83437_Auto_Edited.docx

WORD COUNT 5422
TIME SUBMITTED 06-APR-2023 09:07PM
PAPER ID 98437610

Name of Journal: *World Journal of Gastroenterology*
Manuscript NO: 83437
Manuscript Type: ORIGINAL ARTICLE

Observational Study

Spatial Cluster Mapping and Environmental Modeling in Pediatric Inflammatory Bowel Disease

Spatial-environmental analysis of pediatric IBD

Mielle Michaux, Justin M Chan, Luke Bergmann, Luis F Chaves, Brian Klinkenberg, Kevan Jacobson

Abstract

BACKGROUND

Geographical (geospatial) clusters have been observed in inflammatory bowel disease (IBD) incidence and linked to environmental determinants of disease, but pediatric spatial patterns in North America are unknown. We hypothesized that we would identify geospatial clusters in the pediatric IBD (PIBD) population of British Columbia (BC), Canada and associate incidence with ethnicity and environmental exposures.

AIM

To identify PIBD clusters and model how spatial patterns are associated with population ethnicity and environmental exposures.

METHODS

1,183 patients were included from a BC Children's Hospital clinical registry who met the criteria of diagnosis with IBD <= age 16.9 from 2001–2016 with a valid postal code on file. A spatial cluster detection routine was used to identify areas with similar

incidence. An ecological analysis employed Poisson rate models of IBD, Crohn's disease (CD), and ulcerative colitis (UC) cases as functions of areal population ethnicity, rurality, average family size and income, average population exposure to green space, air pollution, and vitamin-D weighted ultraviolet light from the Canadian Environmental Health Research Consortium, and pesticide applications.

RESULTS

Hot spots (high incidence) were identified in Metro Vancouver (IBD, CD, UC), southern Okanagan regions (IBD, CD), and Vancouver Island (CD). Cold spots (low incidence) were identified in Southeastern BC (IBD, CD, UC), Northern BC (IBD, CD), and on BC's coast (UC). No high incidence hot spots were detected in the densest urban areas. Modeling results were represented as incidence rate ratios (IRR) with 95% confidence intervals (CI). Novel risk factors for PIBD included fine particulate matter ($PM_{2.5}$) pollution (IRR = 1.294, CI = 1.113 - 1.507, P <0.001) and agricultural application of petroleum oil to orchards and grapes (IRR = 1.135, CI = 1.007 - 1.270, *P* = 0.033). South Asian population (IRR = 1.020, CI = 1.011 - 1.028, P <0.001) was a risk factor and Indigenous population (IRR = 0.956, 95%CI = 0.941 - 0.971, P <0.001), family size (IRR = 0 .467, CI = 0.268 - 0.816, *P* = 0.007), and summer ultraviolet (IBD = 0.9993, CI = 0.9990 – 0.9996, P <0.001) protective factors as previously established. Novel risk factors for CD, as for IBD, included: $PM_{2.5}$ air pollution (IRR = 1.230, CI = 1 .056 - 1.435, *P* = 0.008) and agricultural petroleum oil (IRR = 1.159, CI = 1.002 - 1.326, *P* = 0.038). Indigenous population (IRR = 0.923, CI = 0.895 – 0.951, P <0.001), as previously established, was a protective factor. For UC, rural population (UC IRR = 0.990, CI = 0.983 - 0.996, *P* = 0.004) was a protective factor and South Asian population (IRR = 1.054, CI = 1.030 – 1.079, P <0.001) a risk factor as previously established.

CONCLUSION

PIBD spatial clusters were identified and associated with known and novel environmental determinants. The identification of agricultural pesticides and $PM_{2.5}$ air pollution needs further study to validate these observations.

