
23

24 364 A main recognized limitation of reportable disease data, such as the Panorama data in this
25
26 365 study, is the under-ascertainment of foodborne infections. Here, this limitation means that
27
28 366 individuals with foodborne infections who do not seek care nor get tested will be misclassified as
29
30 367 unexposed. We will assess the impacts of such potential misclassification via sensitivity analyses
31
32 368 that illustrate how our findings could be impacted by different misclassification rates, using
33
34 369 estimates of misclassification from the literature,[e.g., 2, 71] and from our data (e.g., individuals
35
36 370 with non-specific gastroenteritis). An additional limitation is that if sequelae develop over
37
38 371 longer timeframes than our 10-year study (e.g., over decades),[41] our study cannot assess this
39
40 372 scenario.
41
42
43
44
45 373

46 374 **Patient and public involvement statement**

47 375 Patients were not involved in the development of this protocol, nor were members of the
48
49 376 public.
50
51
52
53
54 377

1
2
3 378 **ETHICS AND DISSEMINATION**
4

5 379 This study has received approval by a University of Waterloo Research Ethics Committee
6
7 380 (#30645), the University of British Columbia Behavioral Research Ethics Board (#H16-00021),
8
9 381 and McGill University's Institutional Review Board (#A03-M12-19A). In addition to conference
10
11 382 presentations and dissemination to public health practitioners and other knowledge users, results
12
13 383 will be published in peer-reviewed journals, and where such publications are not open access,
14
15 384 they will also be stored on UWSpace, the University of Waterloo's Institutional Repository
16
17 385 (<https://uwspace.uwaterloo.ca>).
18
19
20
21
22 386
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

387 **ACKNOWLEDGEMENTS**

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

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

396 **AUTHORS' CONTRIBUTIONS**

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

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

407 **FUNDING**

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

1
2
3 410 from the University of Waterloo's Research Incentive Fund (PI: Shannon Majowicz) and the
4
5 411 British Columbia Centre for Disease Control Foundation for Public Health's Open Award
6
7 412 Program (PI: Eleni Galanis).
8
9

10 413

11 12 414 **DISCLAIMER**

13
14
15 415 All inferences, opinions, and conclusions drawn in this research protocol are those of the
16
17 416 authors, and do not reflect the opinions or policies of the Data Steward(s).
18
19 417

20 21 418 **COMPETING INTERESTS**

22
23
24 419 Drs. Majowicz and Galanis report funding for this study as per the funding statement.

25
26 420 Dr. Majowicz reports other relationships; she is an Associate Editor at Epidemiology and
27
28 421 Infection (for which she receives a small honorarium); she has served as a paid Expert on behalf
29
30 422 of the Attorney General of Canada in legal proceedings, providing evidence on the public health
31
32 423 risks and benefits of unpasteurized milk; and she is an expert on the Joint FAO/WHO Expert
33
34 424 Meetings on Microbiological Risk Assessment (JEMRA) Roster of Experts. Dr. Kaplan reports
35
36 425 honoraria for speaking or consultancy from Abbvie, Janssen, Pfizer, and Takeda. He has
37
38 426 received research support from Ferring, Janssen, Abbvie, GlaxoSmith Kline, Merck, and Shire.
39
40 427 He shares ownership of a patent: TREATMENT OF INFLAMMATORY DISORDERS,
41
42 428 AUTOIMMUNE DISEASE, AND PBC. UTI Limited Partnership, assignee. Patent
43
44
45 429 WO2019046959A1. PCT/CA2018/051098. 7 Sept. 2018. Dr. Galanis' spouse works for QHR
46
47 430 Technologies, a Canadian medical records company; these records were not used in this study.
48
49
50 431 All other authors have nothing to disclose.
51
52
53
54 432

1
2
3 433 **DATA SHARING**
4

5 434 Open access for these data is not permitted by the data stewards; further details on the
6
7 435 legislation and agreements can be found at: <https://www.popdata.bc.ca/dataaccess/rdaf/history>
8
9
10 436 and <https://www.popdata.bc.ca/dataaccess/rdaf/expectations>. To access the data used for this
11
12 437 study, researchers must submit a Data Access Request through Population Data B.C. (for all
13
14 438 databases except Panorama) and the Panorama Data Governance Committee (Panorama data),
15
16 439 who have record of the files and fields provided (cite project: Majowicz Galanis 15-180). We
17
18 440 will make the programming code used to clean and analyse the data available (on request, or via
19
20 441 publications where possible).
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

442 **REFERENCES**

- 443 1. Havelaar AH, Kirk MD, Torgerson PR, et al. World Health Organization global estimates
444 and regional comparisons of the burden of foodborne disease in 2010. *PLoS Med*
445 2015;12(12):e1001923.
- 446 2. Thomas MK, Murray R, Flockhart L, et al. Estimates of the burden of foodborne illness in
447 Canada for 30 specified pathogens and unspecified agents, circa 2006. *Foodborne Pathog*
448 *Dis.* 2013;10(7):639-48.
- 449 3. Awofisayo A, Amar C, Ruggles R, et al. Pregnancy-associated listeriosis in England and
450 Wales. *Epidemiol Infect.* 2015;143(2):249-56.
- 451 4. Charlier C, Perrodeau É, Leclercq A, et al. Clinical features and prognostic factors of
452 listeriosis: the MONALISA national prospective cohort study. *Lancet Infect Dis.*
453 2017;17(5):510-9.
- 454 5. Karmali MA, Petric M, Lim C, et al. The association between idiopathic hemolytic uremic
455 syndrome and infection by verotoxin-producing *Escherichia coli*. *J Infect Dis.*
456 1985;151(5):775-82.
- 457 6. Gould LH, Demma L, Jones TF, et al. Hemolytic uremic syndrome and death in persons with
458 *Escherichia coli* O157: H7 infection, Foodborne Diseases Active Surveillance Network sites,
459 2000–2006. *Clin Infect Dis.* 2009;49(10):1480-5.
- 460 7. Rodríguez LA, Ruigómez A, Panés J. Acute gastroenteritis is followed by an increased risk
461 of inflammatory bowel disease. *Gastroenterol.* 2006;130(6):1588-94.
- 462 8. Small CL, Xing L, McPhee J, et al. Acute infectious gastroenteritis potentiates a Crohn's
463 disease pathobiont to fuel ongoing inflammation in the post-infectious period. *PLoS Pathog.*
464 2016. <http://dx.doi.org/10.1371/journal.ppat.1005907>.

- 1
2
3 465 9. Ruzante JM, Majowicz SE, Fazil A, et al. Hospitalization and deaths for select enteric
4
5 466 illnesses and associated sequelae in Canada, 2001-2004. *Epidemiol Infect.* 2011;139(6):937-
6
7 467 45.
- 8
9
10 468 10. Tam CC, Rodrigues LC, Petersen I, et al. Incidence of Guillain-Barré syndrome among
11
12 469 patients with *Campylobacter* infection: a general practice research database study. *J Infect*
13
14 470 *Dis.* 2006;194(1):95-7.
- 15
16
17 471 11. Helms M, Vastrup P, Gerner-Smidt P, et al. Short and long term mortality associated with
18
19 472 foodborne bacterial infections: registry-based study. *Br Med J.* 2003;326:1-5.
- 20
21
22 473 12. Bremell T, Bjelle A, Svedhem A. Rheumatic symptoms following an outbreak of
23
24 474 campylobacter enteritis: a five year follow up. *Ann Rheum Dis.* 1991;50(12):934-8.
- 25
26 475 13. Marshall JK, Thabane M, Garg AX, et al. Incidence and epidemiology of irritable bowel
27
28 476 syndrome after a large waterborne outbreak of bacterial dysentery. *Gastroenterol.*
29
30 477 2006;131(2):445-50.
- 31
32
33 478 14. Garg AX, Pope JE, Thiessen-Philbrook H, et al. Arthritis risk after acute bacterial
34
35 479 gastroenteritis. *Rheumatol.* 2008;47(2):200-4.
- 36
37
38 480 15. Hanevik K, Wensaas KA, Rortveit G, et al. Irritable bowel syndrome and chronic fatigue 6
39
40 481 years after giardia infection: a controlled prospective cohort study. *Clin Infect Dis.*
41
42 482 2014;59(10):1394-400.
- 43
44
45 483 16. Rehn M, Wallensten A, Widerström M, et al. Post-infection symptoms following two large
46
47 484 waterborne outbreaks of *Cryptosporidium hominis* in Northern Sweden, 2010–2011. *BMC*
48
49 485 *Public Health.* 2015;15(1):529.
- 50
51
52 486 17. Hannu T, Mattila L, Rautelin H, et al. Campylobacter-triggered reactive arthritis: a
53
54 487 population-based study. *Rheumatol.* 2002;41(3):312-8.
- 55
56
57
58
59
60

- 1
2
3 488 18. Hunter PR, Hughes S, Woodhouse S, et al J. Health sequelae of human cryptosporidiosis in
4
5 489 immunocompetent patients. *Clin Infect Dis*. 2004;39(4):504-10.
6
7
8 490 19. Wang LH, Fang XC, Pan GZ. Bacillary dysentery as a causative factor of irritable bowel
9
10 491 syndrome and its pathogenesis. *Gut*. 2004;53(8):1096-101.
11
12 492 20. Hannu T, Mattila L, Siitonen A, et al. Reactive arthritis attributable to *Shigella* infection: a
13
14 493 clinical and epidemiological nationwide study. *Ann Rheum Dis*. 2005;64(4):594-8.
15
16
17 494 21. Saps M, Pensabene L, Di Martino L, et al. Post-infectious functional gastrointestinal
18
19 495 disorders in children. *J Pediatr*. 2008;152(6):812-6.
20
21
22 496 22. Rosner BM, Werber D, Höhle M, et al. Clinical aspects and self-reported symptoms of
23
24 497 sequelae of *Yersinia enterocolitica* infections in a population-based study, Germany 2009–
25
26 498 2010. *BMC Infect Dis*. 2013;13(1):236.
27
28
29 499 23. Litleskare S, Rortveit G, Eide GE, et al. Prevalence of irritable bowel syndrome and chronic
30
31 500 fatigue 10 years after *Giardia* infection. *Clin Gastroenterol Hepatol*. 2018;16(7):1064-72.
32
33 501 24. Helms M, Simonsen J, Mølbak K. Foodborne bacterial infection and hospitalization: a
34
35 502 registry-based study. *Clin Infect Dis*. 2006;42(4):498-506.
36
37
38 503 25. Rodríguez LA, Ruigómez A, Panés J. Acute gastroenteritis is followed by an increased risk
39
40 504 of inflammatory bowel disease. *Gastroenterol*. 2006;130(6):1588-94.
41
42
43 505 26. Ternhag A, Törner A, Svensson Å, et al. Short-and long-term effects of bacterial
44
45 506 gastrointestinal infections. *Emerg Infect Dis*. 2008;14(1):143.
46
47 507 27. Gradel KO, Nielsen HL, Schönheyder HC, et al. Increased short-and long-term risk of
48
49 508 inflammatory bowel disease after *Salmonella* or *Campylobacter* gastroenteritis.
50
51 509 *Gastroenterol*. 2009;137(2):495-501.
52
53
54
55
56
57
58
59
60

