

2020

A geostatistical study of socioeconomic status (SES), rurality, seasonality and index test results as drivers of free private groundwater testing in southern Ontario, 2012–2016

Shahryar Qayyum

Paul Hynds

Harriet Richardson

See next page for additional authors

Follow this and additional works at: <https://arrow.tudublin.ie/ehsiart>



Part of the [Environmental Health and Protection Commons](#), [Environmental Public Health Commons](#), [Natural Resources and Conservation Commons](#), [Other Environmental Sciences Commons](#), and the [Water Resource Management Commons](#)

This Article is brought to you for free and open access by the ESHI Publications at ARROW@TU Dublin. It has been accepted for inclusion in Articles by an authorized administrator of ARROW@TU Dublin. For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie, gerard.connolly@tudublin.ie.



This work is licensed under a [Creative Commons Attribution-NonCommercial-Share Alike 4.0 License](#)
Funder: Canadian Foundation for Infectious Diseases (CFID)

Authors

Shahryar Qayyum, Paul Hynds, Harriet Richardson, Kevin McDermott, and Anna Majury



A geostatistical study of socioeconomic status (SES), rurality, seasonality and index test results as drivers of free private groundwater testing in southern Ontario, 2012–2016

Shahryar Qayyum ^a, Paul Hynds ^b, Harriet Richardson ^{a,c}, Kevin McDermott ^d, Anna Majury ^{d,e,*}

^a Department of Public Health Sciences, Queen's University, Kingston, Ontario, Canada

^b Environmental Health & Sustainability Institute, Technological University Dublin, Dublin, Ireland

^c Division of Cancer Care and Epidemiology, Cancer Research Institute at Queen's University, Kingston, Ontario, Canada

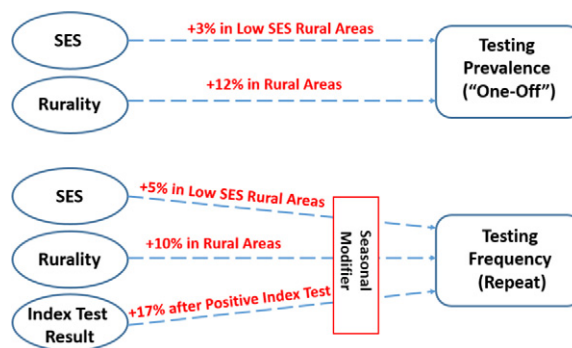
^d Public Health Ontario, Kingston, Ontario, Canada

^e Department of Biology and Molecular Sciences, Queen's University, Kingston, Ontario, Canada

HIGHLIGHTS

- Dataset comprising 417,406 private wells with a 27.5% test rate
- *E. coli* detection rate of 3.2% in categorically rural areas
- Wells located in lower socioeconomic regions more likely to be tested.
- Positive index test associated with increased likelihood of repeat testing.
- Results may be used to inform future groundwater awareness and testing campaigns.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 11 December 2019

Received in revised form 6 February 2020

Accepted 6 February 2020

Available online 7 February 2020

Editor: José Virgílio Cruz

Keywords:

Groundwater
E. coli
 Testing
 Behaviour
 Private Wells
 Socioeconomic status

ABSTRACT

Approximately 12% of the Canadian population uses private wells for daily water consumption; however, well water testing rates are on the decline, resulting in an increased risk of waterborne acute gastrointestinal illness. To date, limited research has explored the determinants influencing well testing practices. Accordingly, the current study sought to investigate the drivers of “one-off” and repeat well water testing in southern Ontario during the 5-year period 2012–2016, using the world's largest private groundwater testing data-frame. Data from >400,000 wells were geospatially integrated with all tests conducted by the provincial laboratory in southern Ontario. The Ontario Marginalization Index (ON-Marg) was used as a proxy measure of socioeconomic status (SES), with rurality, based on population density, season, and index (1st) test results assessed as effect modifiers. Multivariate analysis was undertaken using log-binomial regression. Approximately 27.5% of wells ($n = 417,406$) were tested during the study period, 66.7% of which were sampled more than once; 3% of all samples tested positive for *E. coli* (>0 colony forming unit/100 mL). In rural regions (<150 people/km²), wells located in low SES areas were 13% more likely to be tested compared to high SES areas (95% CI: 1.11, 1.15). In urban (>400 people/km²) and peri-urban regions (>150 and <400 people/km²), wells located in low SES areas were 14% (95% CI: 0.78, 0.95) and 15% (95% CI: 0.76, 0.94) less likely to be tested compared to high SES areas. Wells located in low SES areas were 6% more likely to be re-tested (95% CI: 1.04, 1.07). Positive index tests were associated

* Corresponding author at: 181 Barrie St., Kingston, ON K7L 3K2, Canada.
 E-mail address: anna.majury@oahpp.ca (A. Majury).

with a 17% increased likelihood of repeat testing (95% CI: 1.16, 1.18). Accordingly, the authors conclude that location and SES are significant predictors of well water testing, with index test status being the most influential predictor of repeat well testing.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

Approximately 7 million Canadian residents (19.4% of national population) rely on groundwater for daily consumption, of whom 62% (4.35 million) use a private well for domestic extraction (Murphy et al., 2016). Private water wells are those located on a private property, supplying water to a single household (Government of Ontario, 2017). While groundwater is afforded a level of natural protection (i.e. overlying subsoils) relative to surface water, it is not impervious to contamination (Charrois, 2010). Consumption of groundwater from microbially contaminated private wells represents a potential source of exposure to enteric pathogens, which may cause endemic and epidemic acute gastrointestinal illness (AGI) in both developed and developing countries (Murphy et al., 2017). For example, approximately 14.6% of all waterborne outbreaks in Canada between 1993 and 2007 were associated with contaminated private groundwater supplies, with this considered an underestimate of the true burden of illness given that a significant proportion of AGI cases go unreported (Moffat and Struck, 2011). AGI is one of the most frequently diagnosed illnesses in Canada with over 4 million cases (of both known and unknown aetiologies) being reported in 2006 (Government of Canada, 2013; Thomas et al., 2015).

Regular well water testing (based on local/regional guidance, hydrogeological setting, infrastructure, and climate) for faecal indicator organisms (FIOs) including *E. coli* is recommended to assess potability, thereby reducing the probability of consuming contaminated groundwater and subsequent AGI contraction (Kreutzwiser et al., 2011; Government of Canada, 2013; Thomas et al., 2015; Murphy et al., 2016; Murphy et al., 2017). The province of Ontario offers free testing of private water wells for FIOs, including *E. coli* (Maier et al., 2014). *E. coli* is a particularly useful microbial indicator, as the presence of just 1 colony forming unit (CFU) per 100 mL implies recent faecal ingress, thus indicating that the water may be unsafe for human consumption (An et al., 2002). However, despite the availability of free testing, previous studies have found that a significant proportion of well owners/users in Ontario continue to test their domestic water source infrequently or not at all (Kreutzwiser et al., 2011; Maier et al., 2014).

Many factors can influence a well owner's decision to test their well water, however, to date, studies examining these factors have primarily employed survey-based approaches and thus, may be spatio-temporally limited and based on relatively small sample numbers. Previous studies indicate that knowledge of the risks associated with the consumption of contaminated groundwater, knowledge of well testing, attitudes towards testing, rurality, weather, history of previous contamination/illness and complacency are the most important determinants of well water testing practices (Charrois, 2010; Kreutzwiser et al., 2011; Imgrund et al., 2011; Hynds et al., 2013).

Few studies have researched the relationship between socioeconomic status (SES) and well testing practices, although SES has previously been associated with groundwater quality (Safe Drinking Water Act compliance) from private wells in the United States (Switzer and Teodoro, 2017). For example, SES (both area and individual level) has been linked to both contamination risk in the area and knowledge of risks (Charrois, 2010; Kreutzwiser et al., 2011; Flanagan et al., 2015; Switzer and Teodoro, 2017; American Psychological Association, 2018). It has also been observed that higher SES can influence a well owner's risk perception and increase their general awareness of risks and testing procedures (Flanagan et al., 2015; Switzer and Teodoro, 2017). However, the evidence is not consistent, with Kreutzwiser et al. (2011) reporting an inverse relationship between SES and well water

testing procedures. It has been hypothesized that regions of the province with low area-level SES may reflect certain employment sectors, such as agriculture, and particularly in rural regions. Residents in these regions may have an inherent reliance on groundwater, and consequently may be more aware of the risks associated with not regularly testing well water (Charrois, 2010; Kreutzwiser et al., 2011; Imgrund et al., 2011).

Accordingly, the current study sought to investigate the association between SES and the prevalence of free well testing between 2012 and 2016 by the provincial laboratory, inclusive, in southern Ontario, while adjusting for the effect of rurality (Fig. 1). The association between SES and repeated well water testing was examined during the same study period (2012 and 2016), with rurality, season and index test status included as potential effect modifiers (Fig. 2).

2. Methods

2.1. Study population and data sources

The province of Ontario maintains a database of test results for private drinking water samples submitted to the provincial laboratory for bacteriological testing, herein referred to as the Well Testing Data Base (WTDB). Well water (200 mL) samples for submission to the provincial laboratory are collected using sample bottles (BioNuclear Diagnostics, Inc., Toronto, Canada) containing 0.8 mM sodium thiosulfate. Sample collection kits, including a water sample collection bottle and a drinking water testing requisition, are available free of charge at all provincial public health laboratory sites ($n = 11$) and each of the 35 local public health authority offices and their depots in Ontario. Private drinking water users may submit as many water samples as they wish, the only limitations being that the water must be from a private drinking water source and this source must be from within the province of Ontario. Once collected, the well users is required to return the water sample, and the completed water sample requisition, to a provincial public health laboratory or the municipal health unit of their choice. Samples not received directly by the laboratory are transferred in coolers, to a provincial public health laboratory, at approximately 2–8 °C and tested (via vacuum filtration and culture on Differential Coliform Agar) within 48 h of collection.