Key Words: Inflammatory bowel diseases; Crohn disease; Ulcerative colitis; Pesticides; Air pollution; South Asian people

Michaux M, Chan JM, Bergmann L, Chaves LF, Klinkenberg B, Jacobson K. Spatial Cluster Mapping and Environmental Modeling in Pediatric Inflammatory Bowel Disease. *World J Gastroenterol* 2023; In press

Core Tip: Utilizing spatial mapping methodology, high and low incidence clusters of pediatric inflammatory bowel disease (IBD) were identified in British Columbia, Canada. Associating geographical location with IBD, rurality was negatively associated with ulcerative colitis. Notably, no high incidence hot spots were detected in the densest urban areas, suggesting unexplored urban protective factors. Novel risk factors for PIBD and specifically Crohn's disease included fine particulate matter ($PM_{2.5}$) pollution and agricultural applications of petroleum oil to orchards and grapes. Spatial distribution was partially explained by rurality, population ethnicity, family size, pesticide applications, air pollution, ultraviolet exposure, and residential greenness.

INTRODUCTION

Canada has one of the world's highest rates of pediatric inflammatory bowel disease (IBD), which includes Crohn's disease (CD) and ulcerative colitis (UC)[1]. Incidence within countries can be quite varied, but local patterns of IBD incidence are unknown for most countries[1]. Results from studies in Finland, Norway, Northern France, and Manitoba, Canada suggest that IBD may have a clustered spatial distribution, but more research is necessary for understanding local spatial patterns[2–5]. To date, no study has focused on empirically detecting and evaluating disease clusters in Canadian or North

American PIBD populations. Notably, detecting local spatial clusters of IBD would allow for a better understanding of clinical populations and identify areas where services are needed. Moreover, spatial epidemiology can be used to identify novel environmental risk and protective factors. Although it is evident that environmental factors are important determinants for disease development, additional study is necessary.

A variety of population and environmental factors are thought to influence IBD risk. High incidence has been observed in Canadian populations of Jewish ethnicity and South Asian pediatric populations in British Columbia (BC), and low incidence has been reported in Canadian Indigenous communities[4,6,7]. Higher socioeconomic status has been associated with IBD[4,8]. A variety of environmental exposures, including rural residence, green space, ultraviolet (UV) radiation, and air pollution have also been studied[9–13]. No study known to us has examined pesticides as a potential determinant of IBD. Pesticides can be present in food and the environment, and a variety of pesticides have been linked with changes to the gut microbiota which could have implications for IBD development[14]. Based on Canadian immigration and rural residence studies of IBD, exposures during early life appear to be important[9,15].

The aims of this exploratory study were 1) to determine spatial patterning of PIBD and identify location of disease hot and cold spots in the Canadian province of BC and 2) to model the association between IBD case counts and population-level ethnicity, average income, rural residence, and known as well as novel environmental determinants. We hypothesized that we would identify IBD clusters that would be associated with ethnicity and environmental exposures. Modeling potential population risk factors provided context to the spatial analysis and helped identify areas where additional novel environmental risk or protective factors might have meaningfully affected disease incidence.

MATERIALS AND METHODS

Study area

BC is Canada's westernmost province. It is divided into five Regional Health Authorities, which can be further subdivided into 89 Local Health Areas (LHAs) (see Figure 1 and Supplementary Figure 1). The majority of BC's population live in urban areas located in the Vancouver Coastal and Fraser Health Authorities near the United States border[16]. Northern, Interior, and Vancouver Island Health serve largely more rural populations.

Participants

Patients for this study were selected from a clinical registry of IBD patients maintained by the BC Children's Hospital (BCCH) Division of Gastroenterology, Hepatology and Nutrition who were diagnosed with or received care for IBD at BCCH in Vancouver[17]. Author KJ is data steward for this registry. As BCCH is the only tertiary care pediatric institution with academic pediatric gastroenterologists in the province, it is where most children with IBD are diagnosed. Our recent study comparing PIBD incidence derived from the BCCH registry with incidence derived from population-wide provincial health administrative data between 1996 and 2008 found similar overall rates, particularly from 2001 onwards reaffirming the validity of the registry as a reflection of population-based cases[18]. Notably, a small number of cases are diagnosed in the community, use health services from a different province or, in the case of older patients (> 16.9 years), are diagnosed by adult GI physicians. Registry patients were excluded from the study if they were diagnosed outside the study period (2001 to 2016) or over age 16.9 years, did not have a valid postal code on file, or had a postal code associated with BCCH as this was likely not their permanent address. Postal codes were cleaned to provide consistent formatting. See Table 1 for BC pediatric incidence, and Supplementary Table 1 for incidence by patient age. Patient six-digit postal code point location at diagnosis was associated with latitude and longitude coordinates from DMTI Spatial Inc. obtained from the Canadian Urban Environmental Health Research Consortium (CANUE)[19]. Incidence was pooled for the full study period to maximize the number of cases in each analysis.