- 1
2
3 510 28. Jess T, Simonsen J, Nielsen NM, et al. Enteric *Salmonella* or *Campylobacter* infections and
4
5 511 the risk of inflammatory bowel disease. *Gut*. 2011;60(3):318-24.
6
7
8 512 29. Porter CK, Choi D, Cash B, et al. Pathogen-specific risk of chronic gastrointestinal disorders
9
10 513 following bacterial causes of foodborne illness. *BMC Gastroenterol*. 2013;13(1):46.
11
12 514 30. Riddle MS, Murray JA, Cash BD, et al. Pathogen-specific risk of celiac disease following
13
14 515 bacterial causes of foodborne illness: a retrospective cohort study. *Digestive Diseases and*
15
16 516 *Sciences*. 2013;58(11):3242-5.
17
18
19 517 31. Nakao JH, Collier SA, Gargano JW. Giardiasis and subsequent irritable bowel syndrome: a
20
21 518 longitudinal cohort study using health insurance data. *J Infect Dis*. 2017;215(5):798-805.
22
23
24 519 32. Painter JE, Collier SA, Gargano JW. Association between *Giardia* and arthritis or joint pain
25
26 520 in a large health insurance cohort: could it be reactive arthritis? *Epidemiol Infect*.
27
28 521 2017;145(3):471-7.
29
30
31 522 33. Kaakoush NO, Castaño-Rodríguez N, Mitchell HM, et al. Global epidemiology of
32
33 523 *Campylobacter* infection. *Clin Microbiol Rev*. 2015;28(3):687-720.
34
35 524 34. Tricco AC, Duval B, De Serres G, et al. A review of interventions triggered by hepatitis A
36
37 525 infected food-handlers in Canada. *BMC Health Serv Res*. 2006;6(1):157.
38
39
40 526 35. Carpentier B, Cerf O. Persistence of *Listeria monocytogenes* in food industry equipment and
41
42 527 premises. *Int J Food Microbiol*. 2011;145(1):1-8.
43
44
45 528 36. British Columbia Centre for Disease Control. British Columbia Annual Summary of
46
47 529 Reportable Diseases, 2017. [http://www.bccdc.ca/resource-](http://www.bccdc.ca/resource-gallery/Documents/Statistics%20and%20Research/Statistics%20and%20Reports/Epid/Annual%20Reports/2017CDAnnualReportFinal.pdf)
48
49 530 [gallery/Documents/Statistics%20and%20Research/Statistics%20and%20Reports/Epid/Annual](http://www.bccdc.ca/resource-gallery/Documents/Statistics%20and%20Research/Statistics%20and%20Reports/Epid/Annual%20Reports/2017CDAnnualReportFinal.pdf)
50
51 531 [l%20Reports/2017CDAnnualReportFinal.pdf](http://www.bccdc.ca/resource-gallery/Documents/Statistics%20and%20Research/Statistics%20and%20Reports/Epid/Annual%20Reports/2017CDAnnualReportFinal.pdf) (accessed 12 November 2019).
52
53
54
55
56
57
58
59
60

- 1
2
3 532 37. Scallan E, Majowicz SE, Hall G, et al. Prevalence of diarrhoea in the community in
4
5 533 Australia, Canada, Ireland, and the United States. *Int J Epidemiol.* 2005;34(2):454-60.
6
7
8 534 38. Government of British Columbia. [https://www2.gov.bc.ca/gov/content/health/health-drug-](https://www2.gov.bc.ca/gov/content/health/health-drug-coverage/msp/bc-residents/eligibility-and-enrolment/are-you-eligible)
9
10 535 [coverage/msp/bc-residents/eligibility-and-enrolment/are-you-eligible](https://www2.gov.bc.ca/gov/content/health/health-drug-coverage/msp/bc-residents/eligibility-and-enrolment/are-you-eligible) (accessed 12 November
11
12 536 2019).
13
14
15 537 39. Population Data BC. About PopData. <https://www.popdata.bc.ca/aboutus> (accessed 12
16
17 538 November 2019).
18
19 539 40. Government of British Columbia. Public Health Act [SBC 2008] Chapter 28.
20
21 540 http://www.bclaws.ca/civix/document/id/complete/statreg/08028_01 (accessed 12 November
22
23 541 2019).
24
25
26 542 41. Moorin RE, Heyworth JS, Forbes GM, et al. Long-term health risks for children and young
27
28 543 adults after infective gastroenteritis. *Emerg Infect Dis.* 2010;16(9):1440.
29
30
31 544 42. Population Data BC. Data Linkage. <https://www.popdata.bc.ca/datalinkage> (accessed 12
32
33 545 November 2019).
34
35 546 43. British Columbia Centre for Disease Control. Case Definitions. [http://www.bccdc.ca/health-](http://www.bccdc.ca/health-professionals/clinical-resources/case-definitions)
36
37 547 [professionals/clinical-resources/case-definitions](http://www.bccdc.ca/health-professionals/clinical-resources/case-definitions) (accessed 12 November 2019).
38
39
40 548 44. World Health Organization. ICD-10 Version:2016. <https://icd.who.int/browse10/2016/en>
41
42 549 (accessed 12 November 2019).
43
44
45 550 45. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. Springer
46
47 551 Science & Business Media 2013.
48
49 552 46. Cook RJ, Lawless JF. The Statistical Analysis of Recurrent Events. New York, NY:
50
51 553 Springer-Verlag, 2007.
52
53
54
55
56
57
58
59
60

- 1
2
3 554 47. Batz MB, Henke E, Kowalcyk B. Long-term consequences of foodborne infections. *Infect*
4
5 555 *Dis Clinics N Am.* 2013;27(3):599-616.
6
7
8 556 48. Hu W. Diagnostic Codes in MSP Claim Data, Summary Report. Victoria: Medical Services
9
10 557 Plan; 1996.
11
12 558 49. Galanis E, Taylor M, Romanowski K, et al. Evaluating the timeliness of enteric disease
13
14 559 surveillance in British Columbia, Canada, 2012-13. *Can J Infect Dis Med Microbiol.* 2017
15
16 560 Article ID 9854103, 7 pages, 2017.
17
18
19 561 50. Klein JP, Moeschberger ML. Survival Analysis: Techniques for Censored and Truncated
20
21 562 Data, 2nd ed. New York, NY: Springer-Verlag New York Inc., 2003.
22
23
24 563 51. Qin J, Shen Y. Statistical methods for analyzing right-censored length-biased data under Cox
25
26 564 model. *Biometrics.* 2010;66(2):382-392.
27
28
29 565 52. von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational
30
31 566 Studies in Epidemiology (STROBE) statement: guidelines for reporting observational
32
33 567 studies. *BMJ.* 2007;335(7624):806-8.
34
35
36 568 53. Benchimol EI, Smeeth L, Guttman A, et al. The REporting of studies Conducted using
37
38 569 Observational Routinely-collected health Data (RECORD) Statement. *PLoS Med.* 2015;12(10):
39
40 570 e1001885. <https://doi.org/10.1371/journal.pmed.1001885>
41
42 571 54. R Core Team (2019). R: A language and environment for statistical computing. R
43
44 572 Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
45
46
47 573 55. Cox DRJ. Regression models and life-tables (with discussion). *J R Stat Soc, B.* 1972;34:187-
48
49 574 220.
50
51 575 56. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic
52
53 576 comorbidity in longitudinal studies: development and validation. *J Chron Dis.*
54
55 577 1987;40(5):373-83.
56
57
58
59
60

- 1
2
3 578 57. Charlson M, Szatrowski TP, Peterson J, et al. Validation of a combined comorbidity index. *J*
4
5 579 *Clin Epidemiol.* 1984;47(11):1245-51.
6
7
8 580 58. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in
9
10 581 ICD-9-CM and ICD-10 administrative data. *Med Care.* 2005;43(11):1130-9.
11
12 582 59. Hude Quan, Bing Li, Chantal M, et al. Updating and validating the Charlson comorbidity
13
14 583 index and score for risk adjustment in hospital discharge abstracts using data from 6
15
16 584 countries. *Am J Epidemiol.* 2011;173(6):676–82.
17
18
19 585 60. Kaplan EL, Meier PJ. Nonparametric estimation for incomplete observations. *J Am Stat*
20
21 586 *Assoc.* 1958;53:457-81.
22
23
24 587 61. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J*
25
26 588 *Am Stat Assoc.* 1999;94(446):496-509.
27
28
29 589 62. Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. *Am J*
30
31 590 *Epidemiol.* 2009;170(2):244-256.
32
33 591 63. Government of British Columbia. Medical Services Commission Payment Schedule.
34
35 592 December 1, 2015. [https://www2.gov.bc.ca/assets/gov/health/practitioner-pro/medical-](https://www2.gov.bc.ca/assets/gov/health/practitioner-pro/medical-services-plan/msc_payment_schedule.pdf)
36
37 593 [services-plan/msc_payment_schedule.pdf](https://www2.gov.bc.ca/assets/gov/health/practitioner-pro/medical-services-plan/msc_payment_schedule.pdf) (accessed 12 November 2019).
38
39
40 594 64. Poole B, Robinson S, MacKinnon M. Resource Intensity Weights™ and Canadian hospital
41
42 595 costs: some preliminary data. *Health Manage Forum.* 1998;11(1):22-26.
43
44
45 596 65. Canadian Institute for Health Information. Patient Cost Estimator: Methodological Notes and
46
47 597 Glossary. Ottawa, ON: CIHI; 2016. https://www.cihi.ca/en/pce_methodology_notes_en.pdf
48
49 598 (accessed 26 November 2019).
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 599 66. Statistics Canada, Consumer Price Index (CPI).
4
5 600 <http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=2301> (accessed 12
6
7 601 November 2019).
8
9
10 602 67. Holmes WM. Using Propensity Scores in Quasi-Experimental Designs. Los Angeles, Calif.:
11
12 603 Sage Publishing, 2014.
13
14 604 68. Stürmer T, Wyss R, Glynn RJ, et al. Propensity scores for confounder adjustment when
15
16 605 assessing the effects of medical interventions using nonexperimental study designs. *J Int*
17
18 606 *Med.* 2014;275(6):570-80.
19
20
21 607 69. Brookhart MA, Schneeweiss S, Rothman KJ, et al. Variable selection for propensity score
22
23 608 models. *Am J Epidemiol.* 2006;163(12):1149-56.
24
25
26 609 70. Kuenzig ME, Barnabe C, Seow CH, et al. Asthma is associated with subsequent development
27
28 610 of inflammatory bowel disease: a population-based case-control study. *Clin Gastroenterol*
29
30 611 *Hepatol.* 2017;15(9):1405-12.
31
32
33 612 71. MacDougall L, Majowicz S, Doré K, et al. Under-reporting of infectious gastrointestinal
34
35 613 illness in British Columbia, Canada: who is counted in provincial communicable disease
36
37 614 statistics? *Epidemiol Infect.* 2008;136(2):248-56.
38
39
40 615 72. BC Ministry of Health [creator] (2018): Medical Services Plan (MSP) Payment Information
41
42 616 File. Population Data BC [publisher]. Data Extract. MOH (2018).
43
44 617 <http://www.popdata.bc.ca/data>
45
46
47 618 73. Canadian Institute for Health Information [creator] (2019): Discharge Abstract Database
48
49 619 (Hospital Separations). Population Data BC [publisher]. Data Extract. MOH
50
51 620 (2018). <http://www.popdata.bc.ca/data>
52
53
54
55
56
57
58
59
60

