INFLAMMATORY BOWEL DISEASE

*Open*

# Rural and Urban Residence During Early Life is Associated with Risk of Inflammatory Bowel Disease: A Population-Based Inception and Birth Cohort Study

This article has been corrected since Advance Online Publication and a erratum is also printed in this issue

Eric I. Benchimol, MD, PhD[1,2,3,4,5], Gilaad G. Kaplan, MD, MPH[6,7], Anthony R. Otley, MD, MSc[8], Geoffrey C. Nguyen, MD, PhD[5,9], Fox E. Underwood, MSc[6,7], Astrid Guttmann, MDCM, MSc[5,10], Jennifer L. Jones, MD, MSc[11], Beth K. Potter, PhD[4], Christina A. Catley, PhD[5], Zoann J. Nugent, PhD[12], Yunsong Cui, MSc[11], Divine Tanyingoh, MSc[6], Nassim Mojaverian, MSc[5], Alain Bitton, MD[13], Matthew W. Carroll, BMed, MHSc[14], Jennifer deBruyn, MD, MSc[15], Trevor J.B. Dummer, BA, MSc, PhD[16], Wael El-Matary, MD, MSc[17], Anne M. Griffiths, MD[10], Kevan Jacobson, MBBCh[18,19], M. Ellen Kuenzig, PhD[1,2,5], Desmond Leddin, MD[11], Lisa M. Lix, PhD[20], David R. Mack, MD[1,2,3], Sanjay K. Murthy, MD, MSc[5,21], Juan Nicolás Peña Sánchez, MD, MPH, PhD[22], Harminder Singh, MD, MPH[12,23], Laura E. Targownik, MD, MSHS[12,23], Maria Vutcovici, MD, MSc[13], Charles N. Bernstein, MD[12,23] on behalf of the Canadian Gastro-Intestinal Epidemiology Consortium

**OBJECTIVES:** **To determine the association between inflammatory bowel disease (IBD) and rural/urban household at the time of diagnosis, or within the first 5 years (y) of life.**

**METHODS:** **Population-based cohorts of residents of four Canadian provinces were created using health administrative data. Rural/urban status was derived from postal codes based on population density and distance to metropolitan areas. Validated algorithms identified all incident IBD cases from administrative data (Alberta: 1999–2008, Manitoba and Ontario: 1999–2010, and Nova Scotia: 2000–2008). We determined sex-standardized incidence (per 100,000 patient-years) and incident rate ratios (IRR) using Poisson regression. A birth cohort was created of children in whom full administrative data were available from birth (Alberta 1996–2010, Manitoba 1988–2010, and Ontario 1991–2010). IRR was calculated for residents who lived continuously in rural/urban households during each of the first 5 years of life.**

**RESULTS:** **There were 6,662 rural residents and 38,905 urban residents with IBD. Incidence of IBD per 100,000 was 33.16 (95% CI 27.24–39.08) in urban residents, and 30.72 (95% CI 23.81–37.64) in rural residents (IRR 0.90, 95% CI 0.81–0.99). The protective association was strongest in children <10 years (IRR 0.58, 95% CI 0.43–0.73) and 10–17.9 years (IRR 0.72, 95% CI 0.64–0.81). In the birth cohort, comprising 331 rural and 2,302 urban residents, rurality in the first 1–5 years of life was associated with lower risk of IBD (IRR 0.75–0.78).**

**CONCLUSIONS:** **People living in rural households had lower risk of developing IBD. This association is strongest in young children and adolescents, and in children exposed to the rural environment early in life.**

SUPPLEMENTARY MATERIAL is linked to the online version of the paper at http://www.nature.com/ajg