This study was approved by the University of British Columbia Children's and Women's Research Ethics Board (H19-00739).

Methods

Guidelines from the REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement were adapted for this ecological study. The statistical methods of this study were reviewed by biostatistician Jeffrey Bone from the BC Children's Hospital Research Institute. Data is summarized in Table 2, with additional details on environmental exposures presented in the Supplementary materials. Ethnicity and family size (married or common-law spouses, single parents, and at least one child) data from the Canadian census were obtained at the Dissemination Area level and resampled to LHAs by population-weighted overlay[16,20,21]. Average family income for 2005 and 2015 was available for LHAs[22]. To approximate average environmental exposures for each LHA, census population age 0-19 was used to population-weight postal code exposure data in a process that captured 92.6% of BC's youth population (excluded population resided in areas not covered by single link postal codes). Environmental data provided by CANUE for six digit postal codes included Normalized Difference Vegetation Index (NDVI)[19,23–26] greenness[27,28], vitamin D UV[19,29,30], nitrogen dioxide (NO_2)[19,31,32], ozone (O_3),[33–37], and fine particulate matter ($PM_{2.5}$)[19,38]. Data on metam, petroleum oil, and glyphosate pesticides was resampled from Global Pesticide Grids[39,40]. The least populous LHA was excluded from regression modeling due to missing data.

Spatial cluster analysis

Standardized incidence ratios (SIRs) for each of BC's LHAs were calculated from incident registry cases using BC incidence as the reference rate[41]. DataBC provided geographic data[42]. To reduce the instability in disease rates caused by low case counts and populations, rates were averaged for the study period. In addition, spatial linear empirical Bayes estimation was used to smooth SIRs in areas with average pediatric populations under 10,000[43]. The smoothing process calculated averages between a

LHA's SIR and the average SIR value of adjacent LHAs. Spatial relationships for smoothing and clustering were defined by direct adjacency (queen contiguity) to approximately model geographic connectivity.

Two forms of spatial analysis were used to identify spatial patterns in PIBD. The Global Moran's I statistic measured the degree to which spatial patterning of IBD, CD, and UC in BC was clustered, dispersed, or had no detectable spatial pattern. The Local Moran's I statistic was used as a Local Indicator of Spatial Association (LISA) to identify the location of hot spots (clusters of comparatively high SIRs) and cold spots (clusters of comparatively low SIRs) among LHAs[44]. To approximate the likelihood that a cluster would arise by chance, we used 999 Monte Carlo simulations to compute a pseudo-*P* value[44], which was then adjusted with a Holm correction for multiple comparisons. We chose LISA statistics over spatial and spatiotemporal SCAN methods given the irregular nature of the spatial areal units under analysis and the relative rarity of the disease, conditions under which LISA statistics have offered high sensitivity and specificity in cluster detection[45,46]. Data was aggregated over time (i.e. no spatio-temporal cluster analyses) given the aforementioned relative rarity of PIBD cases in BC.