- 1
2
3 621 74. BC Ministry of Health [creator] (2019): PharmaNet. BC Ministry of Health [publisher]. Data
4
5 622 Extract. Data Stewardship Committee (2018). <http://www.popdata.bc.ca/data>
6
7 623 75. BC Vital Statistics Agency [creator] (2019): Vital Statistics Deaths. Population Data BC
8
9 624 [publisher]. Data Extract BC Vital Statistics Agency (2018). <http://www.popdata.bc.ca/data>
10
11 625 76. BC Vital Statistics Agency [creator] (2019): Vital Statistics Stillbirths. Population Data BC
12
13 626 [publisher]. Data Extract. BC Vital Statistics Agency (2018). <http://www.popdata.bc.ca/data>
14
15 627 77. British Columbia Ministry of Health [creator] (2019): Consolidation File (MSP Registration
16
17 628 & Premium Billing). Population Data BC [publisher]. Data Extract. MOH
18
19 629 (2018). <http://www.popdata.bc.ca/data>
20
21
22 630 78. Statistics Canada [creator] (2009): Statistics Canada Income Band Data. Catalogue Number:
23
24 631 13C0016. Population Data BC [publisher]. Data Extract. Population Data BC (2018).
25
26 632 <http://www.popdata.bc.ca/data>
27
28
29
30
31 633
32
33 634
34
35 635
36
37
38 636
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

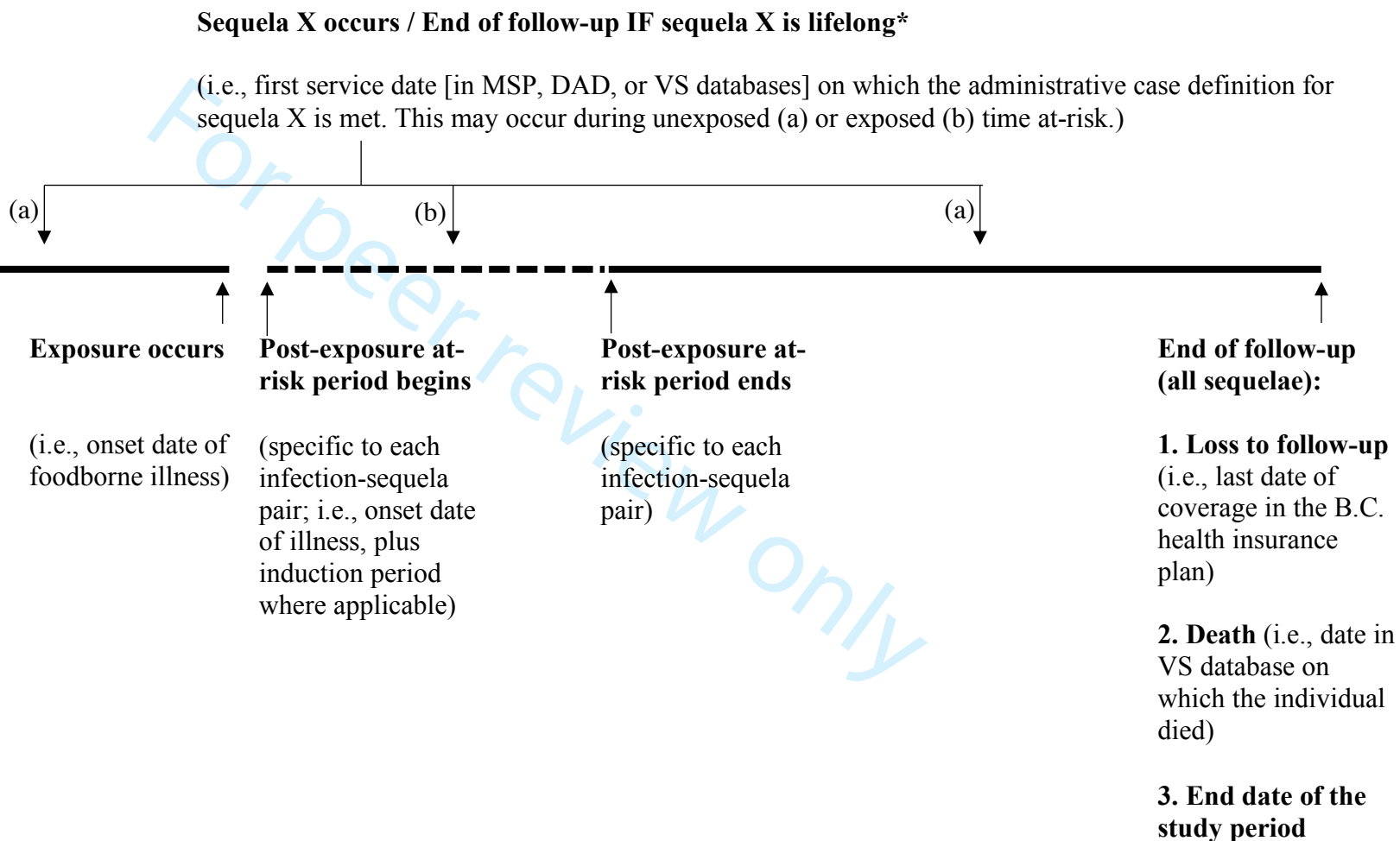
1
2
3 **637 FIGURE LEGEND**
4

5
6 **638**

7
8 **639 Figure 1. Study follow-up period and time-at-risk for development of Sequela X (solid**
9
10 **640 lines: unexposed time at-risk; dashed lines: exposed time at-risk; B.C.,**
11
12 **641 British Columbia, Canada; MSP: Medical Services Plan; DAD: Discharge**
13
14 **642 Abstracts Database; VS: Vital Statistics)**
15
16

17 **643**
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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:	<i>BMJ Open</i>
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

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



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.

1
2
3 1 **Determining the long-term health burden and risk of sequelae for 14 foodborne infections,**
4
5 2 **in British Columbia, Canada: protocol for a retrospective population-based cohort study**
6
7
8 3

9
10 4 Shannon E. Majowicz^{1*}, Dimitra Panagiotoglou², Marsha Taylor³, Mahmood R. Gohari¹, Gilaad
11
12 5 G. Kaplan⁴, Ashok Chaurasia¹, Scott T. Leatherdale¹, Richard J. Cook⁵, David M. Patrick^{3,6},
13
14 6 Steen Ethelberg^{7,8}, Eleni Galanis^{3,6}
15
16

17 7 ¹ School of Public Health and Health Systems, University of Waterloo (Waterloo, Canada)
18

19 8 ² Department of Epidemiology, Biostatistics and Occupational Health, McGill University
20
21 (Montreal, Canada)
22
23

24 10 ³ British Columbia Centre for Disease Control (Vancouver, Canada)
25

26 11 ⁴ Departments of Medicine & Community Health Sciences, University of Calgary (Calgary,
27
28 Canada)
29
30

31 13 ⁵ Department of Statistics and Actuarial Sciences, University of Waterloo (Waterloo, Canada)
32

33 14 ⁶ School of Population and Public Health, University of British Columbia (Vancouver, Canada)
34

35 15 ⁷ Department of Infectious Disease Epidemiology and Prevention, Statens Serum Institut
36
37 (Copenhagen, Denmark)
38
39

40 17 ⁸ Global Health Section, Department of Public Health, University of Copenhagen (Copenhagen,
41
42 Denmark)
43
44

45 19
46
47 20 *Corresponding Author:
48 21 200 University Avenue West
49 22 Waterloo, Ontario, Canada, N2L 3G1
50 23 Email: smajowicz@uwaterloo.ca
51 24 Telephone: +1 519 888 4567 extension 31790
52
53

54 26 KEY WORDS: foodborne infections, sequelae, cohort study, public health
55
56
57
58
59
60

27 ABSTRACT

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

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

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

1
2
3 50 Results will be disseminated via presentations to academics, public health practitioners, and
4
5 51 knowledge users, and publication in peer-reviewed journals. Where such publications are not
6
7 52 open access, manuscripts will also be available via the University of Waterloo's Institutional
8
9
10 53 Repository (<https://uwspace.uwaterloo.ca>).