[1]Children's Hospital of Eastern Ontario IBD Centre, Division of Gastroenterology, Hepatology and Nutrition, Children's Hospital of Eastern Ontario, Ottawa, Canada; [2]Children's Hospital of Eastern Ontario Research Institute, Ottawa, Canada; [3]Department of Pediatrics, University of Ottawa, Ottawa, Canada; [4]School of Epidemiology and Public Health, University of Ottawa, Ottawa, Canada; [5]Institute for Clinical Evaluative Sciences, Toronto, Canada; [6]Department of Medicine, University of Calgary, Calgary, Canada; [7]Department of Community Health Sciences, University of Calgary, Calgary, Canada; [8]Department of Pediatrics, Dalhousie University, Halifax, Canada; [9]Mount Sinai Hospital Centre for Inflammatory Bowel Disease, Department of Medicine, University of Toronto, Toronto, Canada; [10]Department of Paediatrics, University of Toronto, Toronto, Canada; [11]Department of Medicine, Dalhousie University, Halifax, Canada; [12]University of Manitoba IBD Clinical and Research Centre, University of Manitoba, Winnipeg, Canada; [13]McGill University Health Centre, Division of Gastroenterology, Montreal, Canada; [14]Division of Pediatric Gastroenterology and Nutrition, Department of Pediatrics, University of Alberta, Edmonton, Canada; [15]Division of Gastroenterology, Department of Pediatrics, University of Calgary, Calgary, Canada; [16]School of Population and Public Health, University of British Columbia, Vancouver, Canada; [17]Department of Pediatrics and Child Health, University of Manitoba, Winnipeg, Canada; [18]Department of Pediatrics, University of British Columbia, Vancouver, Canada; [19]Child and Family Research Institute, University of British Columbia, Vancouver, Canada; [20]Department of Community Health Sciences, University of Manitoba, Winnipeg, Canada; [21]The Ottawa Hospital IBD Centre, Department of Medicine, University of Ottawa, Ottawa, Canada; [22]Department of Community Health and Epidemiology, University of Saskatchewan, Saskatoon, Canada; [23]Department of Internal Medicine, University of Manitoba, Winnipeg, Canada. **Correspondence:** Eric I. Benchimol, MD, PhD, Division of Gastroenterology, Hepatology and Nutrition, Children's Hospital of Eastern Ontario, 401 Smyth Road, Ottawa, Ontario, K1H 8L1, Canada. E-mail: ebenchimol@cheo.on.ca
**Received 10 March 2017; accepted 16 June 2017**

## INTRODUCTION

Over the past 50 years, the incidence of inflammatory bowel disease (IBD) has risen in various areas of the world, particularly in children living in developed nations and in newly industrialized countries experiencing increased urbanization (1–3). In addition to genetic risk factors, environmental factors have been associated with the risk of developing IBD. Early life exposure to these risk factors may be critical in IBD pathogenesis (4). Smaller family size (5,6), and early life exposure to antibiotics (7–9) are associated with higher risk of IBD, while early life exposure to farm animals (10) is associated with a lower risk of IBD.

Increased urbanization is one hypothesis for the rising incidence of IBD. Urban residence is associated with higher incidence of both Crohn's disease (CD) and ulcerative colitis (UC), as demonstrated by a systematic review and meta-analysis of 40 studies (11). However, significant heterogeneity in study design and results were reported. This heterogeneity may have partly resulted from including studies using different definitions of rurality. Moreover, the relationship between timing of exposure to an urban environment and age of IBD onset has not been assessed.

By determining the association between age of IBD onset and age of exposure to the rural/urban environment, researchers could potentially identify the gene-environment-microbe interaction to find causes of the disease, and intervene at the age of greatest benefit to prevent IBD. We conducted a population-based assessment of the risk of IBD in residents of Canada using validated health administrative data. We evaluated the incidence of IBD based on rural or urban residence. We determined the association of rural/urban residence with age at diagnosis. In addition, we assessed whether prolonged continuous exposure to the rural/urban environment after birth increased the risk of developing IBD later in life.

## METHODS

### Study design, setting, and participants

This study was approved by the Research Ethics Boards of the Children's Hospital of Eastern Ontario, The Ottawa Hospital, University of Calgary, University of Manitoba, IWK Health Centre, and Dalhousie University. This population-based, retrospective cohort study used routinely collected health administrative data from Alberta, Manitoba, Nova Scotia, and Ontario. These data contain all residents of each province who qualified for universal government health care insurance (>99% of the population). In 2011, these four provinces had a combined population of 19,231,900 people or 56.0% of the Canadian population (12).