A second database (Well Water Information System (WWIS), <https://www.ontario.ca/page/well-records>), also maintained by the province, includes well location data, and is updated quarterly. The current study employed fields from both databases, including test dates and result(s) from 2012 to 2016, in addition to geocoded location.

Data from the 2006 update of the Ontario Marginalization Index (ON-Marg) and 2006 Canadian Census were employed for covariate development, including rurality and SES, with the population of interest being all private wells located in southern Ontario, equating to >80% of all private wells in Ontario (Fig. 3) (Statistics Canada, 2012).

The 2006 version of the ON-Marg was used to determine the SES of dissemination areas (DA); DAs are the smallest geographic area for which census data are disseminated and comprise a population of 400 to 700 people (Statistics Canada, 2015; Matheson and Ingen, 2017). The 2006 version of the ON-Marg was employed instead of the 2011 version due to the lack of a long-form census in 2011 (Green and Milligan, 2010). The 2006 National Canadian Census was also used to provide information on rurality, with this version of the census chosen to ensure consistency with the 2006 ON-Marg. The ON-Marg was another factor for limiting the study locale to southern Ontario, as the

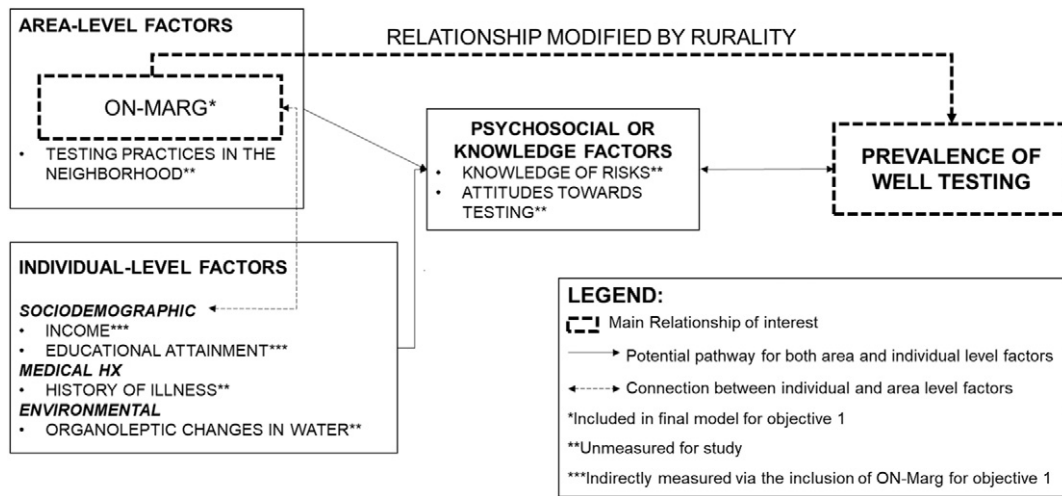


Fig. 1. A conceptual framework for prevalence of well testing with the relationship(s) of interest investigated in the current study highlighted (Hatched); “Testing practised in the neighbourhood” refers to the potential effect of local actions/behaviours (i.e. friends or neighbours) on individual/household actions/behaviours.

index is missing data for many northern DAs (Matheson and Ingen, 2017).

2.2. Dataset development

Based upon the WWIS database, in which well locations are designated geo-coordinates, the postal code for each well location in Ontario was determined using the Google Maps Reverse Geo-coding application program interface (API). Postal codes were required to determine the associated DAs (ON-Marg based on DAs), with reverse geo-coding undertaken using the Googleway package in R (Google Developers, 2017; Cooley, 2018; R Core Team, 2018).

Development of the dataset after reverse geo-coding is outlined in Fig. 4. Using Version 9.4 of the SAS System for Windows 10 (SAS Institute Inc., Cary, NC, USA), 60,202 wells were removed due to missing postal codes (10,111 tested wells and 50,091 untested wells). The remaining 442,202 wells had their postal codes converted to DA codes based on the 2006 Census using the Postal Code Conversion File Plus (PCCF+) (Statistics Canada, Postal Code Conversion File Plus (Version 6A1), 2017). This dataset was then merged with the ON-Marg dataset based on DA codes, with an additional 24,798 wells removed due to

their postal codes not matching any DA or the matched DA not being present in the ON-Marg dataset (790 Tested wells and 24,006 untested wells). Accordingly, the final dataset contained 417,406 wells.

2.3. E. coli and total coliform testing

Groundwater samples submitted to the province of Ontario are tested for *E. coli* and total coliforms (TC), using the modified culture-based membrane filtration method, MECP E3407, approved by the Ontario Ministry of Environment, Conservation and Parks (MECP) (Krolik et al., 2014). Test results are reported to the owner of well and recorded in the provincial database (WTDB), along with well location and test date (Public Health Ontario, 2013; Public Health Ontario, 2017a, 2017b).

2.4. Covariates

2.4.1. Creating a unified measure of socioeconomic status

Several methods exist to determine household SES, with family income being the most frequently employed. The approach to determining SES typically depends on the research question, analytical

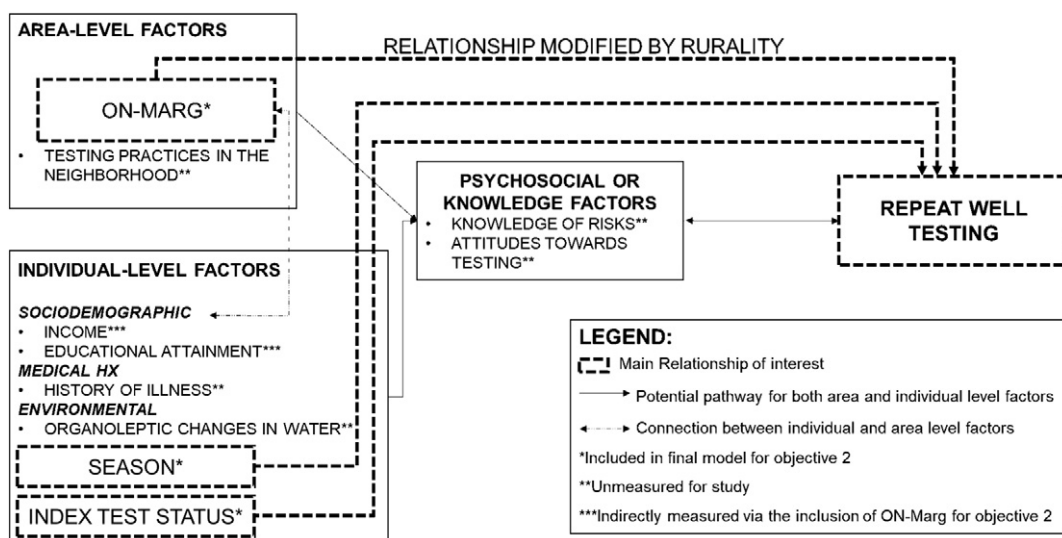


Fig. 2. A conceptual framework for repeat testing with the relationship(s) of interest investigated in the current study highlighted (Hatched); “Testing practised in the neighbourhood” refers to the potential effect of local actions/behaviours (i.e. friends or neighbours) on individual/household actions/behaviours.