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

54 STRENGTHS AND LIMITATIONS

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

68 INTRODUCTION

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

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

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

1
2
3 91 Although infections such as *Salmonella* spp. and STEC can be transmitted via several
4
5 92 routes (e.g., person-to-person, water), their transmission via food and their presence throughout
6
7 93 the food system (e.g., food animals as a reservoir for *Campylobacter* spp.,[33] food handlers
8
9 94 shedding hepatitis A virus,[34] the ability of *Listeria monocytogenes* to persist in food
10
11 95 production equipment [35]) mean that these infections are often termed “foodborne” although
12
13 96 some fraction will not be transmitted via food directly. Here, we apply the term “foodborne
14
15 97 infection” to 14 infections that can be transmitted via food (*Clostridium botulinum*,
16
17 98 *Campylobacter*, *Cryptosporidium*, *Cyclospora*, *Giardia*, hepatitis A virus, *Listeria*, non-
18
19 99 typhoidal *Salmonella* spp., *Salmonella* Typhi, *Salmonella* Paratyphi, STEC, *Shigella*, *Vibrio*
20
21 100 *parahaemolyticus*, and *Yersinia* excluding *pestis*), recognizing that not all result from direct
22
23 101 foodborne transmission.
24
25
26
27

28 102 The overall goal of this study is to estimate the health burden and costs of these 14
29
30 103 infections, considering the acute illness and subsequent sequelae and associated mortality, for the
31
32 104 population of British Columbia (B.C.), Canada. Our specific objectives are to:
33
34

- 35 105 1. determine the risk of developing sequelae following infection;
36
37 106 2. describe the epidemiology and clinical progression across the range of outcomes, including
38
39 107 acute illness, sequelae, and death;
40
41 108 3. quantify the direct healthcare costs due to these infections and their various outcomes; and
42
43 109 4. determine the risk of sequelae in the population attributable to these infections.
44
45
46

47 110

48 111 **METHODS AND ANALYSIS**

49 112 **Study setting**

50
51
52
53
54
55
56
57
58
59
60

1
2
3 113 B.C. is Canada's westernmost and third most populous province (~4.7 million circa
4
5 114 2014). The annual incidence of foodborne infections in B.C. is comparable to annual incidences
6
7
8 115 in Canada, the United States, Australia, and Western Europe.[e.g., 2, 36-37]
9

10 116 All B.C. residents (defined as citizens or permanent residents of Canada who are
11
12 117 physically present in B.C. for at least six months in a calendar year), their dependents, and
13
14 118 certain other individuals (e.g., some holders of study or work permits) are covered by the
15
16 119 province's health insurance program.[38] Enrolment is mandatory, and this program covers
17
18 120 nearly all of the B.C. population (with the exception of members of the Canadian military, Royal
19
20 121 Canadian Mounted Police, and some First Nations individuals covered by federal insurance
21
22 122 programs). Physician visits, laboratory tests, hospitalizations, out-patient hospital services, and
23
24 123 drugs for certain populations are among the publicly-funded benefits. The administrative datasets
25
26 124 that contain these health care use data, along with vital statistics (e.g., births, deaths),
27
28 125 demographic data, and other datasets, are held by the B.C. Ministry of Health and are accessible
29
30 126 to researchers via Population Data B.C.,[39] a central repository and "multi-university, data and
31
32 127 education resource" that "support[s] research access to individual-level, de-identified
33
34 128 longitudinal data on British Columbia's 4.7 million residents".[40]
35
36
37
38
39

40 129 The B.C. Public Health Act mandates that reportable diseases, including several
41
42 130 foodborne infections,[41] be reported by health professionals and laboratories to the local and
43
44 131 provincial public health authorities, and these data are managed provincially by the B.C. Centre
45
46 132 for Disease Control (BCCDC). These data are housed within Panorama, the provincial public
47
48 133 health database of reportable diseases. In the Panorama database, as well as the administrative
49
50 134 health and vital statistics databases, individuals are recorded by their unique Personal Health
51
52
53 135 Number (PHN), allowing information from these data sources to be linked by individual.
54
55
56
57
58
59
60

136

137 Study design, population, and timeframe

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

145 The study population is all individuals in B.C. registered with the provincial health
146 insurance program at any point during the study period, i.e., all individuals with the following
147 from 2005 to 2014 inclusive: one or more record in one or more of the Medical Services Plan
148 (MSP), Discharge Abstracts Database (DAD), Vital Statistics Deaths, or PharmaNet; or record of
149 coverage under the provincial insurance program within the Consolidation File database (see
150 Table 1). The 10-year study period is January 1, 2005 to December 31, 2014, inclusive, with
151 additional two-year wash-in (January 1, 2003 to December 31, 2004) and wash-out (January 1,
152 2015 to December 31, 2016) periods. During these periods we will identify occurrences of
153 foodborne infections, sequelae, and death. The 10-year study period was selected to more than
154 encompass timeframes for initial sequelae development and ensuing healthcare use currently
155 reflected in the literature (i.e., days to years), although there is some evidence that sequelae can
156 develop over longer timeframes (e.g., over decades).[42]

1
2
3 157 We assume that enrolment in the provincial health insurance program (i.e., entry into the
4
5 158 study population) and reasons for exit from the cohort (e.g., moving away from B.C.) are not
6
7
8 159 related to the exposures nor the outcomes of interest.
9
10 160

11 161 **Data sources and linkage**

12
13
14 162 The study will use individually-linked, longitudinal data from eight databases to
15
16 163 investigate both acute and longer-term health outcomes following foodborne infection (Table 1).
17
18 164 In totality, these data contain information on 14 reportable foodborne infections, physician and
19
20 165 hospital visits, prescription medications, vital statistics, and various demographic descriptors, for
21
22 166 the B.C. population across the study period. All data will be stored and analyzed within
23
24 167 Population Data B.C.'s virtual Secure Research Environment.
25
26
27
28 168

29
30
31 169 **Table 1. Population-level administrative and reportable disease databases that will be**
32
33 170 **used in this study (British Columbia [B.C.], Canada)**

34 35 Database (Reference)	36 37 Database description and summary of variables included for this study	38 39 Date range
40 <i>Health Care and Health Services Data</i>		
41 42 Medical Services Plan Payment Information File [43]	43 44 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.	45 46 2003/01/01 to 2016/12/31
47 48 Discharge Abstracts Database (Hospital Separations) [44]	49 50 51 52 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.	53 54 2003/01/01 to 2016/12/31
55 56 PharmaNet [45]	57 58 All prescriptions (for drugs and medical supplies) dispensed from community pharmacies, and from hospital outpatient	59 60 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.: <i>Clostridium botulinum</i> , <i>Campylobacter</i> , <i>Cryptosporidium</i> , <i>Cyclospora</i> , <i>Giardia</i> , hepatitis A virus, <i>Listeria</i> , non-typhoidal <i>Salmonella</i> spp., <i>Salmonella</i> Typhi, <i>Salmonella</i> Paratyphi, STEC, <i>Shigella</i> , <i>Vibrio parahaemolyticus</i> , and <i>Yersinia</i> . 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

1
2
3 178 data are first linked to individuals (and their PHNs) using their 6-digit postal code. Once the
4
5 179 linkage is complete, each individual is then assigned a unique study identifier. All individually-
6
7 180 linked, de-identified databases are provided by Population Data B.C. within their Secure
8
9 181 Research Environment, a centralized online platform, accessible via virtual private network
10
11 182 within Canada, for accessing and analyzing research data, with security standards that meet Data
12
13 183 Steward requirements.
14
15
16
17 184

19 185 **Measuring exposure and outcomes**

20
21 186 The exposures of interest are infections with the 14 foodborne pathogens. These 14
22
23 187 infections were selected because they are (a) considered a priority in terms of prevention
24
25 188 potential and health impacts, and (b) capture nearly all reportable foodborne infections in B.C.
26
27 189 Note that brucellosis and paralytic shellfish poisoning were also reportable foodborne infections
28
29 190 in B.C. during the study period. However, since brucellosis is very rare and nearly always travel-
30
31 191 related, and paralytic shellfish poisoning (also rare) is syndromic and diagnosis is uncertain,
32
33 192 these two were not considered for inclusion in this study.
34
35
36
37

38 193 Individuals will be considered exposed when and if they have a laboratory-confirmed and
39
40 194 provincially-reported case of any of the following, recorded in the Panorama database during the
41
42 195 study period: *Clostridium botulinum*, *Campylobacter*, *Cryptosporidium*, *Cyclospora*, *Giardia*,
43
44 196 hepatitis A virus, *Listeria*, *Salmonella* spp. (non-typhoidal, Typhi, Paratyphi), STEC, *Shigella*,
45
46 197 *Vibrio parahaemolyticus*, and *Yersinia* (excluding *pestis*). Case definitions for each of these
47
48 198 infections are specified by the BCCDC.[51]. Individuals without a reported foodborne infection,
49
50 199 but who have International Classification of Disease (ICD) codes either for one of our infections
51
52 200 (e.g., A02.0, ‘*Salmonella* enteritis’) or for non-specific gastroenteritis (e.g., A08.4, ‘viral
53
54
55
56
57
58
59
60

1
2
3 201 intestinal infection, unspecified' [52]) within the MSP and DAD databases, will be considered
4
5 202 potentially exposed. We will describe these individuals as a separate group in our descriptive,
6
7 203 economic, and population attributable fraction analyses, but will remove them from the main
8
9 204 analyses of sequelae risk. We will, however, estimate the risk of sequelae among those who are
10
11 205 potentially exposed as a secondary analysis.
12
13

14
15 206 It is possible for individuals to have more than one reported foodborne infection during
16
17 207 the study period, either as a simultaneously occurring co-infection, or as two or more distinct
18
19 208 events. For these individuals, we will treat this as a complex exposure problem; sequelae will be
20
21 209 associated with the most plausible explanatory infection, considering biology and timing, and we
22
23 210 will adjust for the presence of concurrent foodborne infections if applicable.[53-54]
24
25

26 211 Our primary outcomes of interest are those sequelae for which the link to a given
27
28 212 foodborne infection is either established or is possible (Table 2). We selected sequelae (a) with
29
30 213 evidence of an association with any of the 14 individual foodborne infections,[e.g., 55] and (b)
31
32 214 that occur via direct effects of pathogens or their toxin, or via auto-immune or chronic
33
34 215 inflammatory processes that can be triggered by the infection. We will classify individuals as
35
36 216 having the sequelae via administrative case definitions, using International Classification of
37
38 217 Disease (ICD) diagnostic codes within the MSP, DAD, and VS data, with the exception of
39
40 218 stillbirths which will be determined using recorded events in the VS-Stillbirths database. The
41
42 219 ICD codes in the MSP data are generally considered accurate to the third digit.[56] Although
43
44 220 ICD codes in the VS and DAD data are ordered by most probable diagnosis, we will consider all
45
46 221 codes, regardless of order.
47
48
49
50

51 222 Validation of the ICD codes is currently in progress, via a literature review to identify
52
53 223 administrative case definitions that have been validated in the Canadian context, medical expert
54
55
56
57
58
59
60

1
2
3 224 consultation, and, for those sequelae without a relevant validated definition, a targeted chart
4
5 225 review in B.C. In addition to the ICD codes, we may also use PharmaNet data to improve
6
7
8 226 sequelae classification if needed, by identifying pharmaceuticals given to patients (e.g., Tumor
9
10 227 Necrosis Factor inhibitor use as an indication of reactive arthritis; intravenous immune globulin
11
12 228 use for GBS).