We conducted two distinct studies. The first was a retrospective cohort study to determine the association between IBD incidence and rural/urban residence at the time of diagnosis. Residents diagnosed with IBD who were >6 months old at diagnosis were included (Alberta: 1999–2008; Manitoba: 1999–2010; Ontario: 1999–2010; and Nova Scotia: 2000–2008). We determined the incidence per 100,000 person-years of longitudinal follow-up time for IBD, CD, and UC by rural/urban residence in each year of the study period as well as in the overall study period. We further stratified persons by age group to examine age-specific associations between residential setting and IBD diagnosis. Full six-digit postal codes (Manitoba, Ontario), or the first three postal code digits (Alberta, Nova Scotia), were used to define rural/urban status. People who changed their residence from rural to urban or vice versa in the year prior to diagnosis were excluded from the study. We also excluded people who left their initial home province and then moved back in, due to missing data in the period they were not in their home province, and therefore our inability to determine their exact diagnosis date. We excluded residents who were missing date of birth, sex, or postal code in the data.

The second study was a birth cohort study to determine the association between length of early life exposure to the rural/urban environment and the subsequent development of IBD. This study included all residents where full administrative data and birth location were available in three participating provinces (Alberta: 1996–2010; Manitoba: 1988–2010; and Ontario: 1991–2010). We determined rural/urban status at birth, and on the first through fifth birthdays. We then determined the association between subsequent diagnosis with IBD, CD, and UC incidence and the length of exposure to the rural/urban environment in the first 5 years of life.

Due to provincial privacy laws, we were not permitted to pool individual-level data across provinces. Therefore, the Canadian Gastro-Intestinal Epidemiology Consortium (CanGIEC), a national network of investigators and analysts who use provincial health administrative data for research, conducted a distributed analysis in which the same research methods were applied to each provincial database. Summary results from each province were then meta-analyzed to produce multi-province estimates and to assess statistical heterogeneity.

### Data sources

We used the health administrative data from four Canadian provinces: Alberta, Manitoba, Nova Scotia, and Ontario. We used cohorts of all IBD patients living in each province identified from within health administrative data using validated algorithms of health care contacts to identify patients with IBD and classify their disease as CD, UC, or unclassifiable, and to distinguish incident from prevalent cases. These cohorts included the Alberta IBD Surveillance Cohort (13), the University of Manitoba IBD Epidemiology Database (14), the Ontario Crohn's and Colitis Cohort (15,16), and a patient cohort derived from Health Data Nova Scotia. The cohorts from Alberta, Manitoba, and Ontario used algorithms validated in their own province to identify and classify patients with IBD. The Nova Scotia cohort used the identification and classification algorithms validated in Manitoba. **Supplementary Table S1** online includes information on each cohort, the accuracy of the algorithms used, and information on the source, database, and study populations. This table also includes information on health administrative data used to identify subjects and derive their socio-demographic characteristics (date of birth, postal code). In the case of all four provinces, the entire IBD population was available to investigators for analysis. In Ontario and

Nova Scotia, the complete non-IBD population of the province was available to investigators. In Alberta and Manitoba, general population estimates from the 2001 and 2006 Canadian censuses were used to determine the denominator of incidence and prevalence rates, and to calculate inter-censal population estimates. We used 2006 Canadian Census estimates to standardize incidence rates by age and sex. Postal codes were taken from Statistics Canada's Postal Code Conversion File (PCCF) (17).

### Definition of rural/urban status

Residents were assigned a category of rural/urban based on their postal code at the time of diagnosis, or in each of their first five years of life for the birth cohort study. We assessed 14 different definitions of rural/urban status based on different spatial census units and five core definitions determined by Statistics Canada (18). We chose our main definition for rural/urban status because it most closely reflected the proportion of people in rural/urban households in Canada according to Statistics Canada (19). This definition used Metropolitan Area and Census Agglomeration Influenced Zones (MIZ), which gauges the level of influence that metropolitan areas exert upon non-metropolitan areas and incorporates that into their definitions of rurality (i.e., areas that are outside of cities but where a substantial proportion of the population commutes to the city for work would be considered urban). In the case of the main definition, rural/urban was considered dichotomous.