Count regression modeling

Poisson Rate generalized linear models (PR-GLM) were used in an ecological analysis of LHAs to quantify the impact of population ethnicity, income, family size, rurality, air pollution, greenness, UV, and pesticide exposure on raw (unsmoothed) IBD, CD, and UC case counts. These models adjust for the population of each LHA and have an equal mean and variance. Model selection was based on the minimization of the Akaike Information Criterion using a stepwise algorithm combining backward elimination and forward addition[47]. Model selection was based on the minimization of the Akaike Information Criterion using a stepwise algorithm combining backward elimination and forward addition. PR-GLMs performed better than Negative Binomial, zero-inflated Poisson, and hurdle Poisson models. Considerations for reducing collinearity, conducting model selection, testing spatial independence of residuals, examining best model diagnostics, and considerations around the mapping of Incidence

Rate Ratios (IRRs) are presented in the Supplement materials. Multivariate models and high-quality Canadian census and exposure data were used to minimize sources of potential bias.

RESULTS

Spatial clustering results

The registry consisted of 1,232 eligible patients diagnosed at or before age 16.9. Of these, 49 were associated with invalid or BCCH postal codes. In the spatial cluster modeling of the remaining 1,183 PIBD patients, we observed a statistically significant (pseudo-P < 0.05) clustered distribution with a Moran's I statistic of 0.65 (P = 0.001). For the 780 CD patients we observed a Moran's I of 0.56 (P = 0.001) and for the 288 UC patients a Moran's I of 0.28 (P = 0.001). See Figure 2 for local cluster locations. Hot spots were observed in the Lower Mainland, the main urban centre of BC, for IBD ($P \leq 0.031$), CD ($P \leq 0.031$), and UC ($P \leq 0.007$), in the Okanagan region for IBD ($P \leq 0.029$) and CD (P = 0.030), and on Vancouver Island for CD (P = 0.034). A cold spot cluster was detected in southeastern BC for IBD (P values < 0.001), CD ($P \leq 0.0499$), and UC ($P \leq 0.040$), in Northern BC for IBD (P = 0.026) and CD (P = 0.038), and on BC's coast for UC (P = 0.036). We observed no LHA which had a significantly different trend in incidence than neighboring LHAs (spatial outliers) for IBD, CD, or UC.

Environmental modeling results

Summary statistics for variables are presented in Table 3 for BC and Supplementary Table 2 by Health Authority (see Figure 1 for reference and Table 1 for incidence). Variables (Table 3) were modeled individually and combined in a multivariate model. Lack of spatial clustering observed in the residuals indicates that the model assumption of spatial independence was met, and suggests that within the parameters of the analysis there were either no important missing variables with spatial patterning or that the existing predictors capture similar spatial variability of important unmeasured

variables. The final models included variables shown in Table 4, and logged average population as an offset variable.

For the whole PIBD population, indigenous ethnic origin, average family size, summer UV radiation, and metam pesticide application were identified as significant protective factors. In contrast, South Asian origin, greenness, $PM_{2.5}$, and petroleum oil application were significant risk factors. When broken down by disease subtype, Indigenous ethnic origin, NO_2, and O_3 were statistically significantly negatively correlated with CD, while maximum growing season greenness, $PM_{2.5}$, and petroleum oil application were significant risk factors. Rurality, and alfalfa and corn glyphosate were statistically significant protective factors for UC, while significant risk factors included South Asian origin, mean growing season greenness, O_3 pollution, and wheat glyphosate.

A 1% increase in the Indigenous population of a LHA was associated with a 4.4% decrease in the number of IBD cases (IRR = 0.956, $P < 0.001$) and a 7.7% decrease in CD (IRR = 0.923, $P < 0.001$), while a 20% increase in Indigenous population was associated with a 59% decrease in IBD. Each additional family member added to family size was associated with a 53.3% decrease in IBD cases (IRR = 0.467, $P = 0.007$). An increase of 1 J/m^{-2} in summer UV was associated with 0.001% decrease in IBD cases (IRR = 0.999, $P < 0.001$) while an increase of 500 J/m^{-2} was associated with a 29.4% decrease. A 1% increase in rural population was associated with a 1% decrease in UC (IRR = 0.990, $P = 0.004$). Each unit increase of metam pesticide was associated with a 6.9% decrease in IBD cases (IRR = 0.931, $P < 0.001$ for 1 kg/ha). Each ppb increase in O_3 was associated with a 9.5% increase in UC (IRR = 1.095, $P = 0.001$) and a 5.3% decrease in CD (IRR = 9.47, $P = 0.010$). Each ppb increase in NO_2 was associated with a 4.8% decrease in CD cases (IRR = 0.952, $P = 0.006$). Each unit increase of glyphosate applied to wheat was associated with an increase in UC (IRR = 121.196, $P = 0.019$ for 1 kg/ha), while the same pesticide applied to corn (IRR = 0.828, $P = 0.021$ for 1 kg/ha) and alfalfa (IRR = 0.001, $P = 0.024$ for 1 kg/ha) was a significant protective factor. It is important to note that glyphosate applications to corn (range 0.00 – 6.06, mean 0.49 kg/ha) were much higher