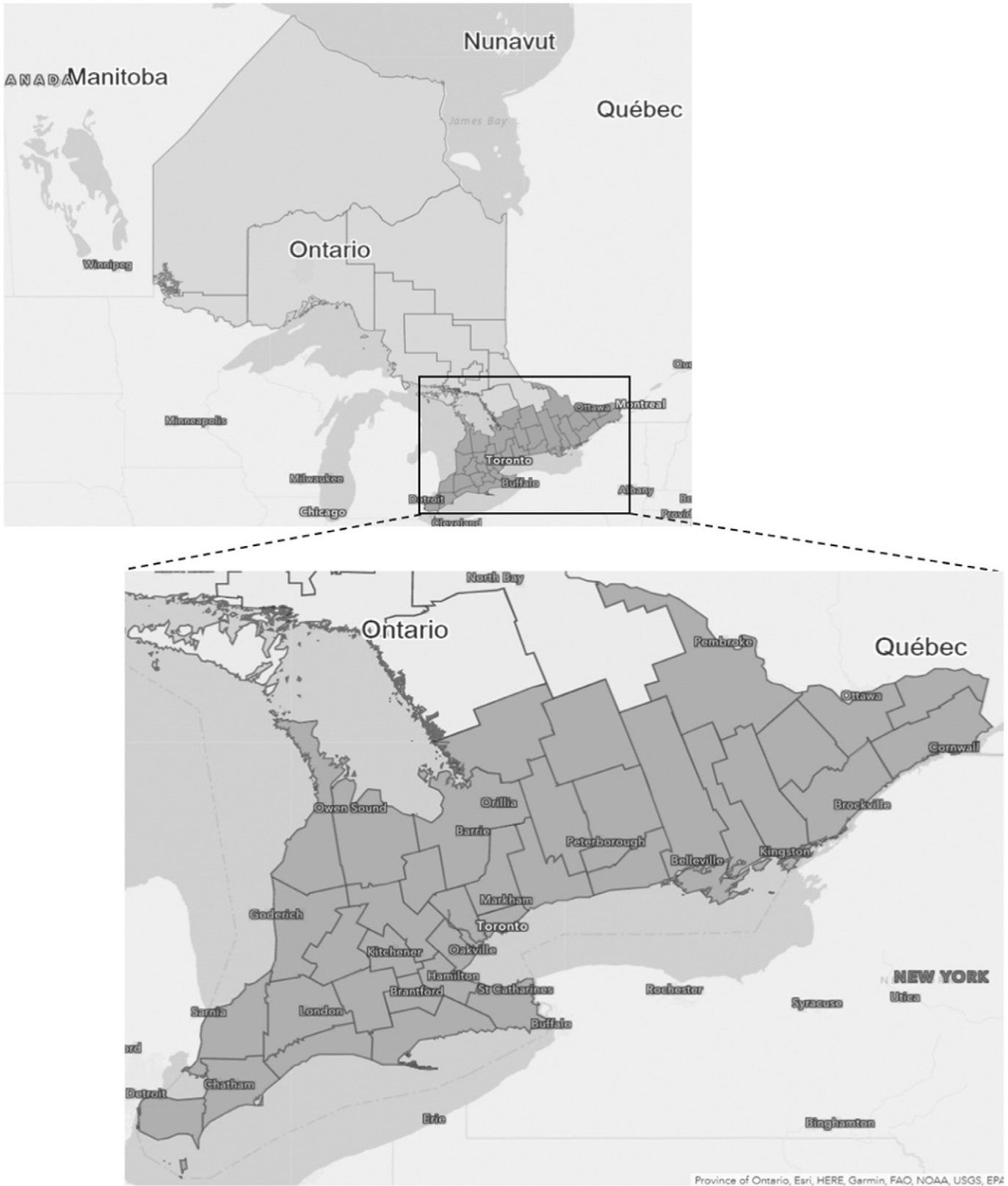


Fig. 3. Map of Ontario with the study area (southern Ontario) highlighted in grey.

approach and data availability (Pampalon and Raymond, 2000; Pampalon et al., 2010; Berzofsky et al., 2014).

In Canada, several SES measures are used by researchers and government organizations, with one of the most frequently employed being the Canadian Marginalization Index (CAN-Marg), an area-based

measure of SES that uses data from the Canadian Census and the Canadian Community Health Survey (Matheson et al., 2012a, 2012b; Dunn et al., 2012). The Ontario Marginalization Index (ON-Marg) was developed from a subset of the CAN-Marg dataset, both of which are composite SES measures (Matheson et al., 2012a, 2012b).

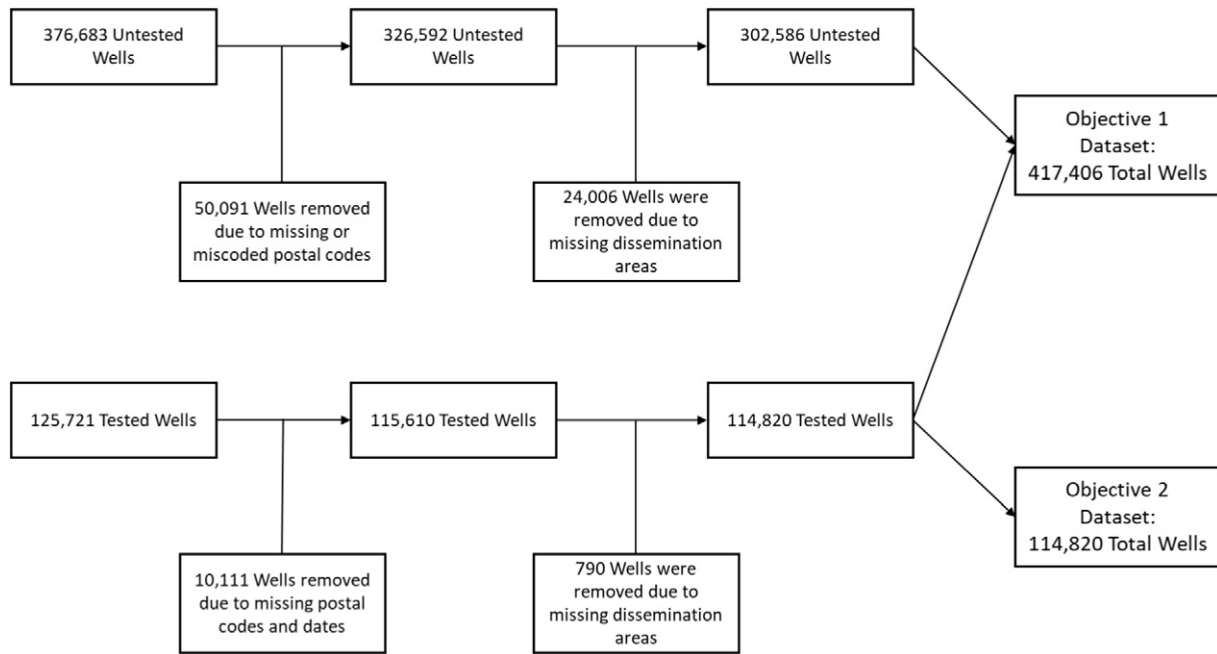


Fig. 4. Dataset Development Process for both study objectives. This process occurs after the reverse geo-coding for both provincial datasets.

The ON-Marg comprises eighteen indicators categorized into four dimensions; namely residential instability, material deprivation, dependency, and ethnic concentration (Table 1) and was chosen over other potential area-level SES measures as these indicators represent many of the predictors associated with well testing practices and thus may best represent the effect of area-level SES on these behaviours (Figs. 1 and 2).

Factor scores are computed for all indicators and subsequently discretized into quintiles and assigned a final discrete score from 1 (low marginalization [high SES]) to 5 (high marginalization [low SES]) for each dimension. Dimension scores were added to create a summary score for each DA, which was non-parametrically distributed and remained as such after several transformation approaches were employed. Thus, summary score was categorized based on quartiles (Table 2), where level 1 represents low marginalization (high SES) and level 4 represents high marginalization (low SES) and was investigated as a categorical indicator of SES in the data analysis.

2.4.2. Rurality

Longer distances to sample drop-off centres (primarily occurring in low population density, rural areas) have been found to reduce the likelihood of well testing, such that location alters the relationship between other predictors, and specifically those associated with location such as SES (Kreutzwiser et al., 2011; Imgrund et al., 2011). Thus, rurality was evaluated as a potential effect modifier of the relationship between SES and both the prevalence (well test status) and frequency (repeat test status) of well testing. The 2006 Canadian Census was used to provide the population and spatial extent of each DA. Based on this, DA population density was calculated and ranged from 0.49 people/km² to 34,058 people/km², with a mean DA population density of 228.27 people/km².

Statistics Canada and the Organization of Economic Co-operation and Development (OECD) both classify rurality based on population density. Statistics Canada uses 400 people/km² as the cut-off for rurality while the OECD uses 150 people/km² (OECD, 2011; Statistics Canada, 2017a, 2017b). To ensure both definitions were considered, three categories for rurality based on population density were

Table 1
Indicators used to create the four dimensions of the Ontario Marginalization Index, adapted from Matheson et al. (2012a, 2012b).

Dimensions	Indicators						
Residential instability	Proportion of the population living alone	Proportion of the population who are 16+	Mean number of persons per dwelling	Proportion of dwellings that are apartment buildings	Proportion of the population who are single/divorced/widowed	Proportion of dwellings that are not owned	Proportion of the population who moved during the past 5 years
Material deprivation	Proportion of the population aged 20+ without a high-school diploma	Proportion of lone parent families	Proportion of the population receiving government transfer payments	Proportion of the population aged 15+ who are unemployed	Proportion of the population considered low income	Proportion of dwellings requiring major repair	
Dependency	Proportion of the population who are 65+	Dependency ratio (total population 0–14 and 65+ /total population 15 to 64)	Proportion of the population not participating in labour force (aged 15+)				
Ethnic concentration	Proportion of the population who are recent immigrants (arrived in the 5 years prior to census)	Proportion of the population who self-identify as a visible minority					

Table 2
Descriptive summary of the potential predictors of (“ever”) well testing during the study period (N = 417,406 Wells).