13
14
15 229 Some of our sequelae of interest are lifelong (e.g., Graves' Disease), and some are
16
17 230 transient in that complete recovery is possible (e.g., GBS, stillbirth). For lifelong sequelae, we
18
19 231 will consider the individual as having the sequela on the earliest date they meet the
20
21 232 administrative case definition for that sequela (with subsequent records considered as a
22
23 233 continuation of the original event). For sequelae from which recovery and subsequent return to
24
25 234 being at-risk is possible, we will consider the individual as first having the sequela on the earliest
26
27 235 date they meet the administrative case definition for that sequela; we will then apply a post-
28
29 236 sequela recovery time to determine the date on which the individual can be considered to be at-
30
31 237 risk for a new, subsequent occurrence of that sequela.

32
33
34
35 238 Individuals may develop more than one sequela during the study period, either because
36
37 239 they develop multiple different sequelae (e.g., HUS and GBS), or because they develop multiple
38
39 240 occurrences of a single sequela from which complete recovery is possible (e.g., GBS, stillbirth).
40
41 241 In all instances, the occurrence of multiple sequelae will be recorded and described. When
42
43 242 individuals develop multiple different sequelae during the study period (e.g., HUS and GBS), we
44
45 243 will treat these as distinct outcomes in our risk estimates. When individuals develop multiple
46
47 244 occurrences of the same sequela, we will treat these as distinct outcomes but account for
48
49 245 recurrent events.[54]
50
51
52
53
54
55
56
57
58
59
60

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

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

257 **Table 2. Established (E) and possible (P) sequelae of foodborne infections, that will be**
 258 **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
<i>Campylobacter</i>	E	P			E	P	P	P			E	

<i>Cryptosporidium</i>												P	
<i>Cyclospora</i>					P							P	
<i>Giardia</i>								P				P	
Hepatitis A virus	E				P								
<i>Listeria monocytogenes</i>									E ⁵	E			
<i>Salmonella</i> (non-typhoidal)	E				P	P	P	P				E	
<i>Salmonella</i> Paratyphi	E				P	P	P	P				E	
<i>Salmonella</i> Typhi	E				P	P	P	P				E	
STEC	E					E						P	P
<i>Shigella</i>	E					E		P				E	P
<i>Yersinia</i> (excluding <i>pestis</i>)	E		E	P				P				E	

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

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

262 ³ considered here as a sequela of maternal *Listeria* infection

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

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

265

266

267 **Measuring time at-risk**

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

1
2
3 271 may contribute to both the exposed time-at-risk (during the post-infection ‘at-risk’ period, see
4
5 272 below) and the unexposed time-at-risk (prior to, and after, the post-infection ‘at-risk’ period),
6
7
8 273 while individuals without foodborne infections will only contribute to the unexposed time-at-
9
10 274 risk. Time-at-risk for a specific sequela will be measured in days, from the date of entry into the
11
12 275 study, until: the development of that sequela, death, loss to follow-up, or the end of the study.
13
14
15 276 We define loss to follow-up as the last date of coverage in the provincial health insurance plan,
16
17 277 calculated using the start day registered in the most recent year plus the total days registered in
18
19 278 that year.

20
21
22 279 During the unexposed time-at-risk, we will treat all individuals as having the potential to
23
24 280 develop any of the sequelae (with the exception of neonatal listeriosis and stillbirth, for which
25
26 281 only those who are pregnant are at risk). For those who develop a foodborne infection,
27
28 282 unexposed time at-risk will end on the onset date of the infection. Infection onset date will be
29
30
31 283 determined using the onset date reported in Panorama, and where this is missing, the date that the
32
33 284 infection was reported minus the number of days between onset to reporting (e.g., estimated
34
35 285 using the Panorama data or from the literature).[57]

36
37
38 286 Exposed time at-risk will be measured starting from the infection onset date, plus any
39
40 287 additional induction periods (specific to each sequela and currently being determined via
41
42 288 literature review and medical expert consultation). The end of the exposed time at-risk period is
43
44
45 289 currently being determined via literature review and medical expert consultation. During the
46
47 290 exposed time-at-risk, individuals will be classified as having a sequela specific to their infection
48
49 291 (Table 2) on the date within the ‘at-risk’ period on which they meet the administrative case
50
51 292 definition for that sequela (e.g., the date of the physician visit or hospitalization). After the post-
52
53
54 293 infection at-risk period ends, individuals will revert to contributing to the unexposed time-at-risk.

1
2
3 294 These data are subject to censoring and truncation. Individuals will be censored for the
4
5 295 sequela in the event of: death, last date of coverage in the provincial health insurance plan (i.e.,
6
7 296 loss to follow up; calculated as above), or the end of the study period, whichever comes first.[58-
8
9 297 59] In our descriptive and economic analyses, we will include all related health care use and
10
11 298 prescription medication costs over the course of the infection and sequela(e), and in our
12
13 299 estimates of mortality we will include any deaths recorded during the study period, following the
14
15 300 sequela. We will determine whether health care use is related to infection and sequelae using
16
17 301 ICD diagnosis codes, and we will determine whether prescription medication use is related via
18
19 302 medical expert consultation.
20
21
22
23
24
25

26 304 **Analysis plan**

27
28 305 Data will be analyzed and results reported following the STROBE and RECORD
29
30 306 guidelines.[60-61]. Analyses will be performed using SAS 9.4 (SAS Institute Inc., Cary, NC,
31
32 307 USA) and R.[62] The nature and extent of missing data will be described. If imputation is used
33
34 308 to complete missing data, specific methods and assumptions will be reported. We will emphasize
35
36 309 estimation over tests of statistical significance by reporting relative measures of effect along with
37
38 310 associated 95% confidence intervals.
39
40
41

42 311 The datasets in Table 1 contain the variable “sex” (identified via government records),
43
44 312 that denotes whether individuals are ‘male’ or ‘female’, thereby capturing a composite of sex and
45
46 313 gender. To reflect for potential sex- and gender-differences, we will report and interpret findings
47
48 314 stratified by this variable, in addition to overall findings.
49
50

51 315 *Objective 1:* To determine the risk of developing sequelae following foodborne infection,
52
53 316 we will estimate hazard ratios using Cox regression models,[63] that adjust for confounders and
54
55
56
57
58
59

1
2
3 317 comorbidities (see below), along with the possible effect modifying role of age, sex,
4
5 318 comorbidity, and medication use (e.g., antibiotics). We will include comorbidities as a composite
6
7 319 score, using the revised Charlson comorbidity index and its associated coding algorithms.[64-67]
8
9
10 320 Following foodborne infection, we will compare the cumulative risk of first diagnosis of each
11
12 321 infection-specific sequela using life-table and Kaplan-Meier approaches.[63, 68] In the event an
13
14 322 individual dies, we will use competing risk analysis.[69-70] For those who experience more than
15
16 323 one foodborne infection across the study period, we will explore the impacts of having multiple
17
18 324 infections on the risk of sequelae and mortality. We will also estimate the likelihood of dying
19
20 325 following foodborne infection using the same methods described above.
21
22

23
24 326 *Objective 2:* To describe the epidemiology and clinical progression across the range of
25
26 327 outcomes, including acute illness (typically diarrhea and other gastrointestinal symptoms),
27
28 328 sequelae, and death, for each foodborne infection we will calculate incidence rates, demographic,
29
30 329 geographic, and temporal distributions, timing and progression of outcomes, and case fatality
31
32 330 rates, for both the acute stage, and sequelae associated to the foodborne infection.
33
34

35 331 *Objective 3:* To quantify the direct healthcare costs due to these infections and their
36
37 332 various outcomes, we will determine health service use (i.e., patterns of use by type, frequency,
38
39 333 timing of physician visits and hospitalizations), for both the acute foodborne infection and any
40
41 334 sequelae. We will estimate direct healthcare costs of out-of-hospital physician visits,
42
43 335 hospitalizations, and prescription medications. Out-of-hospital costs will be estimated using the
44
45 336 MSP variables 'Fee Item' and 'Paid Service' and fee rates from the B.C. fee schedule.[71] Costs
46
47 337 of in-patient and day-case hospitalizations will be calculated using established case-mix
48
49 338 methodology (i.e., using the 'Resource Intensity Weight' of each hospitalization),[72] and the
50
51 339 B.C. Ministry of Health unit costs for hospital stays.[73] Total prescription medication costs will
52
53
54
55
56
57
58
59

1
2
3 340 be calculated using the drug cost claimed by the pharmacist which includes the ingredient cost,
4
5 341 professional dispensing fee and other special service fees (if applicable). Because these costs are
6
7 342 captured directly in the PharmaNet data, they will be tallied directly. We will also apply these
8
9 343 methods to determine the direct costs per sequelae, regardless of exposure. Costs will be adjusted
10
11 344 for inflation using the Canadian Consumer Price Index.[74] Results will be reported to allow
12
13 345 comparability with other estimates (e.g., 2010/2011 Canadian and US dollars).
14

15
16
17 346 *Objective 4:* To determine the risk of sequelae in the population attributable to foodborne
18
19 347 infections, we will calculate population attributable fractions using standard formulae.[75] We
20
21 348 will also describe the proportion of cases of each sequela with specific foodborne infections.
22
23 349 Here, the total number of cases of each sequela occurring in B.C. during the study period will be
24
25 350 the denominator (e.g., total number of cases of acute kidney injury), and the numerators will be
26
27 351 the numbers of cases of each sequela occurring in those with specific foodborne infections (e.g.,
28
29 352 total number of cases of hepatitis A virus, and of STEC, with acute kidney injury). We will also
30
31 353 describe the proportion of individuals with sequela who do not have a foodborne infection, but
32
33 354 who do have an ICD code for prior gastroenteritis, and use this to estimate the additional
34
35 355 proportions of sequelae that may have an unidentified foodborne infection. We will calculate
36
37 356 population attributable fractions and proportions for both established and possible sequelae, but
38
39 357 clearly distinguish between the two when reporting findings.
40
41
42
43
44