than those to afalfa (range 0.00 – 0.17, mean 0.02 kg/ha) and wheat (range 0.00 – 0.21, mean 0.02 kg/ha), and the extremely high and low IRRs for glyphosate are more reflective of the small data values than actual effect size.

A 1% increase in the percent of South Asian residents was associated with a 2% increase in IBD (IRR = 1.020, $P < 0.001$) and a 5.4% increase in UC (IRR = 1.054, $P < 0.001$), while a 20% increase in South Asian residents was associated with a 47.5% increase in IBD. A 1% increase in maximum growing season greenness was associated with a 6% increase in IBD and a 3.8% increase in CD (IRR = 1.060, $P < 0.001$, and IRR = 1.038, $P = 0.002$, respectively). A 1% increase in mean growing season greenness was associated with a 4.3% increase in UC (IRR = 1.043, $P = 0.006$). Each 1 $\mu g/m^3$ concentration increase in $PM_{2.5}$ air pollution was associated with a 29.4% increase in IBD cases (IRR = 1.294, $P < 0.001$) and a 23% increase in CD cases (IRR = 1.230, $P = 0.008$). Finally, a 1 kg/ha increase in the application of petroleum oil in grapes and orchards was associated with a 13.5% increase in IBD (IRR = 1.135, $P = 0.033$) and a 15.9% increase in CD (IRR = 1.159, $P = 0.038$)

DISCUSSION

Based on our exploratory analysis describing spatial patterning of PIBD in BC, incidence varied substantially across the province in both rural and urban areas, likely reflective of BC's diverse population and environments. The spatial distribution of IBD, CD, and UC was significantly clustered during the period of 2001 to 2016, with substantial overlap between cluster locations for each of the three. The Lower Mainland hotspot and southeastern BC cold spot were consistent across the study period in the IBD population. As previously reported in PIBD, UC represented a lower percentage of IBD cases than CD and in addition, displayed a less clustered distribution with fewer shared clusters. It is unlikely that these observed clusters resulted from chance. Notably, patients living near the eastern border (Alberta) may receive care in that province, potentially explaining the cold spot in southeastern BC.

In our modeling of the association between IBD case counts and population and environment variables, higher proportions of South Asians tended to be associated with higher IBD and UC case counts. In Ontario, Canada, similar incidence of IBD has been observed for children of immigrants of South Asian origin born in Canada and children of non-immigrants[15]. However, higher rates of IBD have been documented in South Asian populations in BC, the United States, the United Kingdom, New Zealand, and Singapore and Malaysia [7,48–51]. Our observation of lower IBD and CD cases associated with higher proportions of Indigenous residents is consistent with previous Canadian studies[4,52]. When interpreting modeling results, locations where IBD rates were well predicted by population ethnicity should still be considered places where environmental risk or protective factors were also present. Larger family size has been associated with a protective effect on CD, which is consistent with our results for IBD but not CD[6]. Higher socioeconomic status is an established risk factor[4,8], but average income quantile was not a significant predictor in any model (though it improved CD model fit). While several identified hot spots were located in the highest income areas of BC and cold spots in the lowest income areas, there were also hot spots with low income and cold spots with high income. It may be that average household income is less relevant than other socioeconomic indicators in BC, or that an effect was not observable for aggregated populations.