Variable	Total, n (%)	Testing status, n (%)	
		Tested	Untested
SES			
High SES	130,586 (31.29)	36,151 (8.66)	94,435 (22.62)
High/medium SES	135,104 (32.37)	35,988 (8.62)	99,116 (23.75)
Low/medium SES	91,929 (22.02)	25,662 (6.15)	66,267 (15.88)
Low SES	59,787 (14.32)	17,019 (4.08)	42,768 (10.25)
Rural status			
Rural	335,536 (80.39)	95,459 (22.87)	240,077 (57.52)
Peri-urban	31,421 (7.53)	8223 (1.97)	23,198 (5.56)
Urban	50,449 (12.09)	11,138 (2.67)	39,311 (9.42)
Dependency^a			
1 (least)	28,005 (6.71)	6935 (1.66)	21,070 (5.05)
2	87,066 (20.86)	23,101 (5.53)	63,965 (15.32)
3	114,223 (27.36)	30,620 (7.34)	83,603 (20.03)
4	103,109 (24.70)	28,236 (6.76)	74,873 (17.94)
5 (most)	85,003 (20.36)	25,928 (6.21)	59,075 (14.15)
Residential instability^a			
1 (least)	132,647 (31.78)	36,267 (8.69)	96,380 (23.09)
2	154,342 (36.98)	41,263 (9.89)	113,079 (27.09)
3	95,959 (22.99)	27,726 (6.64)	68,233 (16.35)
4	30,559 (7.32)	8459 (2.03)	22,100 (5.29)
5 (most)	3899 (0.93)	1105 (0.26)	2794 (0.67)
Material deprivation^a			
1 (least)	100,802 (24.15)	27,581 (6.61)	73,221 (17.54)
2	127,646 (30.58)	34,381 (8.24)	93,265 (22.34)
3	107,045 (25.65)	29,467 (7.06)	77,578 (18.59)
4	61,823 (14.81)	17,509 (4.19)	44,314 (10.62)
5 (most)	20,090 (4.81)	5882 (1.41)	14,208 (3.40)
Ethnic concentration^a			
1 (least)	150,547 (36.07)	43,787 (10.49)	106,760 (25.58)
2	147,199 (35.27)	40,721 (9.76)	106,478 (25.51)
3	83,831 (20.08)	22,485 (5.39)	61,346 (14.70)
4	28,109 (6.73)	6839 (1.64)	21,270 (5.10)
5 (most)	7720 (1.85)	988 (0.24)	6732 (1.61)

Note: SES, socioeconomic status; CI, confidence interval.

^a 1 represents the least marginalized or highest SES areas, whereas 5 represents the most marginalized or lowest SES areas.

created, as follows: Rural DA: <150 people/km²; Peri-urban DA: ≥150 people/km² to <400 people/km²; Urban DA: ≥400 people/km². The peri-urban category was created to represent mid-density areas such as small towns and large villages. Based upon this classification of rurality, a significant majority (80.4%) of private water wells are located in rural DAs, 7.5% are in peri-urban DAs and 12.1% are in urban DAs (Table 2).

Additionally, the Rurality Index of Ontario (RIO), developed by the Ontario Medical Association, was examined as a potential measure of rurality. RIO measures access to healthcare by accounting for the effects of population density and travel time to basic/advanced healthcare centres (Kralj, 2009). This measure was developed to account for another facet of rurality, specifically as a proxy measure of testing service accessibility, which has been observed to be a barrier to well water testing (Charrois, 2010).

2.4.3. Season

Based upon the test date recorded in the provincial database, a season variable was created using the northern meteorological definition of seasons as follows: Winter (December to February), Spring (March to May), Summer (June to August) and Fall (September to November) (National Oceanic and Atmospheric Administration, 2017).

2.4.4. Index test status

Previous studies have shown that well owners are more likely to test their well if it has a history of contamination (Kreutzwisser et al.,

2011). Thus, if the first test conducted for a well water sample is positive (>0 CFU for *E. coli*), then it is more likely to be tested subsequently (Kreutzwisser et al., 2011). Index test status (first test for a well in the dataset) was determined by combining the test results for *E. coli* and total coliforms in the provincial datasets. A well was assigned a positive test status if it had >0 CFU per 100 mL for either *E. coli* or total coliforms and was assigned a negative test status if it had 0 CFUs per 100 mL for both indicators. This definition was chosen, as the authors believe that many well owners may not know the difference between *E. coli* and total coliforms and thus may keep testing until both indicators are zero. It is important to note that, the presence of an index test, as defined by the authors, is bounded by the start- (January 1st 2012) and end-dates (December 31st 2016) of the employed dataset. As such, the index test variable used for analyses does not account for any samples sent for analyses prior to January 1st 2012, and thus it is entirely reasonable to assume that samples may be tested outside the study period.

Table 3
Descriptive summary of the potential predictors of “repeated” well testing (N = 114,820 Wells).

Variable	Total, n (%)	Repeat testing status, n (%)	
		Multiple tests	1 test only
SES			
High SES	36,151 (31.48)	23,744 (20.68)	12,407 (10.81)
High/medium SES	35,988 (31.34)	24,156 (21.04)	11,832 (10.30)
Low/medium SES	25,662 (22.35)	17,161 (14.95)	8501 (7.40)
Low SES	17,019 (14.82)	11,543 (10.05)	5476 (4.77)
Rural status			
Rural	95,459 (83.14)	64,376 (56.07)	31,083 (27.07)
Peri-urban	8223 (7.16)	5447 (4.74)	2776 (2.42)
Urban	11,138 (9.70)	6781 (5.91)	4357 (3.79)
Season			
Winter	20,378 (17.75)	14,074 (12.26)	6304 (5.49)
Spring	36,177 (31.51)	25,502 (22.21)	10,675 (9.30)
Summer	34,848 (30.35)	23,115 (20.13)	11,733 (10.22)
Fall	23,417 (20.39)	13,913 (12.12)	9504 (8.28)
Index test status			
Negative	82,690 (72.02)	52,689 (45.89)	30,001 (26.13)
Positive	32,130 (27.98)	23,915 (20.83)	8215 (7.15)
Dependency^a			
1 (least)	6935 (6.04)	4400 (3.83)	2535 (2.21)
2	23,101 (20.12)	15,041 (13.10)	8060 (7.02)
3	30,620 (26.67)	20,234 (17.62)	10,386 (9.05)
4	28,236 (24.59)	18,870 (16.43)	9366 (8.16)
5 (most)	25,928 (22.58)	18,059 (15.73)	7869 (6.85)
Residential instability^a			
1 (least)	36,267 (31.59)	23,849 (20.77)	12,418 (10.82)
2	41,263 (35.94)	27,368 (23.84)	13,895 (12.10)
3	27,726 (24.15)	19,066 (16.61)	8660 (7.54)
4	8459 (7.37)	5671 (4.94)	2788 (2.43)
5 (most)	1105 (0.96)	650 (0.56)	455 (0.40)
Material deprivation^a			
1 (least)	27,581 (24.02)	18,214 (15.86)	9367 (8.16)
2	34,381 (29.94)	23,017 (20.05)	11,364 (9.90)
3	29,467 (25.66)	19,804 (17.25)	9663 (8.42)
4	17,509 (15.25)	11,718 (10.21)	5791 (5.04)
5 (most)	5882 (5.12)	3851 (3.35)	2031 (1.77)
Ethnic concentration^a			
1 (least)	43,787 (38.14)	29,924 (26.06)	13,863 (12.07)
2	40,721 (35.47)	26,880 (23.41)	13,841 (12.05)
3	22,485 (19.58)	14,871 (12.95)	7614 (6.63)
4	6839 (5.96)	4392 (3.83)	2447 (2.13)
5 (most)	988 (0.86)	537 (0.47)	451 (0.39)

Note: SES, socioeconomic status; CI, confidence interval.

^a 1 represents the least marginalized or highest SES areas, whereas 5 represents the most marginalized or lowest SES areas.

2.5. Statistical analysis

Statistical analyses were performed using SAS 9.4 for Windows 10 (SAS Institute Inc., Cary, NC, USA). Standard descriptive statistics were computed for all categorical variables (e.g. rurality). Spearman's Rho was employed to assess the presence of (multi) collinearity, while chi-square tests and crude relative risks (calculated using log-binomial regression) were used to assess the presence of bivariate relationships between predictors and outcomes. The outcome of interest (dependent variable) for Objective 1 was the occurrence of a well test (well test status [dichotomous: yes/no]), while the outcome of interest for Objective 2 was the occurrence of ≥ 1 well test (repeat test status [dichotomous: yes/no]).

Most predictors exhibited very weak ($R_{sp} < \pm 0.2$) correlations with each other, however, as expected, health access and rurality were significantly correlated ($R_{sp} = 0.76, p < 0.0001$). Accordingly, these variables were not concurrently included in individual multivariate models. Rurality based on population density was chosen as the preferred method of rural classification, as multivariate models comprising rurality exhibited narrower confidence intervals for point estimates, compared to health access.

Log-binomial regression was employed for multivariate analyses as the outcomes of interest were shown to be statistically "common" ($>10\%$ prevalence) and thus calculated relative risks are considered more accurate than odds ratios (e.g. logistic regression), due to a lack of bias away from the null hypothesis (Cummins, 2009).

Season and index test status were only included as model covariates for repeat testing (Objective 2) as these are only valid in the presence of

≥ 1 test during the study period. All independent variables had significant bivariate relationships with the outcomes and therefore no predictors were removed prior to multivariate analysis.

Likelihood ratio tests ($p < 0.05$) were used to determine if rurality was an effect modifier of the relationship between SES and both primary outcomes (1. Well Test Status, and 2. Repeat Test Status). Rurality was not assessed as a potential effect modifier for the relationship between index test status, season and the primary outcomes, as they do not vary by location.

For well test status, the categorized combined total score of the ON-Marg (SES), stratified by rurality was the only predictor of interest. For repeat test status, the categorized combined total score of the ON-Marg (SES), season during which the index test was conducted, and the index test result were included as predictors of interest.

Log-binomial regression does not have any established model fit or diagnostic tests, thus tests established for logistic or Poisson regression to assess model fit and diagnostics are typically employed, such as the Hosmer-Lemeshow test (Blizzard and Hosmer, 2006). All variables in presented models were significant at the 0.05 significance level, with the Hosmer-Lemeshow goodness of fit test used to assess model fit.