### Sensitivity analysis—other definitions of rural/urban status

Rurality in Canada has been defined in multiple distinct ways by Statistics Canada, based on population size, population density, or the economic and social influence of a city on neighboring regions (18). Although the most frequently used definition in Canada is based on both population size and density (19), other definitions may be employed in specific circumstances. We assigned rural/urban status of all residents of each province using 13 additional different definitions, based on previous work by Statistics Canada (see **Supplementary Table S2**) (18). This sensitivity analysis was conducted to assess the impact of using different definitions of rurality at the time of diagnosis with IBD on the incidence rate ratio (IRR). In addition to the definition used in the main analysis, other definitions that used MIZ included Definitions 2 and 3 which split rural/urban status into multiple categories. We also defined rurality by population density (Definitions 4 to 8). Definitions from the Organisation for Economic Co-operating and Development (OECD) for Rural Communities (Definitions 9–11) classified census divisions into 3 groups based on the percent of rural and non-rural areas of consolidated census divisions within each census division. Finally, we used Modified Beale Codes for Canadian Non-Metropolitan Analysis (Definitions 12–14).

In the case of Nova Scotia, due to the province's smaller population size, many of the definitions used in other provinces could not be applied. Therefore, for incidence estimates in the sensitivity analysis, only Definition 12 was applied to data for Nova Scotia residents.

### Statistical analysis

Descriptive statistics were reported as means with s.d., medians with interquartile range, or proportions with 95% confidence intervals (CIs) where appropriate. Sex-standardized incidence was calculated per 100,000 person-years of follow-up using 2006 Canadian census data, with 95% CI based on the gamma distribution. Incidence estimates were standardized by sex only, because age changed during the study period. To compare incidence between rural and urban residents, we calculated IRRs with 95% CI, adjusted for age and sex, using the urban population as the reference group. We also adjusted for mean neighborhood income quintile using 2001 Canadian census data and the Postal Code Conversion File Plus (20), a validated proxy for individual-level income (21). In the case of the sensitivity analyses, in definitions of rural/urban residence with more than two levels, the most urban level was taken as the reference group. We tested for significant differences in standardized incidence between groups using a Poisson regression analysis. We reported the incidence and IRR of IBD in rural/urban residents separately for each province.

In order to report a multi-province estimate and IRR, we conducted a meta-analysis across provinces using random-effects models. Meta-analysis of provincial data used stratified incidence rate differences and stratified IRR, which has been demonstrated to be an effective method of meta-analyzing rate data that accounts for heterogeneity (22). Heterogeneity was tested using the $I^2$ statistic and Cochrane $\chi^2$ test (Q test), which describe the percentage of total variation across incidence estimates due to heterogeneity rather than chance. Analyses were conducted using SAS version 9.3 (SAS Institute Inc., Cary, NC, USA). Meta-analysis was conducted using Stata release 14 (StataCorp LP, College Station, TX, USA). All tests of statistical significance were conducted using a nominal $\alpha = 0.05$.

## RESULTS

### Descriptive statistics

A total of 45,567 patients diagnosed with IBD were included in this study, consisting of 6,662 living in rural residences and 38,905 living in urban residences. Population descriptive statistics are presented in **Table 1**. Compared to urban patients, rural patients were diagnosed at an older age, and were more likely to be diagnosed in the earlier time periods.

The characteristics of the birth cohort of children who were exposed to the rural/urban environment for one continuous year from birth are presented in **Table 2**. This cohort comprised 331 rural patients and 2,302 urban patients diagnosed with IBD. Median age at diagnosis was older in rural patients, but there was no significant difference in sex distribution, diagnosis (CD, UC, or unclassified), or length of follow-up time between rural and urban patients in any of these exposure cohorts for any province.