Our results suggest that at the LHA level, residential environmental exposures at diagnosis may also be significant potential determinants. Links with vitamin D and sunlight have been inconsistent, due in part to variability in study design and exposure assessment[12]. Our findings indicate that summer vitamin D UV radiation may confer a protective effect on IBD development. An individual unit of UV (1 J/m^{-2}) is quite small which resulted in a low measured IRR. Mean UV in BC is significantly higher in summer (6,053.15 J/m^{-2}) than winter (274.98 J/m^{-2}) and varies within the province (summer maximum of 7,272.52 and minimum of 4,488.80 J/m^{-2}). An increase in UV of 500 J/m^{-2}, which is more representative of actual geographic and seasonal UV variation, was associated with a 29.4% decrease in IBD cases. In the analysis stratified by disease

type, summer and winter UV vitamin D were nonsignificant protective factors that improved model fit for CD and UC, respectively. Smaller sample size in stratified analysis may have contributed to nonsignificant results. Low winter UV across BC could be particularly relevant for people of South Asian descent who have a higher level of skin melanin and require substantially more UV exposure to synthesize sufficient vitamin D[30]. $PM_{2.5}$ air pollution was a significant risk factor for PIBD and CD. Italian (IBD) and Chinese (UC but not CD) studies of middle and older adults have identified $PM_{2.5}$ as a risk factor for incident cases as well as IBD and UC hospitalizations in China [53–55]. In contrast, an Ontario pediatric study found no association and a European adult study which found a negative association with $PM_{2.5}$[13,56]. Though population exposure for most LHAs met 2020 national and provincial air quality objectives, BC experiences seasonal wildfire events which can cause short-term high $PM_{2.5}$ concentrations that would be obscured in the yearly average values used in this study[57]. Regular high exposure events should be investigated further, especially as climate change is projected to increase wildfire potential[58]. O_3 air pollution was a significant risk factor for UC which is consistent with a Chinese study which measured an association between O_3 and IBD and UC hospital visits[54]. We observed statistically significant negative associations with CD for NO_2, and O_3. This is in contrast to the lack of association observed for IBD in Ontario and Europe, a United Kingdom study which found a positive association between NO_2 pollution and CD onset before age 23, and a Chinese study which found a positive association between NO_2 and UC incidence in middle and older aged adults[13,55,56,59]. Differential effects on CD and UC have been observed for environmental exposures such as smoking and appendectomy[60].

BC had few areas of low residential greenness and high incidence (Supplementary Figures 3 – 5), and we observed statistically significant positive associations between IBD, CD, UC, and greenness. Our findings contrast with a pediatric cohort study in Ontario, which found a protective effect for maximum growing season NDVI at 250 m[11]. This inconsistency could be due to methodological differences between the two studies. NDVI is a measure of vegetation greenness only

and may not capture other elements of green spaces, such as vegetation composition, environmental microbiome, or pesticide and herbicide applications, which may differ in BC. Measures of greenness in BC are highly dependent on specific indicators, as a Metro Vancouver study comparing green space metrics found that NDVI at 250 m from postal codes diverged significantly from other measures such as street tree density, total percentage of green space, and park quality[61].

A previous Canadian study found a protective effect of rurality on pediatric CD and UC, while our study only observed an effect for UC[9]. BC may have more diverse environments than other provinces; for example, the largely rural Interior Health Authority had many of the highest average $PM_{2.5}$ and O_3 exposures while the majority-urban Lower Mainland region included significant sections of Agricultural Land Reserve. A hot spot near Vancouver included mostly suburban LHAs rather than the main urban center, which is similar to results observed in Oslo, Norway[62]. Perhaps some suburban and rural areas lack protective effects conferred by other rural regions while also missing potential health-promoting features of dense urban areas (e.g. public transportation, public parks, and access to amenities). Some rural and peri-urban areas can also be associated with potential risk factors such as petroleum pesticides.