3. Results

3.1. Descriptive analysis

Overall, the analysed dataset comprised 512,733 individual private well samples received and processed by the provincial laboratory during

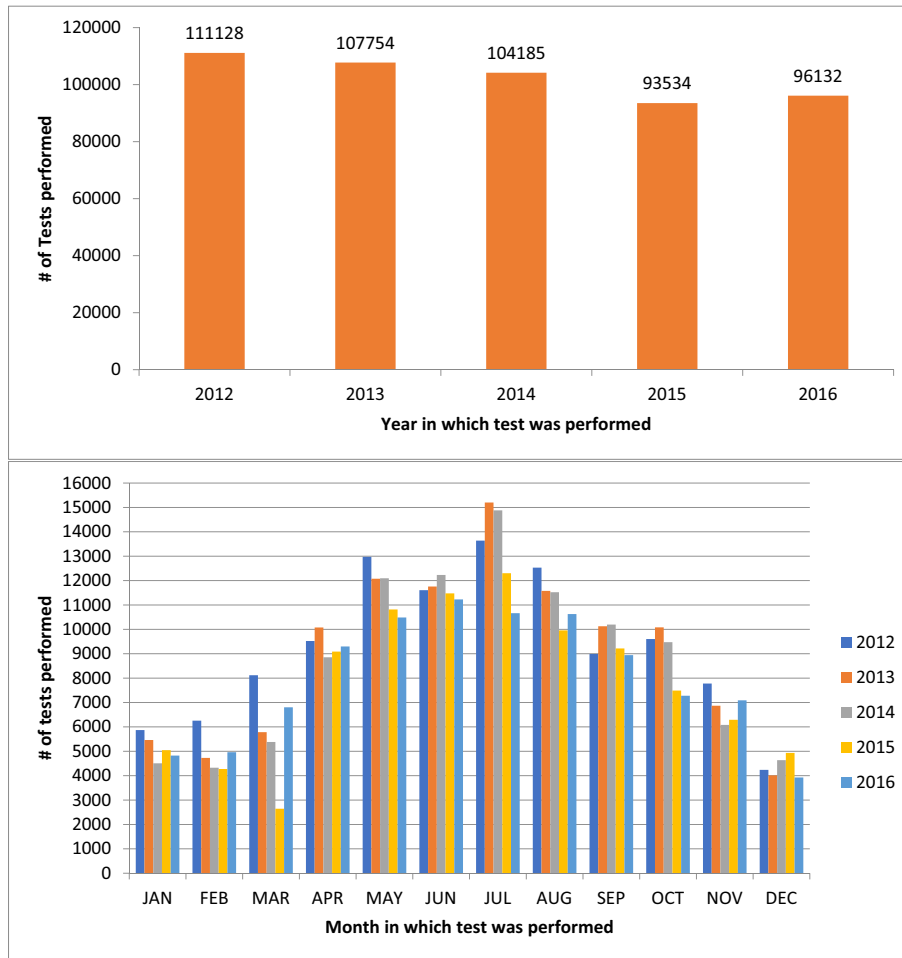


Fig. 5. All bacteriological tests included in the dataset that were performed by the province of Ontario between January 1st, 2012 and December 31st, 2016, inclusive. Displayed on both a year by year (Panel a) and month by month (Panel b) basis.

the 5-year period 2012–2016. Of 114,820 unique geo-referenced groundwater supplies, 33.3% ($n = 38,216$) were associated with one submitted sample, 18.2% ($n = 20,923$) were associated with two samples (i.e. index test + 1), with the remaining 48.5% ($n = 55,681$) temporally referenced to >2 samples (i.e. index test + >1). The largest sample number attributed to one well over the study period was 290 (Mean > 1 sample/week).

Descriptive summaries of each potential predictor included for analyses are presented in Tables 2 and 3. As shown (Table 2), 80.4% of all wells are located in rural DAs, while 63.7% of all wells are located in high and high/medium SES DAs. Tested wells were typically located in rural DAs (83.1%), with 31.5% and 31.3% of tested wells located in high and high/medium SES DAs, respectively. A majority of index tests (72.0%) were negative for both *E. coli* and total coliforms (0 CFU per 100 mL), with 31.5% of index tests undertaken during spring compared to 17.8% in winter. Mean *E. coli* detection rates (including both “one-off” samples and index tests among repeat sampled wells) of 3.25%, 2.32% and 2.86% were found among rural, peri-urban and urban wells, respectively. As shown (Fig. 5a), sample submissions declined year on year between 2012 and 2015, with a slight increase noted in 2016. Testing was most frequent during spring and summer and least frequent during winter (Fig. 5b), with highest *E. coli* detection rates encountered during summer months (Fig. 6).

3.2. Predictors of well water testing

As shown (Table 4), all variables of interest were significantly ($p < 0.05$) associated with well test status. Wells located in urban and peri-urban DAs were 28% and 9% less likely to be tested compared to those located in rural DAs, respectively (crude RR (urban) = 0.78; 95% CI: 0.76, 0.79 and crude RR (peri-urban) = 0.92; 95% CI: 0.90, 0.94). Wells located in low SES DAs were slightly (3%) more likely to be tested than wells located in high SES DAs (Crude RR = 1.03; 95% CI: 1.01, 1.04).

The results of the well test status (prevalence) model are presented in Table 6. The Hosmer–Lemeshow goodness of fit test indicates that this model is a good fit ($p = 0.84$), with rurality found to be an effect modifier of the relationship between SES and well testing status ($p < 0.0001$). As shown (Table 6), within rural areas, wells located in low SES DAs were 13% more likely to be tested than wells located in high SES DAs (RR = 1.13; 95% CI: 1.11, 1.15). Conversely, in peri-urban and urban areas, wells located in low SES DAs were approximately 15% and 14% less likely to be tested compared to wells in high SES DAs (RR = 0.85; 95% CI: 0.76, 0.94 and RR = 0.86; 95% CI: 0.78, 0.95), respectively.

Table 4

Bivariate analysis of associations between predictors of groundwater testing and test status, using log-binomial regression (N = 417,406 Wells).

Variable	Total, n (%)	Crude relative risk (95% CI)	Chi-square value (p-value)
SES			
High SES	130,586 (31.29)	1.00 (reference)	88.55
High/medium SES	135,104 (32.37)	0.96 (0.95, 0.97)	(<0.0001)
Low/medium SES	91,929 (22.02)	1.01 (0.99, 1.02)	
Low SES	59,787 (14.32)	1.03 (1.01, 1.04)	
Rural status			
Rural	335,536 (80.39)	1.00 (reference)	923.42
Peri-urban	31,421 (7.53)	0.92 (0.90, 0.94)	(<0.0001)
Urban	50,449 (12.09)	0.78 (0.76, 0.79)	
Dependency^a			
1 (least)	28,005 (6.71)	1.00 (reference)	558.46
2	87,066 (20.86)	1.07 (1.05, 1.10)	(<0.0001)
3	114,223 (27.36)	1.08 (1.06, 1.11)	
4	103,109 (24.70)	1.11 (1.06, 1.13)	
5 (most)	85,003 (20.36)	1.23 (1.20, 1.26)	
Residential instability^a			
1 (least)	132,647 (31.78)	1.00 (reference)	142.33
2	154,342 (36.98)	0.98 (0.97, 0.99)	(<0.0001)
3	95,959 (22.99)	1.06 (1.04, 1.07)	
4	30,559 (7.32)	1.01 (0.99, 1.03)	
5 (most)	3899 (0.93)	1.04 (0.99, 1.09)	
Material deprivation^a			
1 (least)	100,802 (24.15)	1.00 (reference)	74.22
2	127,646 (30.58)	0.98 (0.97, 1.00)	(<0.0001)
3	107,045 (25.65)	1.01 (0.99, 1.02)	
4	61,823 (14.81)	1.04 (1.02, 1.05)	
5 (most)	20,090 (4.81)	1.07 (1.05, 1.10)	
Ethnic concentration^a			
1 (least)	150,547 (36.07)	1.00 (reference)	1189.46
2	147,199 (35.27)	0.95 (0.94, 0.96)	(<0.0001)
3	83,831 (20.08)	0.92 (0.91, 0.93)	
4	28,109 (6.73)	0.84 (0.82, 0.86)	
5 (most)	7720 (1.85)	0.44 (0.42, 0.47)	

Note: SES, socioeconomic status; CI, confidence interval.

^a 1 represents the least marginalized or highest SES areas, whereas 5 represents the most marginalized or lowest SES areas.

3.3. Predictors of repeat well water testing

For repeat test status (Table 5), all variables of interest were significantly associated with repeat test status (all p -values < 0.05). Wells characterised by a positive index test were 17% more likely

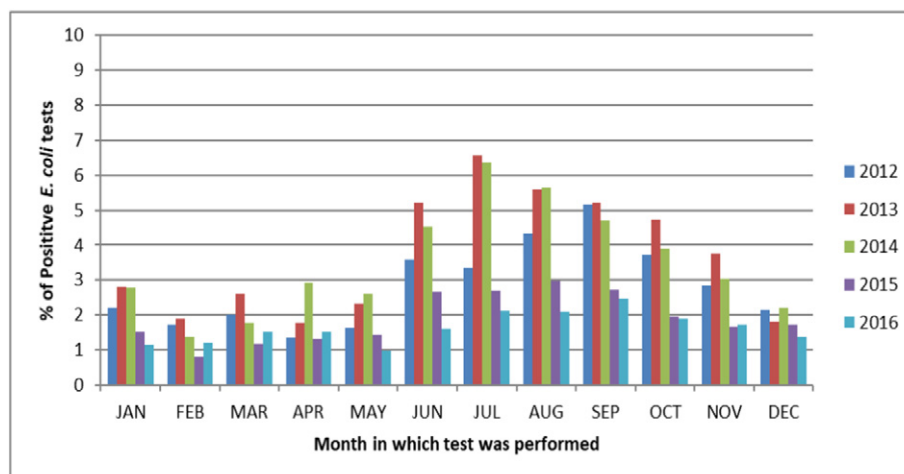


Fig. 6. Percentage of tests that were positive for *E. coli* presence (>1 CFU per 100 mL), for all tests performed by the province of Ontario between January 1st, 2012 and December 31st, 2016, displayed on a month by month basis.

to be re-tested (RR = 1.17; 95% CI: 1.16, 1.18). Wells tested for the first time (relative to study period start-date) during the Fall were 14% less likely to be tested again compared to wells tested for the first time (relative to study period start-date) in winter (RR = 0.86; 95% CI: 0.85, 0.87).