45 358

46 47 359 **Potential confounders and their adjustment**

48
49 360 We will use propensity score matching, and inclusion of potential confounders as
50
51 361 covariates in our analyses, as our primary methods to adjust for confounding.[76] The databases
52
53 362 in Table 1 include direct measurements of important known, strong confounders (e.g., age, sex),
54
55
56
57
58
59

1
2
3 363 as well as other potential factors (e.g., use of protein pump inhibitors and antibiotics, disease
4
5 364 severity). We will consider the following variables as potential confounders: age, sex, local
6
7
8 365 health area, income band/area income (as a proxy for socioeconomic status), month/year,
9
10 366 seasonality, immune status (e.g., indication of immunosuppressant drugs, presence of conditions
11
12 367 like cancer, pregnancy), use of medications like antibiotics, and Charlson comorbidity index.
13
14 368 Because none of our data sources include ethnicity nor race, we are unable to adjust, or conduct
15
16 369 sub-analyses, for these factors. Although ethnicity may impact health care seeking behaviours,
17
18 370 we anticipate these impacts will apply equally regardless of exposure, and thus we expect any
19
20 371 bias in our findings to be negligible. Nevertheless, to assess residual confounding, we will
21
22 372 conduct sensitivity analyses,[77-78] and perform indirect adjustments.[79].
23
24
25
26 373

27 28 374 **Planned sensitivity analyses and study limitations**

29
30 375 We are planning several sensitivity analyses to explore assumptions, methodological
31
32 376 decisions, limitations in the data, and robustness of results. We will explore the impact of
33
34 377 propensity score matching on our sequelae and mortality risk estimates by also using (a) the
35
36 378 whole unexposed population, and (b) a random sample of unexposed individuals (matched on
37
38 379 time), instead of propensity score-matched individuals. We may also explore additional matching
39
40 380 and control strategies (e.g., matching on age and sex). We will also explore the impacts of
41
42 381 including individuals with foodborne infections who develop a sequela for which there is no
43
44 382 current evidence of an established or possible link to the specific pathogen (e.g., *Campylobacter*
45
46 383 and stillbirth), in our estimates of sequelae risk. Our primary analyses will use a composite of all
47
48 384 infections (i.e., a report of any of the 14 foodborne infections) and their various sequelae. We
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 385 will also analyze and present results for each of the individual foodborne infections, and for each
4
5 386 of the sequela.

6
7
8 387 A main recognized limitation of reportable disease data, such as the Panorama data in this
9
10 388 study, is the under-ascertainment of foodborne infections. Here, this limitation means that
11
12 389 individuals with foodborne infections who do not seek care, do not get tested, or who test
13
14 390 negative, will be misclassified as unexposed. We will assess the impacts of such potential
15
16 391 misclassification via sensitivity analyses that illustrate how our findings could be impacted by
17
18 392 different misclassification rates, using estimates of misclassification from the literature,[e.g., 2,
19
20 393 80] and from our data (e.g., individuals with non-specific gastroenteritis). An additional
21
22 394 limitation is that if sequelae develop over longer timeframes than our 10-year study (e.g., over
23
24 395 decades),[42] our study cannot assess this scenario.
25
26
27
28
29
30

31 397 **Patient and public involvement statement**

32
33 398 Patients were not involved in the development of this protocol, nor were members of the
34
35 399 public.
36
37

38 400

39 401 **ETHICS AND DISSEMINATION**

40
41
42 402 This study has received approval by a University of Waterloo Research Ethics Committee
43
44 403 (#30645), the University of British Columbia Behavioral Research Ethics Board (#H16-00021),
45
46 404 and McGill University's Institutional Review Board (#A03-M12-19A). In addition to conference
47
48 405 presentations and dissemination to public health practitioners and other knowledge users, results
49
50 406 will be published in peer-reviewed journals, and where such publications are not open access,
51
52
53
54
55
56
57
58
59
60

1
2
3 407 they will also be stored on UWSpace, the University of Waterloo's Institutional Repository
4
5 408 (<https://uwspace.uwaterloo.ca>).
6
7

8 409
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

410 **ACKNOWLEDGEMENTS**

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

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

419 **AUTHORS' CONTRIBUTIONS**

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

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

430 **FUNDING**

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

1
2
3 433 from the University of Waterloo's Research Incentive Fund (PI: Shannon Majowicz) and the
4
5 434 British Columbia Centre for Disease Control Foundation for Public Health's Open Award
6
7
8 435 Program (PI: Eleni Galanis).
9

10 436

11 437 **DISCLAIMER**

12
13
14
15 438 All inferences, opinions, and conclusions drawn in this research protocol are those of the
16
17 439 authors, and do not reflect the opinions or policies of the Data Steward(s).
18

19 440

20 441 **COMPETING INTERESTS**

21
22
23
24 442 Drs. Majowicz and Galanis report funding for this study as per the funding statement.
25
26 443 Dr. Majowicz reports other relationships; she is an Associate Editor at Epidemiology and
27
28 444 Infection (for which she receives a small honorarium); she has served as a paid Expert on behalf
29
30 445 of the Attorney General of Canada in legal proceedings, providing evidence on the public health
31
32 446 risks and benefits of unpasteurized milk; and she is an expert on the Joint FAO/WHO Expert
33
34 447 Meetings on Microbiological Risk Assessment (JEMRA) Roster of Experts. Dr. Kaplan reports
35
36 448 honoraria for speaking or consultancy from Abbvie, Janssen, Pfizer, and Takeda. He has
37
38 449 received research support from Ferring, Janssen, Abbvie, GlaxoSmith Kline, Merck, and Shire.
39
40 450 He shares ownership of a patent: TREATMENT OF INFLAMMATORY DISORDERS,
41
42 451 AUTOIMMUNE DISEASE, AND PBC. UTI Limited Partnership, assignee. Patent
43
44
45 452 WO2019046959A1. PCT/CA2018/051098. 7 Sept. 2018. Dr. Galanis' spouse works for QHR
46
47 453 Technologies, a Canadian medical records company; these records were not used in this study.
48
49
50 454 All other authors have nothing to disclose.
51
52
53

54 455

1
2
3 456 **DATA SHARING**
4

5 457 Open access for these data is not permitted by the data stewards; further details on the
6
7 458 legislation and agreements can be found at: <https://www.popdata.bc.ca/dataaccess/rdaf/history>
9
10 459 and <https://www.popdata.bc.ca/dataaccess/rdaf/expectations>. To access the data used for this
11
12 460 study, researchers must submit a Data Access Request through Population Data B.C. (for all
13
14 461 databases except Panorama) and the Panorama Data Governance Committee (Panorama data),
15
16 462 who have record of the files and fields provided (cite project: Majowicz Galanis 15-180). We
17
18 463 will make the programming code used to clean and analyse the data available (on request, or via
19
20 464 publications where possible).
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

465 **REFERENCES**

- 466 1. Havelaar AH, Kirk MD, Torgerson PR, et al. World Health Organization global estimates
467 and regional comparisons of the burden of foodborne disease in 2010. *PLoS Med*
468 2015;12(12):e1001923.
- 469 2. Thomas MK, Murray R, Flockhart L, et al. Estimates of the burden of foodborne illness in
470 Canada for 30 specified pathogens and unspecified agents, circa 2006. *Foodborne Pathog*
471 *Dis.* 2013;10(7):639-48.
- 472 3. Awofisayo A, Amar C, Ruggles R, et al. Pregnancy-associated listeriosis in England and
473 Wales. *Epidemiol Infect.* 2015;143(2):249-56.
- 474 4. Charlier C, Perrodeau É, Leclercq A, et al. Clinical features and prognostic factors of
475 listeriosis: the MONALISA national prospective cohort study. *Lancet Infect Dis.*
476 2017;17(5):510-9.
- 477 5. Karmali MA, Petric M, Lim C, et al. The association between idiopathic hemolytic uremic
478 syndrome and infection by verotoxin-producing *Escherichia coli*. *J Infect Dis.*
479 1985;151(5):775-82.
- 480 6. Gould LH, Demma L, Jones TF, et al. Hemolytic uremic syndrome and death in persons with
481 *Escherichia coli* O157: H7 infection, Foodborne Diseases Active Surveillance Network sites,
482 2000–2006. *Clin Infect Dis.* 2009;49(10):1480-5.
- 483 7. Rodríguez LA, Ruigómez A, Panés J. Acute gastroenteritis is followed by an increased risk
484 of inflammatory bowel disease. *Gastroenterol.* 2006;130(6):1588-94.
- 485 8. Small CL, Xing L, McPhee J, et al. Acute infectious gastroenteritis potentiates a Crohn's
486 disease pathobiont to fuel ongoing inflammation in the post-infectious period. *PLoS Pathog.*
487 2016. <http://dx.doi.org/10.1371/journal.ppat.1005907>.

- 1
2
3 488 9. Ruzante JM, Majowicz SE, Fazil A, et al. Hospitalization and deaths for select enteric
4
5 489 illnesses and associated sequelae in Canada, 2001-2004. *Epidemiol Infect.* 2011;139(6):937-
6
7 490 45.
- 9
10 491 10. Tam CC, Rodrigues LC, Petersen I, et al. Incidence of Guillain-Barré syndrome among
11
12 492 patients with *Campylobacter* infection: a general practice research database study. *J Infect*
13
14 493 *Dis.* 2006;194(1):95-7.
- 16
17 494 11. Helms M, Vastrup P, Gerner-Smidt P, et al. Short and long term mortality associated with
18
19 495 foodborne bacterial infections: registry-based study. *Br Med J.* 2003;326:1-5.
- 21
22 496 12. Bremell T, Bjelle A, Svedhem A. Rheumatic symptoms following an outbreak of
23
24 497 campylobacter enteritis: a five year follow up. *Ann Rheum Dis.* 1991;50(12):934-8.
- 25
26 498 13. Marshall JK, Thabane M, Garg AX, et al. Incidence and epidemiology of irritable bowel
27
28 499 syndrome after a large waterborne outbreak of bacterial dysentery. *Gastroenterol.*
29
30 500 2006;131(2):445-50.
- 32
33 501 14. Garg AX, Pope JE, Thiessen-Philbrook H, et al. Arthritis risk after acute bacterial
34
35 502 gastroenteritis. *Rheumatol.* 2008;47(2):200-4.
- 37
38 503 15. Hanevik K, Wensaas KA, Rortveit G, et al. Irritable bowel syndrome and chronic fatigue 6
39
40 504 years after giardia infection: a controlled prospective cohort study. *Clin Infect Dis.*
41
42 505 2014;59(10):1394-400.
- 44
45 506 16. Rehn M, Wallensten A, Widerström M, et al. Post-infection symptoms following two large
46
47 507 waterborne outbreaks of *Cryptosporidium hominis* in Northern Sweden, 2010–2011. *BMC*
48
49 508 *Public Health.* 2015;15(1):529.
- 51
52 509 17. Hannu T, Mattila L, Rautelin H, et al. Campylobacter-triggered reactive arthritis: a
53
54 510 population-based study. *Rheumatol.* 2002;41(3):312-8.