### Incidence of IBD: rural/urban status at diagnosis

The overall incidence of IBD was 30.72 (95% CI 23.81–37.64) per 100,000 person-years in the rural population compared with 33.16

(95% CI 27.24–39.08) per 100,000 in the urban population (IRR 0.90, 95% CI 0.81–0.99; see **Figure 1a**). The protective effect of rurality was strongest in people with disease onset <10 years (IRR 0.58, 95% CI 0.43–0.73) and onset 10–17.9 years (IRR 0.72, 95% CI 0.64–0.81). There was no statistically significant association in patients with onset 18–39.9 years (IRR 0.96, 95% CI 0.86–1.06), 40–64.9 years (IRR 0.98, 95% CI 0.90–1.06), or ≥65 years (IRR 0.90, 95% CI 0.73–1.07).

The protective association of rural residence was similar in CD and UC (**Figure 1b,c**). In CD, the incidence was 14.88 (95% CI 10.83–18.92) per 100,000 in the rural population and 15.57 (95% CI 12.60–18.54) per 100,000 in the urban population. The protective effect of rurality was not statistically significant in the overall CD cohort (IRR 0.92, 95% CI 0.82–1.01). However, it was statistically significant in children <10 years (IRR 0.52, 95% CI 0.26–0.78) and children 10–17.9 years (IRR 0.70, 95% CI 0.61–0.79). In UC, the incidence was 13.83 (95% CI 10.61–17.05) per 100,000 in the rural population and 15.55 (95% CI 12.87–18.23) per 100,000 in the urban population. The protective effect of rurality was significant in the overall UC cohort (IRR 0.87, 95% CI 0.76–0.98). The association was strongest in children <10 years (IRR 0.58, 95% CI 0.32–0.85) and children 10–17.9 years (IRR 0.78, 95% CI 0.67–0.90), but also statistically significant in adults 18–39.9 years (IRR 0.86, 95% CI 0.80–0.92).

### Birth cohort study: rural/urban status at birth
When rural/urban status was determined at birth, rurality was associated with lower risk of IBD development later in life (**Figure 2**, see **Supplementary Data S2**). This effect remained stable with longer duration of continuous rural residence. The magnitude of the association was similar for patients with ≥1 years (IRR 0.77, 95% CI 0.58–0.96), ≥2 years (IRR 0.78, 95% CI 0.59–0.96), ≥3 years (IRR 0.77, 95% CI 0.54–1.01), ≥4 years (IRR 0.75, 95% CI 0.50–1.00), and ≥5 years (IRR 0.76, 95% CI 0.51–1.00) of continuous rural residence from birth. The protective effect was similar in CD (≥1 year: IRR 0.80, 95% CI 0.67–0.93) and UC (≥1 year: IRR 0.72, 95% CI 0.57–0.88).

### Sensitivity analysis evaluating different definitions of rurality
The sensitivity analyses demonstrated that the association between rural residence and incidence of IBD varied by the definition of rurality (**Table 3**, see **Supplementary Data S1**).

In the case of the MIZ definitions of rurality (Definitions 2–3), there was a trend towards a protective effect in more rural communities compared to the most urban levels (variable 99 in Definition 2 and variable 8 in Definition 3). However, due to smaller population sizes resulting from categorization into multiple levels of rurality, most of these associations were not statistically significant. The exception were children, in which most comparisons demonstrated lower incidence of IBD in more rural residences. For the population density definitions (Definitions 4–8), most demonstrated no statistically significant association between rural/urban status and incidence. Although the OECD definitions (Definitions 9–11) demonstrated inconsistent associations

between rural household and IBD, Definition 9 again demonstrated protection of rural household in childhood-onset disease. Of note, in Definition 11, while children aged 0–9 years in the most rural residences had lower incidence of IBD, those in the second level of the variable (those living in households between urban and rural) demonstrated a higher incidence of IBD compared to urban residents.