The results of the repeat well test (Retest/No Retest) model are presented in Table 6, with the Hosmer-Lemeshow goodness of fit test for this model indicating that it was a good fit ($p = 0.78$). Rurality was found to be an effect modifier of the relationship between SES and repeat well testing ($p = 0.0015$). As shown (Table 6), within rural areas, wells located in low SES DAs were 6% more likely to have a repeat test than wells located in high SES DAs (RR = 1.06; 95% CI: 1.04, 1.07). Conversely, in peri-urban and urban areas, wells located in low SES DAs were approximately 4% and 3% less likely to have a repeat test compared to wells in high SES DAs (RR = 0.96; 95% CI: 0.86, 1.07 and RR = 0.97; 95% CI: 0.88, 1.07, respectively), although these estimates were not statistically significant. The relationship between season, index test and repeated testing did not vary considerably by rurality and effect estimates were of similar magnitude to calculated crude relative risks and in the same direction.

4. Discussion

There are >500,000 private water wells in Ontario currently used for domestic purposes, serving an estimated 1.5 million Ontarians as their primary source of drinking water (Public Health Ontario, 2017a, 2017b). Well water testing is one of the most effective ways to ensure that contaminated groundwater is not consumed by well owners and their families (Charrois, 2010; Kreutzwiser et al., 2011). This study aimed to investigate the relationship between SES, index test status, rurality, season and well testing practices in southern Ontario using a large integrated dataset. To the authors knowledge, this is the first study to explore these relationships at this scale (>400,000 wells, >500,000 groundwater samples) and primarily focus on the relationship between SES and both the prevalence and frequency of testing while accounting for the effect of rurality and other potentially significant predictors.

Residential instability, material deprivation and dependency (Table 4) exhibited similar relationships (magnitude, directionality) to both outcomes (testing prevalence and frequency), with higher marginalization shown to concur with a greater rate of testing. Conversely, the opposite relationship was observed for ethnic concentration, where

Table 5

Bivariate analysis of associations between predictors of groundwater testing and repeat test status, using log-binomial regression (N = 114,820 Wells).

Variable	Total, n (%)	Crude relative risk (95% CI)	Chi-square value (p-value)	Adjusted relative risk (95% CI) ^a
SES				
High SES	36,151 (31.48)	1.00 (reference)	29.85	1.00 (reference)
High/medium SES	35,988 (31.34)	1.02 (1.01, 1.03)	(<0.0001)	1.02 (1.01, 1.03)
Low/medium SES	25,662 (22.35)	1.02 (1.01, 1.03)		1.02 (1.01, 1.03)
Low SES	17,019 (14.82)	1.03 (1.02, 1.05)		1.05 (1.03, 1.06)
Rural status				
Rural	95,459 (83.14)	1.00 (reference)	194.01	1.00 (reference)
Peri-urban	8223 (7.16)	0.98 (0.97, 0.99)	(<0.0001)	0.98 (0.97, 0.99)
Urban	11,138 (9.70)	0.90 (0.89, 0.92)		0.90 (0.89, 0.92)
Season				
Winter	20,378 (17.75)	1.00 (reference)	847.55	1.00 (reference)
Spring	36,177 (31.51)	1.02 (1.01, 1.03)	(<0.0001)	1.02 (1.01, 1.03)
Summer	34,848 (30.35)	0.96 (0.95, 0.97)		0.95 (0.94, 0.96)
Fall	23,417 (20.39)	0.86 (0.85, 0.87)		0.86 (0.85, 0.87)
Index test status				
Negative	82,690 (72.02)	1.00 (reference)	1195.99	1.00 (reference)
Positive	32,130 (27.98)	1.17 (1.16, 1.18)	(<0.0001)	1.17 (1.16, 1.18)
Dependency^b				
1 (least)	6935 (6.04)	1.00 (reference)	166.51	
2	23,101 (20.12)	1.03 (1.01, 1.05)	(<0.0001)	
3	30,620 (26.67)	1.04 (1.02, 1.06)		
4	28,236 (24.59)	1.05 (1.03, 1.07)		
5 (most)	25,928 (22.58)	1.10 (1.08, 1.12)		
Residential instability^b				
1 (least)	36,267 (31.59)	1.00 (reference)	101.63	
2	41,263 (35.94)	1.01 (1.00, 1.02)	(<0.0001)	
3	27,726 (24.15)	1.05 (1.03, 1.06)		
4	8459 (7.37)	1.02 (1.00, 1.04)		
5 (most)	1105 (0.96)	0.89 (0.85, 0.94)		
Material deprivation^b				
1 (least)	27,581 (24.02)	1.00 (reference)	14.19	
2	34,381 (29.94)	1.01 (1.00, 1.03)	(0.0067)	
3	29,467 (25.66)	1.02 (1.01, 1.03)		
4	17,509 (15.25)	1.01 (1.00, 1.03)		
5 (most)	5882 (5.12)	0.99 (0.97, 1.01)		
Ethnic concentration^a				
1 (least)	43,787 (38.14)	1.00 (reference)	151.73	
2	40,721 (35.47)	0.97 (0.96, 0.98)	(<0.0001)	
3	22,485 (19.58)	0.97 (0.96, 0.98)		
4	6839 (5.96)	0.94 (0.92, 0.96)		
5 (most)	988 (0.86)	0.80 (0.75, 0.84)		

Note: SES, socioeconomic status; CI, confidence interval.

^a The adjusted log-binomial model contains the following variables with no interaction terms (SES, Index Test Status, Season and Rurality).

^b 1 represents the least marginalized or highest SES areas, whereas 5 represents the most marginalized or lowest SES areas.

Table 6
Association between SES and a) groundwater testing and b) repeat testing, stratified by rural status.

Variable	Total, n (%)	Relative risk (95% CI) ^a					
		n (%)	Rural	n (%)	Peri-urban	n (%)	Urban
SES^a							
High SES	130,586 (31.29)	111,637 (26.75)	1.00 (reference)	8156 (1.95)	1.00 (reference)	10,793 (2.59)	1.00 (reference)
High/medium SES	135,104 (32.37)	113,598 (27.22)	1.04 (1.03, 1.06)	10,141 (2.43)	1.03 (0.93, 1.15)	11,365 (2.72)	0.70 (0.63, 0.77)
Low/medium SES	91,929 (22.02)	74,132 (17.76)	0.99 (0.97, 1.00)	6491 (1.56)	0.87 (0.79, 0.96)	11,306 (2.71)	0.70 (0.63, 0.77)
Low SES	59,787 (14.32)	36,169 (8.67)	1.13 (1.11, 1.15)	6633 (1.59)	0.85 (0.76, 0.94)	16,985 (4.07)	0.86 (0.78, 0.95)
SES^b							
High SES	36,151 (31.48)	31,169 (27.15)	1.00 (reference)	2331 (2.03)	1.00 (reference)	2651 (2.31)	1.00 (reference)
High/medium SES	35,988 (31.34)	31,311 (27.27)	1.02 (1.01, 1.03)	2456 (2.14)	0.99 (0.90, 1.08)	2221 (1.93)	0.92 (0.84, 1.02)
Low/medium SES	25,662 (22.35)	21,595 (18.81)	1.02 (1.01, 1.04)	1872 (1.63)	0.98 (0.89, 1.08)	2195 (1.91)	0.93 (0.84, 1.03)
Low SES	17,019 (14.82)	11,384 (9.91)	1.06 (1.04, 1.07)	1564 (1.36)	0.96 (0.86, 1.07)	4071 (3.55)	0.97 (0.88, 1.07)

Note: SES, socioeconomic status; CI, confidence interval.

The Relative Risk was obtained by running a multi-variable log-binomial regression model, with Rural Status as an interaction term.

^a Multi-variable model for Objective 1 (Well Test Status). Model contains only SES, stratified by Rural Status (N = 417,406 Wells).

^b Multi-variable model for Objective 2 (Repeat Test Status). Model contains SES (stratified by Rural Status), index test status and season (N = 114,820 Wells).

higher marginalization equated to significantly lower rates of testing. Despite these differences in associations, the dimensions were combined to create a unified proxy measure of SES, which was significantly associated with the prevalence of well water testing in southern Ontario.