- 1
2
3 511 18. Hunter PR, Hughes S, Woodhouse S, et al J. Health sequelae of human cryptosporidiosis in
4
5 512 immunocompetent patients. *Clin Infect Dis*. 2004;39(4):504-10.
6
7
8 513 19. Wang LH, Fang XC, Pan GZ. Bacillary dysentery as a causative factor of irritable bowel
9
10 514 syndrome and its pathogenesis. *Gut*. 2004;53(8):1096-101.
11
12 515 20. Hannu T, Mattila L, Siitonen A, et al. Reactive arthritis attributable to *Shigella* infection: a
13
14 516 clinical and epidemiological nationwide study. *Ann Rheum Dis*. 2005;64(4):594-8.
15
16
17 517 21. Saps M, Pensabene L, Di Martino L, et al. Post-infectious functional gastrointestinal
18
19 518 disorders in children. *J Pediatr*. 2008;152(6):812-6.
20
21
22 519 22. Rosner BM, Werber D, Höhle M, et al. Clinical aspects and self-reported symptoms of
23
24 520 sequelae of *Yersinia enterocolitica* infections in a population-based study, Germany 2009–
25
26 521 2010. *BMC Infect Dis*. 2013;13(1):236.
27
28
29 522 23. Litleskare S, Rortveit G, Eide GE, et al. Prevalence of irritable bowel syndrome and chronic
30
31 523 fatigue 10 years after *Giardia* infection. *Clin Gastroenterol Hepatol*. 2018;16(7):1064-72.
32
33 524 24. Helms M, Simonsen J, Mølbak K. Foodborne bacterial infection and hospitalization: a
34
35 525 registry-based study. *Clin Infect Dis*. 2006;42(4):498-506.
36
37
38 526 25. Rodríguez LA, Ruigómez A, Panés J. Acute gastroenteritis is followed by an increased risk
39
40 527 of inflammatory bowel disease. *Gastroenterol*. 2006;130(6):1588-94.
41
42 528 26. Ternhag A, Törner A, Svensson Å, et al. Short-and long-term effects of bacterial
43
44 529 gastrointestinal infections. *Emerg Infect Dis*. 2008;14(1):143.
45
46
47 530 27. Gradel KO, Nielsen HL, Schönheyder HC, et al. Increased short-and long-term risk of
48
49 531 inflammatory bowel disease after *Salmonella* or *Campylobacter* gastroenteritis.
50
51 532 *Gastroenterol*. 2009;137(2):495-501.
52
53
54
55
56
57
58
59
60

- 1
2
3 533 28. Jess T, Simonsen J, Nielsen NM, et al. Enteric *Salmonella* or *Campylobacter* infections and
4
5 534 the risk of inflammatory bowel disease. *Gut*. 2011;60(3):318-24.
6
7
8 535 29. Porter CK, Choi D, Cash B, et al. Pathogen-specific risk of chronic gastrointestinal disorders
9
10 536 following bacterial causes of foodborne illness. *BMC Gastroenterol*. 2013;13(1):46.
11
12 537 30. Riddle MS, Murray JA, Cash BD, et al. Pathogen-specific risk of celiac disease following
13
14 538 bacterial causes of foodborne illness: a retrospective cohort study. *Digestive Diseases and*
15
16 539 *Sciences*. 2013;58(11):3242-5.
17
18
19 540 31. Nakao JH, Collier SA, Gargano JW. Giardiasis and subsequent irritable bowel syndrome: a
20
21 541 longitudinal cohort study using health insurance data. *J Infect Dis*. 2017;215(5):798-805.
22
23
24 542 32. Painter JE, Collier SA, Gargano JW. Association between *Giardia* and arthritis or joint pain
25
26 543 in a large health insurance cohort: could it be reactive arthritis? *Epidemiol Infect*.
27
28 544 2017;145(3):471-7.
29
30
31 545 33. Kaakoush NO, Castaño-Rodríguez N, Mitchell HM, et al. Global epidemiology of
32
33 546 *Campylobacter* infection. *Clin Microbiol Rev*. 2015;28(3):687-720.
34
35
36 547 34. Tricco AC, Duval B, De Serres G, et al. A review of interventions triggered by hepatitis A
37
38 548 infected food-handlers in Canada. *BMC Health Serv Res*. 2006;6(1):157.
39
40 549 35. Carpentier B, Cerf O. Persistence of *Listeria monocytogenes* in food industry equipment and
41
42 550 premises. *Int J Food Microbiol*. 2011;145(1):1-8.
43
44
45 551 36. British Columbia Centre for Disease Control. British Columbia Annual Summary of
46
47 552 Reportable Diseases, 2017. [http://www.bccdc.ca/resource-](http://www.bccdc.ca/resource-gallery/Documents/Statistics%20and%20Research/Statistics%20and%20Reports/Epid/Annual%20Reports/2017CDAnnualReportFinal.pdf)
48
49 553 [gallery/Documents/Statistics%20and%20Research/Statistics%20and%20Reports/Epid/Annua](http://www.bccdc.ca/resource-gallery/Documents/Statistics%20and%20Research/Statistics%20and%20Reports/Epid/Annual%20Reports/2017CDAnnualReportFinal.pdf)
50
51 554 [l%20Reports/2017CDAnnualReportFinal.pdf](http://www.bccdc.ca/resource-gallery/Documents/Statistics%20and%20Research/Statistics%20and%20Reports/Epid/Annual%20Reports/2017CDAnnualReportFinal.pdf) (accessed 12 November 2019).
52
53
54
55
56
57
58
59
60

- 1
2
3 555 37. Scallan E, Majowicz SE, Hall G, et al. Prevalence of diarrhoea in the community in
4
5 556 Australia, Canada, Ireland, and the United States. *Int J Epidemiol.* 2005;34(2):454-60.
6
7
8 557 38. Government of British Columbia. [https://www2.gov.bc.ca/gov/content/health/health-drug-](https://www2.gov.bc.ca/gov/content/health/health-drug-coverage/msp/bc-residents/eligibility-and-enrolment/are-you-eligible)
9
10 558 [coverage/msp/bc-residents/eligibility-and-enrolment/are-you-eligible](https://www2.gov.bc.ca/gov/content/health/health-drug-coverage/msp/bc-residents/eligibility-and-enrolment/are-you-eligible) (accessed 12 November
11
12 559 2019).
13
14
15 560 39. Ark TK, Kesselring S, Hills B, et al. Population Data BC: Supporting population data science
16
17 561 in British Columbia. *Int J Pop Data Sci.* 2020;4(2):14.
18
19 562 40. Population Data BC. About PopData. <https://www.popdata.bc.ca/aboutus> (accessed 12
20
21 563 November 2019).
22
23
24 564 41. Government of British Columbia. Public Health Act [SBC 2008] Chapter 28.
25
26 565 http://www.bclaws.ca/civix/document/id/complete/statreg/08028_01 (accessed 12 November
27
28 566 2019).
29
30
31 567 42. Moorin RE, Heyworth JS, Forbes GM, et al. Long-term health risks for children and young
32
33 568 adults after infective gastroenteritis. *Emerg Infect Dis.* 2010;16(9):1440.
34
35 569 43. BC Ministry of Health [creator] (2018): Medical Services Plan (MSP) Payment Information
36
37 570 File. Population Data BC [publisher]. Data Extract. MOH (2018).
38
39 571 <http://www.popdata.bc.ca/data>
40
41
42 572 44. Canadian Institute for Health Information [creator] (2019): Discharge Abstract Database
43
44 573 (Hospital Separations). Population Data BC [publisher]. Data Extract. MOH
45
46 574 (2018). <http://www.popdata.bc.ca/data>
47
48
49 575 45. BC Ministry of Health [creator] (2019): PharmaNet. BC Ministry of Health [publisher]. Data
50
51 576 Extract. Data Stewardship Committee (2018). <http://www.popdata.bc.ca/data>
52
53
54
55
56
57
58
59
60

- 1
2
3 577 46. BC Vital Statistics Agency [creator] (2019): Vital Statistics Deaths. Population Data BC
4
5 578 [publisher]. Data Extract BC Vital Statistics Agency (2018). <http://www.popdata.bc.ca/data>
6
7
8 579 47. BC Vital Statistics Agency [creator] (2019): Vital Statistics Stillbirths. Population Data BC
9
10 580 [publisher]. Data Extract. BC Vital Statistics Agency (2018). <http://www.popdata.bc.ca/data>
11
12 581 48. British Columbia Ministry of Health [creator] (2019): Consolidation File (MSP Registration
13
14 582 & Premium Billing). Population Data BC [publisher]. Data Extract. MOH
15
16 583 (2018). <http://www.popdata.bc.ca/data>
17
18
19 584 49. Statistics Canada [creator] (2009): Statistics Canada Income Band Data. Catalogue Number:
20
21 585 13C0016. Population Data BC [publisher]. Data Extract. Population Data BC (2018).
22
23 586 <http://www.popdata.bc.ca/data>
24
25
26 587 50. Population Data BC. Data Linkage. <https://www.popdata.bc.ca/datalinkage> (accessed 12
27
28 588 November 2019).
29
30
31 589 51. British Columbia Centre for Disease Control. Case Definitions. [http://www.bccdc.ca/health-](http://www.bccdc.ca/health-professionals/clinical-resources/case-definitions)
32
33 590 [professional/clinical-resources/case-definitions](http://www.bccdc.ca/health-professionals/clinical-resources/case-definitions) (accessed 12 November 2019).
34
35 591 52. World Health Organization. ICD-10 Version:2016. <https://icd.who.int/browse10/2016/en>
36
37 592 (accessed 12 November 2019).
38
39
40 593 53. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. Springer
41
42 594 Science & Business Media 2013.
43
44
45 595 54. Cook RJ, Lawless JF. The Statistical Analysis of Recurrent Events. New York, NY:
46
47 596 Springer-Verlag, 2007.
48
49 597 55. Batz MB, Henke E, Kowalcyk B. Long-term consequences of foodborne infections. *Infect*
50
51 598 *Dis Clinics N Am.* 2013;27(3):599-616.
52
53
54
55
56
57
58
59
60