For the Beale Codes definitions of rurality (Definition 12–14), there was no statistically significant association between rural/urban residence and incidence of adult-onset IBD. However, for pediatric-onset disease (<10 years), there was a very strong association between the most rural postal codes and protection against IBD (Definition 12: IRR 0.65, 95% CI 0.40–0.90; Definition 13: IRR 0.56, 95% CI 0.26–0.85; Definition 14: IRR 0.56, 95% CI 0.26–0.85).

## DISCUSSION
Using population-based data from multiple Canadian provinces, we demonstrated that rural residence at diagnosis and at birth was associated with a lower incidence of IBD compared to urban residence, with the strongest protective association in childhood-onset IBD. Varying the definition of rural/urban resulted in varying degrees of association and even resulted in a lack of association for some definitions. However, most definitions demonstrated the protective effect of rural residence on pediatric IBD, particularly in children <10 years. Exposure to the rural environment from birth was consistently associated with a strong protective association with the development of IBD later in life, whether children were exposed continuously for one to five years from birth. These findings demonstrate the importance of early life exposure in altering the risk of IBD, the greater magnitude of effect of this environmental risk factor on the risk of childhood-onset disease, and the importance of adequately defining rurality.

Our findings were consistent with the findings of a systematic review of 40 articles (11). The authors found that urban residence was significantly associated with risk of IBD, and the effect was stronger for CD (IRR 1.42, 95% CI 1.26–1.60) than for UC (IRR 1.16, 95% CI 1.03–1.32) (11). We demonstrated that the association between rural/urban residence and IBD was similar for CD and UC. The authors of the systematic review noted a high degree of statistical heterogeneity amongst study results, potentially resulting due to the use of different definitions of rurality. A likely factor that contributes to this heterogeneity is the use of various spatial census units required by the different rural/urban definitions, as results can vary in geographic analysis when multiple spatial units are employed (23). We also found that even within Canada, using the same definition of rurality, there was heterogeneity in risk across different provinces. However, the finding that rural residence was protective for IBD in children was consistent across most definitions. This finding indicates the importance of early life environmental risk factors in the development of IBD (4), and is consistent with other studies demonstrating that antibiotics (7–9) and air pollution (24) in early life