A novel result in this study is the measured associations between pesticides and IBD. Petroleum oil applied to grapes and orchards was a significant risk factor for IBD and CD. Indeed, exposure to agricultural petroleum oil has been previously associated with systemic autoimmunity (measured with antinuclear antibodies) and rhinitis in a prospective cohort of pesticide applicators in the United States, suggesting that petroleum pesticide products may have inflammatory properties[63,64]. As numerous other pesticides have been linked with dysbiosis, and petroleum is used as a fungicide, agricultural applications of petroleum oil should be investigated for potential impacts on the gut microbiome which could induce CD[14,39]. Interestingly, glyphosate application appeared to have a different effect on UC depending on which crop it was applied to. Wheat crop glyphosate was a significant risk factor, while alfalfa and corn crop glyphosate were significant protective factors. Differences in agricultural practices

between alfalfa, corn, and wheat may be responsible for the negative or both associations observed for UC. The unexpected negative association between IBD and metam pesticide also warrants further investigation. This association could have resulted from an unmeasured confounder such as diet. For example, proximity to fruit and vegetable crops where metam is applied could be correlated with access to these food groups which are known to lower IBD risk[65].

A key strength of this study is the use of high resolution national environmental exposure metrics which increased the quality of the study and will facilitate comparisons between our results and future research. In addition, the clinical registry only contained patients with a confirmed IBD diagnosis, confirmed diagnosis of CD or UC, and accurate date of diagnosis which minimizes risk of misclassification. Moreover, we have demonstrated that our registry data is representative of the BC IBD population[18]. The use of aggregate case counts and reliable population data in both analyses and areal geographic analysis combined with spatial smoothing in the cluster analysis reduced the instability of disease rates caused by low cases numbers. The use of Monte Carlo simulation and a multiple comparison correction for statistical significance increased confidence in the spatial results. Finally, future healthcare service planning in BC would be implemented for health administrative units such as LHAs, so our scale of analysis would allow this research to be directly integrated into planning and intervention efforts.

Despite these strengths, there are several limitations which warrant discussion. Some patients were excluded due to missing data and it is possible that cases were missing from our clinical registry, either of which may have altered spatial clustering or biased modeling results. However, the registry likely included the vast majority of PIBD patients diagnosed in BC during the study period[18]. Patient data was not initially collected for research purposes, which may have affected available variables and may have contributed to missing data. Small numbers of cases could have produced large variation in incidence in sparsely populated areas. In addition, missing early environmental exposure data for several LHAs may have impacted the results, though

all but one of the modeled LHAs included environment data from the majority of the study period. An important unmeasured variable in this analysis is diet, and missing potential confounding variables may have biased the results. This was not a birth cohort study, so there is uncertainty about exposures at gestation or early life. However, specific critical periods for many environmental exposures have not been established; consequently, further prospective studies are required. Previous environmental studies have used average childhood exposure[11,13] or did not have a standardized lookback period dating from diagnosis[56,59]. Accordingly, we used broad metrics of average population-level exposure during the study period. This was an ecological analysis and our findings should not be used to make claims about individual risk.

CONCLUSION

To our knowledge, this is the first spatial hot spot analysis focused on PIBD in North America and the first study identifying an association with pesticides. Spatial cluster detection was a valuable method for exploring patterns of IBD, and we identified $PM_{2.5}$ air pollution, petroleum oil, glyphosate, and metam pesticides as novel determinants of PIBD. Given the inconsistency of IBD incidence in urban areas and relatively high incidence in some suburban and rural areas, future research should move beyond binary urban-rural classifications and use specific characteristics such as built environment and pollutant exposures to characterize environments. Expanded regional and global studies are needed to validate these results and to determine the relationship between timing of exposure and clinical onset of disease. Furthermore, the inclusion of other immune-mediated inflammatory diseases will likely uncover potential shared disease clusters and environmental determinants.

ARTICLE HIGHLIGHTS

Research background

Geospatial patterning has been observed in inflammatory bowel disease (IBD) incidence and linked to environmental determinants of disease. However, knowledge of North

American IBD spatial patterns is limited, and unknown in pediatric IBD (PIBD). A further understanding of geospatial patterns of IBD will help guide distribution of healthcare services and aid in identifying potential environmental risk and protective factors and populations at risk.