In rural areas, wells located in lower SES areas were more likely to be tested compared to wells located in higher SES areas (Table 6), with the opposite being observed in urban and peri-urban regions. Previous studies indicate that contaminated groundwater sources (including microbial contamination) are predominantly located in rural regions, primarily due to land-use (e.g. increased prevalence of contaminant sources including agriculture, septic tanks, etc.) and infrastructural differences (i.e. engineered municipal systems versus rural sources) between regions (Hynds et al., 2014). Findings from the current study would seem to support previous findings i.e. wells located in rural regions were found to exhibit a higher mean detection rate (3.25%) than those in peri-urban (2.32%) and urban regions (2.86%). This may account for a proportion of increased testing rates exhibited across rural areas (i.e. as a precursor to positive index test rates and/or increased levels of experience (e.g. incidence of household AGI) leading to increased levels of risk perception). Rural regions are often classified as low SES areas due to lower income, educational attainment and higher proportion of dependent populations (Corkal et al., 2004; Nahar et al., 2008). Studies on personal motivations for well water testing have shown that many well owners test their well due to changes in the organoleptic characteristics of their water or the occurrence of gastrointestinal symptoms among household members and/or neighbours (Imgrund et al., 2011; Malecki et al., 2017). Therefore, well owners in low SES rural regions may be more motivated to test due to the higher risk of contamination.

Additionally, the greater proportion of farmers in rural areas may have caused the marked observed differences between regions. Low SES rural regions in the current study likely comprise the greatest proportion of farmers due to the classification of SES based on the ON-Marg, which correlates with a lower proportion of post-secondary education completion (approximately two-thirds of farmers do not have any post-secondary education) (Statistics Canada, 2016). Current literature on risk perception among farmers has found that they often have a better understanding of the risks associated with contaminated groundwater due to its impact on their farms and are often more willing to test, if adequately incentivized (Keraita et al., 2008).

While SES was significantly associated with the incidence of testing, it was only weakly associated with repeat testing, specifically in rural areas where wells located in low SES areas were slightly more likely to be tested again compared to high SES areas. Instead, it was observed that a prior positive test result for well water contamination was the strongest predictor of repeat testing. Previous groundwater literature

indicates that tests results directly influence the “peace of mind” of well owners regarding their water quality, and thus positive tests may result in multiple retests, potentially regularly, to achieve and maintain that sense of security (Imgrund et al., 2011; Malecki et al., 2017).

Current literature on the relationship between season and testing suggest that private water well owners are less inclined to travel long distances during winter months (Imgrund et al., 2011; Murphy et al., 2016), while retesting rates are expected to mirror groundwater contamination rates, with warmer seasons (spring and summer) characterised by higher FIO detection rates (Imgrund et al., 2011). Conversely, wells tested for the first time during summer or fall during the current study were less likely to be retested compared to those tested for the first time during winter. This difference may be attributable to index test results, such that wells which tested negative during months typically associated with high contamination rates may lead to a false sense of security (i.e. contrary to expectation) among well owners. Furthermore, wells testing positive during low contamination months may result in the opposite effect, with well owners re-testing their wells to achieve reassurance. Peace of mind and reassurance about water quality have been found to represent strong drivers of repeat testing (Imgrund et al., 2011; Kreutzwiser et al., 2011), potentially explaining the observed relationship.

Index test status was found to be the most significant predictor of retesting in the current study; accordingly, future campaigns should focus on a collaborative approach between public health agencies and non-government organizations (NGOs) that emphasize both the public health significance and the necessity and relative simplicity of testing, albeit this will inevitably vary between jurisdictions. These strategies have been shown to assist in identifying and addressing specific barriers to testing (e.g. psychological factors, awareness/knowledge of importance, awareness/knowledge of testing protocols) among some individuals and populations (Flanagan et al., 2015; Flanagan et al., 2015). Campaigns should also raise awareness on the fluid nature of groundwater contamination (i.e. one negative result does not equate to permanent safety) and subsequently the importance of regular testing. It is important to note that the employed dataset pertains to a region in which the provincial government offers free testing of private domestic supplies, and as such, negates the cost of testing as a behavioural barrier. Previous studies have shown that the cost of groundwater analyses, in addition to the time required for sample collection and drop-off (i.e. inconvenience), represent two significant barriers for well owners (Hexemer et al., 2008; Flanagan et al., 2015; MacDonald Gibson and Pieper, 2017). For example, Hexemer et al. (2008) report that removal of the barriers of cost and inconvenience approximately doubled the private well water background sampling ($\approx 25\%$) rate in Waterloo (Ontario) region during a 1-year intervention study. Conversely, background sampling rates remain significantly lower in the United States

where a majority of private groundwater sampling is only available on a commercial basis, apart from states in which testing is legally required as part of a real estate transaction (e.g. Maine, New Jersey) (Atherholt et al., 2015; Flanagan et al., 2015). Indeed, MacDonald Gibson & Pieper (2017) have recently recommended that appropriate state agencies or foundations should analyse options for provision of financial assistance to low-income private well users for monitoring and maintenance. Notwithstanding, Flanagan et al. (2015) also suggest that cost/convenience barriers do not fully account for low testing rates, thus results of the current study may be employed by governmental agencies to design and optimise similar free well testing campaigns as that operated by Public Health Ontario.

4.1. Study limitations

Studies based on retrospective cohorts, with area-level predictors often encounter common issues, such as discrepancies in study periods based on data sources used and the lack of individual-level data. SES and rurality predictors in the current study are based on the 2006 census, whereas the study period begins in 2012. Due to the discrepancies in timing, a potential demographic shift may have occurred in some DAs, leading to potential SES misclassification in some cases. A potential solution was to use the recent census and ON-Marg update; however, this was not possible due to the compromised data quality of the 2011 update of the ON-Marg, following a revised government policy that did not mandate completion of the long census form (Green and Milligan, 2010). Moreover, a potential limitation of using the ON-Marg as a proxy measure of area-level SES is that it has not been validated for the study outcomes of interest including behavioural and environmental outcomes. Instead, the ON-Marg was developed and validated for investigating health outcomes including morbidity and mortality. As such, the ON-Marg may not fully capture the most relevant dimensions of SES that are important for predicting well water testing behaviour. In defining the presence/absence of an index test, the researchers have excluded all tests carried out prior to January 1st 2012, as well as those tests for which the well address could not be geocoded or linked to a well from the WWIS, and as such, the authors caution that a “first test” as defined within this study may not be representative of the first time a well owner/user tested their well. While the 5-year study duration and sheer volume of well tests considered as part of the current study likely result in an accurate identification of overarching groundwater quality and human behavioural patterns, this caveat should be considered when interpreting study findings. Finally, several previous studies have found that local and site-specific factors including local/regional geology, hydrogeological setting, individual source design, construction and maintenance, and household-level knowledge and behaviours are all important drivers of source susceptibility to microbial contamination (Hynds et al., 2012, 2013, 2014). Due to the geographical and spatial extent of the current study, it was not possible to investigate these inherently local issues. The authors consider that while findings from the current study are indicative of trends at the regional level, future work should seek to integrate high-resolution local or source-specific data (e.g. household surveying) with large geospatial datasets in order to more clearly elucidate the behavioural mechanisms associated with “healthy behaviours”.

5. Conclusion

The presented study employed a geostatistical approach to assess predictors of well water testing and observed that area-level SES played a significant role in the prevalence of well water testing but a lesser role in re-testing. In rural regions, lower SES was associated with an increased likelihood of testing, but the opposite was observed in urban and peri-urban regions. The study also observed that index test status and season were the most significant predictors of conducting multiple tests. The results of the current study could help form the basis of future

research in the area of environmental behaviours and private groundwater testing. Future research on private water well testing should focus on formulating a comprehensive SES predictor that incorporates both individual factors, and area factors. Area-level factors provide information on neighbourhood and regional level influences, which, as shown in this study, can influence both the prevalence and frequency of testing. Due to the significance of index test status and season with respect to the frequency of testing, further research is required to understand the complex relationship between time, place, personal risk perception and testing and its relationship to other factors such as SES. Future studies should seek to employ integrated “mixed methods” approaches, whereby large spatial dataset analyses are merged with higher resolution individual/household level surveying, in order to effectively examine behavioural barriers at multiple scales.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was financially supported by the Canadian Foundation for Infectious Diseases (CFID). The authors would like to thank the Public Health Ontario Laboratory (PHOL) for dataset access, and their tireless work in permitting free access to well water testing for all residents of Ontario. The authors would also like to thank Mr. Tanner Hoffman for his assistance during manuscript preparation.