- 1
2
3 599 56. Hu W. Diagnostic Codes in MSP Claim Data, Summary Report. Victoria: Medical Services
4
5 600 Plan; 1996.
6
7
8 601 57. Galanis E, Taylor M, Romanowski K, et al. Evaluating the timeliness of enteric disease
9
10 602 surveillance in British Columbia, Canada, 2012-13. *Can J Infect Dis Med Microbiol*. 2017
11
12 603 Article ID 9854103, 7 pages, 2017.
13
14
15 604 58. Klein JP, Moeschberger ML. Survival Analysis: Techniques for Censored and Truncated
16
17 605 Data, 2nd ed. New York, NY: Springer-Verlag New York Inc., 2003.
18
19 606 59. Qin J, Shen Y. Statistical methods for analyzing right-censored length-biased data under Cox
20
21 607 model. *Biometrics*. 2010;66(2):382–392.
22
23
24 608 60. von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational
25
26 609 Studies in Epidemiology (STROBE) statement: guidelines for reporting observational
27
28 610 studies. *BMJ*. 2007;335(7624):806-8.
29
30
31 611 61. Benchimol EI, Smeeth L, Guttman A, et al. The REporting of studies Conducted using
32
33 612 Observational Routinely-collected health Data (RECORD) Statement. *PLoS Med*. 2015;12(10):
34
35 613 e1001885. <https://doi.org/10.1371/journal.pmed.1001885>
36
37 614 62. R Core Team (2019). R: A language and environment for statistical computing. R
38
39 615 Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
40
41
42 616 63. Cox DRJ. Regression models and life-tables (with discussion). *J R Stat Soc, B*. 1972;34:187-
43
44 617 220.
45
46
47 618 64. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic
48
49 619 comorbidity in longitudinal studies: development and validation. *J Chron Dis*.
50
51 620 1987;40(5):373-83.
52
53
54 621 65. Charlson M, Szatrowski TP, Peterson J, et al. Validation of a combined comorbidity index. *J*
55
56 622 *Clin Epidemiol*. 1984;47(11):1245-51.
57
58
59
60

- 1
2
3 623 66. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in
4
5 624 ICD-9-CM and ICD-10 administrative data. *Med Care*. 2005;43(11):1130-9.
6
7
8 625 67. Hude Quan, Bing Li, Chantal M, et al. Updating and validating the Charlson comorbidity
9
10 626 index and score for risk adjustment in hospital discharge abstracts using data from 6
11
12 627 countries. *Am J Epidemiol*. 2011;173(6):676–82.
13
14
15 628 68. Kaplan EL, Meier PJ. Nonparametric estimation for incomplete observations. *J Am Stat*
16
17 629 *Assoc*. 1958;53:457-81.
18
19 630 69. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J*
20
21 631 *Am Stat Assoc*. 1999;94(446):496-509.
22
23
24 632 70. Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. *Am J*
25
26 633 *Epidemiol*. 2009;170(2):244-256.
27
28
29 634 71. Government of British Columbia. Medical Services Commission Payment Schedule.
30
31 635 December 1, 2015. [https://www2.gov.bc.ca/assets/gov/health/practitioner-pro/medical-](https://www2.gov.bc.ca/assets/gov/health/practitioner-pro/medical-services-plan/msc_payment_schedule.pdf)
32
33 636 [services-plan/msc_payment_schedule.pdf](https://www2.gov.bc.ca/assets/gov/health/practitioner-pro/medical-services-plan/msc_payment_schedule.pdf) (accessed 12 November 2019).
34
35 637 72. Poole B, Robinson S, MacKinnon M. Resource Intensity Weights™ and Canadian hospital
36
37 638 costs: some preliminary data. *Healthc Manage Forum*. 1998;11(1):22-26.
38
39
40 639 73. Canadian Institute for Health Information. Patient Cost Estimator: Methodological Notes and
41
42 640 Glossary. Ottawa, ON: CIHI; 2016. https://www.cihi.ca/en/pce_methodology_notes_en.pdf
43
44 641 (accessed 26 November 2019).
45
46
47 642 74. Statistics Canada, Consumer Price Index (CPI).
48
49 643 <http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=2301> (accessed 12
50
51 644 November 2019).
52
53
54 645 75. Mansournia MA, Altman DG. Population attributable fraction. *BMJ*. 2018;360:k757.
55
56
57
58
59
60

- 1
2
3 646 76. Holmes WM. Using Propensity Scores in Quasi-Experimental Designs. Los Angeles, Calif.:
4
5 647 Sage Publishing, 2014.
6
7
8 648 77. Stürmer T, Wyss R, Glynn RJ, et al. Propensity scores for confounder adjustment when
9
10 649 assessing the effects of medical interventions using nonexperimental study designs. *J Int*
11
12 650 *Med.* 2014;275(6):570-80.
13
14
15 651 78. Brookhart MA, Schneeweiss S, Rothman KJ, et al. Variable selection for propensity score
16
17 652 models. *Am J Epidemiol.* 2006;163(12):1149-56.
18
19 653 79. Kuenzig ME, Barnabe C, Seow CH, et al. Asthma is associated with subsequent development
20
21 654 of inflammatory bowel disease: a population-based case-control study. *Clin Gastroenterol*
22
23 655 *Hepatol.* 2017;15(9):1405-12.
24
25
26 656 80. MacDougall L, Majowicz S, Doré K, et al. Under-reporting of infectious gastrointestinal
27
28 657 illness in British Columbia, Canada: who is counted in provincial communicable disease
29
30 658 statistics? *Epidemiol Infect.* 2008;136(2):248-56.
31
32

33 659
34
35 660
36
37
38 661
39
40 662
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 663 **FIGURE LEGEND**
4

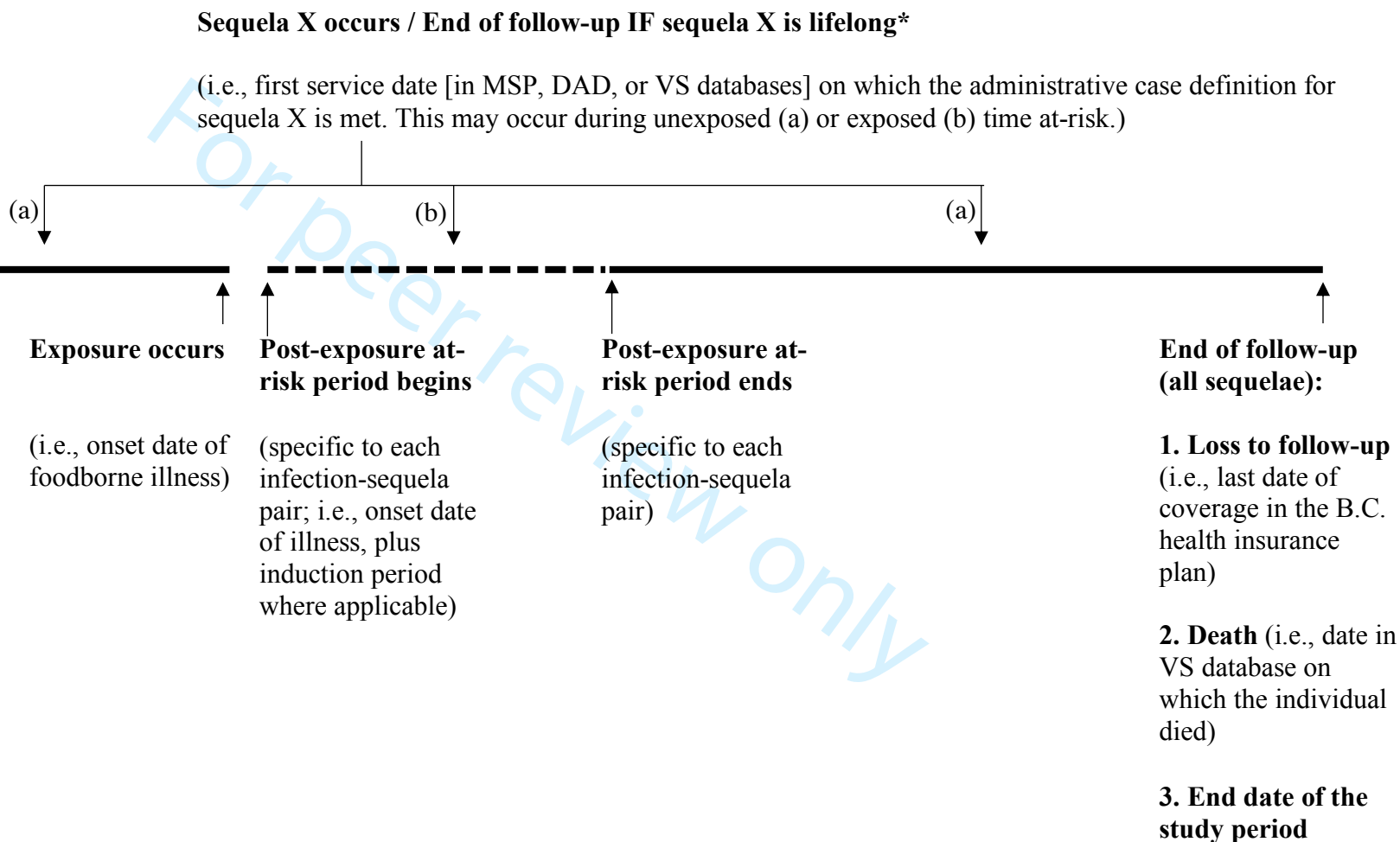
5 664
6

7
8 665 **Figure 1. Study follow-up period and time-at-risk for development of Sequela X (solid**
9
10 666 **lines: unexposed time at-risk; dashed lines: exposed time at-risk; B.C.,**
11
12 667 **British Columbia, Canada; MSP: Medical Services Plan; DAD: Discharge**
13
14 668 **Abstracts Database; VS: Vital Statistics)**
15
16
17 669

18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

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.