Research motivation

There is a lack of knowledge of the spatial distribution and environmental exposures relevant to PIBD in Canada and specifically in the Canadian province of British Columbia (BC).

Research objectives

The main objectives of this study were 1) to determine spatial patterning of PIBD and identify location of disease hot and cold spots in the Canadian province of BC during the period of 2001 - 2016 and 2) to model the association between IBD case counts and population-level ethnicity, average income, rural residence, and known as well as novel environmental determinants. Both objectives were addressed using the methods described below.

Research methods

The Moran's I statistic was used as a Local Indicator of Spatial association to measure the degree, location, and type of geographic clustering of PIBD incidence, a method which improves on visual analysis of mapped incidence by empirically quantifying clustering. Statistical significance of observed clusters was approximated using Monte Carlo simulation. Case counts of IBD, Crohn's disease (CD), and ulcerative colitis (UC) were modeled in Poisson rate models as a function of average population characteristics and average population environmental exposures to assess associations between IBD and rurality, ethnicity, income, family size, and air pollution, green space, ultraviolet (UV) light, and pesticide exposures. Data sources included a BC Children's Hospital clinical registry of patients diagnosed with IBD <= age 16.9, high-quality national

environmental exposure datasets developed for health research, and Canadian census data.

Research results

No high incidence hot spots were detected in the densest urban areas, suggesting unexplored urban protective factors. Rurality was negatively associated with ulcerative colitis. Novel risk factors for PIBD and specifically CD included fine particulate matter ($PM_{2.5}$) pollution and agricultural applications of petroleum oil to orchards and grapes. Spatial distribution was partially explained by rurality, population ethnicity, family size, pesticide applications, air pollution, UV exposure, and residential greenness.

Research conclusions

Pesticide and $PM_{2.5}$ exposure are linked to the development of PIBD. Suburban and low-density urban areas of BC appear to lack protective exposures conferred by rural and dense urban areas.

Research perspectives

Exploring geographic patterns of PIBD facilitated the identification of novel environmental determinants, which has prompted followup studies of environmental exposures and IBD onset.

ACKNOWLEDGEMENTS

M.M. was supported as a MSc student by the University of British Columbia Graduate Support Initiative and International Tuition Award and is currently a research assistant supported by the Moffat Foundation. J.M.C is a PhD candidate supported by the BC Children's Hospital Research Institute Studentship and the Lutsky Foundation. L.B. acknowledges support by the Canada Research Chairs Program and the Canada Foundation for Innovation. L.F.C. acknowledges funding from Indiana University. R.J. is a Senior Clinician Scientist supported by the Children with Intestinal and Liver

Disorders (CHILD) Foundation and the BC Children's Hospital Research Institute Clinician Scientists Award Program, University of British Columbia. NDVI metrics, nitrogen dioxide data, calculated ozone metrics, and $PM_{2.5}$ metrics were indexed to DMTI Spatial Inc. postal codes and provided by CANUE (Canadian Urban Environmental Health Research Consortium). Long-term monthly UV data were accessed *via* the CANUE data portal (https://canuedata.ca). Portions of this methodology were developed for M.M.'s MSc thesis[17].

83437_Auto_Edited.docx

ORIGINALITY REPORT

3%
SIMILARITY INDEX

PRIMARY SOURCES

1	L E Gyllenhammer, T L Alderete, C M Toledo-Corral, M Weigensberg, M I Goran. "Saturation of subcutaneous adipose tissue expansion and accumulation of ectopic fat associated with metabolic dysfunction during late and post-pubertal growth", International Journal of Obesity, 2015 Crossref	63 words — 1%
2	www.dovepress.com Internet	55 words — 1%
3	escholarship.mcgill.ca Internet	19 words — < 1%
4	www.ncbi.nlm.nih.gov Internet	14 words — < 1%
5	cedar.wwu.edu Internet	13 words — < 1%

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	< 12 WORDS
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	< 12 WORDS