References

- American Psychological Association, 2018. Education and socioeconomic status. Retrieved October 1, 2018, from <http://www.apa.org/pi/SES/resources/publications/education.aspx>.
- An, Y., Kampbell, D.H., Breidenbach, G.P., 2002. *Escherichia coli* and total coliforms in water and sediments at lake marinas. *Environ. Pollut.* 120 (3), 771–778.
- Atherholt, T.B., Korn, L.R., Louis, J.B., Procopio, N.A., 2015. Repeat sampling and coliform bacteria detection rates in New Jersey domestic wells. *Ground Water Monit. Remediat.* 35 (2), 70–80.
- Berzofsky, M., Smiley-McDonald, H., Moore, A., Krebs, C., 2014. Measuring socioeconomic status (SES) in the NCVS: background, options, and recommendations (Rep. No. 0213170.001.002.001). Retrieved October 10, 2018, from https://www.bjs.gov/content/pub/pdf/Measuring_SES-Paper_authorship_corrected.pdf.
- Blizzard, L., Hosmer, W., 2006. Parameter estimation and goodness-of-fit in log binomial regression. *Biom. J.* 48, 5–22.
- Charrois, J.W., 2010. Private drinking water supplies: challenges for public health. *Can. Med. Assoc. J.* 182 (10), 1061–1064.
- Cooley, D., 2018. Googleway: accesses Google maps APIs to retrieve data plot maps. R Package version 2.6. Retrieved October 15, 2018, from <https://CRAN.R-project.org/package=googlewayV>.
- Corkal, D., Schutzman, W.C., Hilliard, C.R., 2004. Rural water safety from the source to the on-farm tap. *J. Toxicol. Environ. Health* 67 (20–22), 1619–1642.
- Cummings, P., 2009. The relative merits of risk ratios and odds ratios. *Arch. Pediatr. Adolesc. Med.* 163 (5), 438–445.
- Dunn, J., Matheson, F., Smith, K., 2012. Overview of the ontario marginalization index (ON-Marg). Retrieved August 3, 2017, from <http://www.torontohealthprofiles.ca/onmarg/additionalResources/OverviewOfONMarg06July2012.pdf>.
- Flanagan, S.V., Marvinnay, R.G., Zheng, Y., 2015. Influences on domestic well water testing behavior in a Central Maine area with frequent groundwater arsenic occurrence. *Sci. Total Environ.* 505, 1274–1281.
- Google Developers, 2017. Google maps javascript API. Retrieved October 15, 2018, from <https://developers.google.com/maps/documentation/javascript/geocoding#ReverseGeocoding>.
- Government of Canada, 2013. Guidelines for Canadian drinking water quality: guideline technical document - total coliforms. Retrieved April 30, 2018, from <https://www.canada.ca/en/health-canada/services/publications/healthy-living/guidelines-canadian-drinking-water-quality-guideline-technical-document-total-coliforms.html>.
- Government of Ontario, 2017. Water supply wells requirements and best practices. Retrieved from <https://www.ontario.ca/document/water-supply-wells-requirements-and-best-practices>.
- Green, D., Milligan, K., 2010. The importance of the long form census to Canada. *Can. Public Policy* 36 (3), 383–388.
- Hexemer, A.M., Pintar, K., Bird, T.M., Zentner, S.E., Garcia, H.P., Pollari, F., 2008. An investigation of bacteriological and chemical water quality and the barriers to private well water sampling in a Southwestern Ontario Community. *J. Water Health* 6 (4), 521–525.

- Hynds, P.D., Misstear, B.D., Gill, L.W., 2012. Development of a microbial contamination susceptibility model for private domestic groundwater sources. *Water Resour. Res.* 48 (12).
- Hynds, P.D., Misstear, B.D., Gill, L.W., 2013. Unregulated private wells in the Republic of Ireland: consumer awareness, source susceptibility and protective actions. *J. Environ. Manag.* 127, 278–288.
- Hynds, P., Misstear, B.D., Gill, L.W., Murphy, H.M., 2014. Groundwater source contamination mechanisms: physicochemical profile clustering, risk factor analysis and multivariate modelling. *J. Contam. Hydrol.* 159, 47–56.
- Imgrund, K., Kreutzwiser, R., Loe, R.D., 2011. Influences on the water testing behaviors of private well owners. *J. Water Health* 9 (2), 241–252.
- Keraita, B., Drechsel, P., Konradsen, F., 2008. Perceptions of farmers on health risks and risk reduction measures in wastewater-irrigated urban vegetable farming in Ghana. *J. Risk Res.* 11 (8), 1047–1061.
- Kralj, B., 2009. Measuring rurality - RIO 2008_BASIC: methodology and results. Retrieved from: <https://www.oma.org/wp-content/uploads/2008rio-fulltechnicalpaper.pdf>.
- Kreutzwiser, R., Loe, R.D., Imgrund, K., Conboy, M.J., Simpson, H., Plummer, R., 2011. Understanding stewardship behaviour: factors facilitating and constraining private water well stewardship. *J. Environ. Manag.* 92 (4), 1104–1114.
- Krolik, J., Evans, G., Belanger, P., Maier, A., Joyce, A., Majury, A., 2014. Microbial source tracking and spatial analysis of *E. coli* contaminated private well waters in southeastern Ontario. *J. Water Health* 12 (2), 348–357.
- MacDonald Gibson, J., Pieper, K.J., 2017. Strategies to improve private-well water quality: a North Carolina perspective. *Environ. Health Perspect.* 125 (7), 076001.
- Maier, A., Krolik, J., Randhawa, K., Majury, A., 2014. Bacteriological testing of private well water: a trends and guidelines assessment using five years of submissions data from Southeastern Ontario. *Can. J. Public Health* 105 (3), 203–208.
- Malecki, K.M., Schultz, A.A., Severtson, D.J., Anderson, H.A., Vanderslice, J.A., 2017. Private-well stewardship among a general population-based sample of private well-owners. *Sci. Total Environ.* 601–622, 1533–1543.
- Matheson, F. I., & Van Ingen, T. (2017). 2011 Ontario marginalization index: user guide. Retrieved August 3, 2017, from https://www.publichealthontario.ca/en/DataAndAnalytics/Documents/User_Guide_2011_ON-Marg.pdf
- Matheson, F.I., Dunn, J.R., Smith, K.L., Moineddin, R., Glazier, R.H., 2012a. Development of the Canadian Marginalization Index: a new tool for the study of inequality. *Can. J. Public Health* 103 (Suppl.2), S12–S16.
- Matheson, F.I., Dunn, J.R., Smith, K.L., Moineddin, R., Glazier, R.H., 2012b. ON-Marg user guide. Retrieved August 3, 2017, from: http://www.torontohealthprofiles.ca/onmarg/userguide_data/ON-Marg_user_guide_1.0_FINAL_MAY2012.pdf.
- Moffat, H., Struck, S., 2011. Water-borne disease outbreaks in Canadian small drinking water systems. Retrieved September 20, 2018, from: http://www.ncceh.ca/sites/default/files/SDWS-Water-borne_EN.pdf.
- Murphy, H.M., Thomas, M.K., Medeiros, D.T., McFayden, S., Pintar, K.D., 2016. Estimating the number of cases of acute gastrointestinal illness (AGI) associated with Canadian municipal drinking water systems. *Epidemiol. Infect.* 144 (7), 1371–1385.
- Murphy, H.M., Prioleau, M.D., Borchardt, M.A., Hynds, P.D., 2017. Review: epidemiological evidence of groundwater contribution to global enteric disease, 1948–2015. *Hydrogeol. J.* 25 (4), 981–1001. <https://doi.org/10.1007/s10040-017-1543-y>.
- Nahar, N., Hossain, F., Hossain, M., 2008. Health and socioeconomic effects of groundwater arsenic contamination in rural Bangladesh: new evidence from field surveys. *J. Environ. Health* 70 (9), 42–47.
- National Oceanic and Atmospheric Administration, 2017. Meteorological versus astronomical seasons. Retrieved January 20, 2018, from: <https://www.ncei.noaa.gov/news/meteorological-versus-astronomical-seasons>.
- OECD Directorate for Public Governance and Territorial Development, 2011. OECD regional typology. Retrieved January 20, 2018, from https://www.oecd.org/cfe/regional-policy/OECD_regional_typology_Nov2012.pdf.
- Pampalon, R., Raymond, G., 2000. A deprivation index for health and welfare planning in Quebec. *Chronic Disabil. Can.* 21 (3), 104–113.
- Pampalon, R., Hamel, D., Gamache, P., 2010. Health inequalities in urban and rural Canada: comparing inequalities in survival according to an individual and area-based deprivation index. *Health Place* 16 (2), 416–420.
- Public Health Ontario, 2013. Public health inspectors guide to the principles and practices of environmental microbiology. Retrieved September 3, 2017, from: https://www.publichealthontario.ca/en/eRepository/Public_Health_Inspectors_Guide_2013.pdf.
- Public Health Ontario, 2017a. Beneath the surface: probing the dynamics of private well water. Retrieved October 4, 2018, from: https://www.publichealthontario.ca/en/About/Research/Pages/Research-in-action_private-well-water.aspx.
- Public Health Ontario, 2017b. Drinking water testing-private citizen. Retrieved September 3, 2017, from: https://www.publichealthontario.ca/en/ServicesAndTools/LaboratoryServices/Pages/Drinking_Water_Testing_Private_Citizen.aspx.
- R Core Team, 2018. R: A Language and Environment for Statistical Computing. Austria, Vienna Retrieved October 15, 2018, from: <https://www.R-project.org>.
- Statistics Canada, 2012. Map 1 - 2011 census agricultural regions and census divisions. Retrieved July 25, 2017, from: <http://www.statcan.gc.ca/eng/ca2011/110006/map1on-eng.pdf>.
- Statistics Canada, 2015. Dissemination area (DA). Retrieved August 3, 2017, from: <http://www12.statcan.gc.ca/census-recensement/2011/ref/dict/geo021-eng.cfm>.
- Statistics Canada, 2016. Get to know Canadian farmers and their families. Retrieved November 15, 2018, from: <https://www.statcan.gc.ca/eng/ca2011/ha#a1-3-2>.
- Statistics Canada, 2017a. Archived – from urban areas to population centres. Retrieved January 20, 2018, from: <https://www.statcan.gc.ca/eng/subjects/standard/sgc/notice/sgc-06>.
- Statistics Canada, 2017b. Postal Code Conversion File Plus (Version 6A1) Retrieved from Queen's Research Data Centre.
- Switzer, D., Teodoro, M.P., 2017. Class, race, ethnicity, and justice in safe drinking water compliance. *Soc. Sci. Q.* 99 (2), 524–535.
- Thomas, M.K., Murray, R., Flockhart, L., Pintar, K., Fazil, A., Nesbitt, A., ... Pollari, F., 2015. Estimates of foodborne illness-related hospitalizations and deaths in Canada for 30 specified pathogens and unspecified agents. *Foodborne Pathogens and Disease* 12 (10), 820–827.