INFLAMMATORY BOWEL DISEASE

**Table 1. Descriptive characteristics of the cohort of patients diagnosed with IBD during the study period**

| | Alberta | | | Manitoba | | | Ontario | | | Nova Scotia | | | Overall | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Rural (n=1,879) | Urban (n=7,594) | P value | Rural (n=909) | Urban (n=2,483) | P value | Rural (n=2,506) | Urban (n=26,271) | P value | Rural (n=1,368) | Urban (n=2,557) | P value | Rural (n=6,662) | Urban (n=38,905) | P value |
| Age at diagnosis, mean±s.d. | 41.5±18.7 | 40.0±18.6 | 0.002 | 40.8±18.2 | 39.6±18.3 | 0.11 | 44.1±18.3 | 39.6±18.1 | <0.0001 | 44.6±19.3 | 42.7±18.7 | 0.003 | 42.7±18.6 | 40.2±13.4 | <0.0001 |
| Age at diagnosis, median (IQR) | 40 (26–54) | 38 (25–53) | 0.002 | 39 (26–54) | 37 (25–52) | 0.08 | 44.9 (29.1–57.9) | 38.0 (24.3–52.6) | <0.0001 | 44 (29–59) | 42 (28–56) | 0.004 | 41.9 (22.5–61.2) | 38.7 (20.0–57.5) | <0.0001 |
| **Age at diagnosis** | | | | | | | | | | | | | | | |
| 0–9y | 29 (1.5%) | 165 (2.2%) | 0.04 | 9 (1.0%) | 36 (1.5%) | 0.27 | 18 (0.7%) | 298 (1.1%) | <0.0001 | 18 (1.3%) | 35 (1.4%) | 0.10 | 74 (1.1%) | 534 (1.4%) | <0.0001 |
| 10–17 years | 139 (7.4%) | 594 (7.8%) | | 72 (7.9%) | 210 (8.5%) | | 208 (8.3%) | 2,692 (10.3%) | | 92 (6.7%) | 198 (7.7%) | | 511 (7.7%) | 3,694 (9.5%) | |
| 18–39 years | 743 (39.5%) | 3,190 (42.0%) | | 376 (41.4%) | 1,102 (44.4%) | | 791 (31.6%) | 11,111 (42.3%) | | 456 (33.3%) | 934 (36.5%) | | 2,366 (35.5%) | 16,337 (42.0%) | |
| 40–64 years | 715 (38.1%) | 2,741 (36.1%) | | 343 (37.7%) | 849 (34.2%) | | 1,143 (45.6%) | 9,588 (36.5%) | | 578 (42.3%) | 1,029 (40.2%) | | 2,779 (41.7%) | 14,207 (36.5%) | |
| 65+ years | 253 (13.5%) | 904 (11.9%) | | 109 (12.0%) | 286 (11.5%) | | 346 (13.8%) | 2,582 (9.8%) | | 224 (16.4%) | 361 (14.1%) | | 932 (14.0%) | 4,133 (10.6%) | |
| Female, N (%) | 980 (52.2%) | 4,023 (53.0%) | 0.53 | 495 (54.5%) | 1,297 (52.2%) | 0.27 | 1,298 (51.8%) | 13,548 (51.6%) | 0.83 | 762 (55.7%) | 1,396 (54.6%) | 0.51 | 3,535 (53.1%) | 20,264 (52.1%) | 0.14 |
| Longitudinal follow-up time, mean±s.d. | 9.4±2.8 | 9.4±2.9 | 0.59 | 8.7±3.8 | 8.60±3.68 | 0.42 | 11.9±0.5 | 11.9±0.8 | 0.0001 | 4.8±3.2 | 5.0±3.1 | 0.17 | 9.7±3.4 | 9.7±3.3 | 0.26 |
| **Diagnosis, n (%)** | | | | | | | | | | | | | | | |
| Crohn's | 897 (47.7%) | 3,491 (46.0%) | 0.04 | 400 (44.0%) | 1,118 (45.0%) | 0.62 | 1,066 (42.5%) | 11,969 (45.6%) | 0.011 | 711 (52.0%) | 1,286 (50.3%) | 0.32 | 3,074 (46.1%) | 17,864 (45.9%) | 0.43 |
| Ulcerative colitis | 616 (32.8%) | 2,705 (35.6%) | | 509 (56.0%) | 1,365 (55.0%) | | 1,274 (50.8%) | 12,746 (48.5%) | | 657 (48.0%) | 1,271 (49.7%) | | 3,056 (45.9%) | 18,087 (46.5%) | |
| IBD type unclassifiable | 366 (19.5%) | 1,398 (18.4%) | | 0 | 0 | | 166 (6.6%) | 1,556 (5.9%) | | NA | NA | | 532 (8.0%) | 2,954 (7.6%) | |
| **Diagnosis time period, n (%)** | | | | | | | | | | | | | | | |
| 1999–2003 | 997 (53.1%) | 3,752 (49.4%) | 0.005 | 427 (46.0%) | 1,076 (43.3%) | 0.04 | 982 (39.2%) | 9,302 (35.4%) | <0.0001 | 701 (51.2%) | 1,354 (53.0%) | 0.31 | 3,107 (46.6%) | 15,484 (39.8%) | <0.0001 |
| 2004–2008 | 882 (46.9%) | 3,842 (50.6%) | | 262 (28.8%) | 828 (33.4%) | | 874 (34.9%) | 9,169 (34.9%) | | 667 (48.8%) | 1,203 (47.0%) | | 2,685 (40.3%) | 15,042 (38.7%) | |
| 2009–2010 | NA | NA | | 220 (24.2%) | 579 (23.3%) | | 650 (25.9%) | 7,800 (29.7%) | | NA | NA | | 870 (13.1%) | 8,379 (21.5%) | |

Bold values represent the proportion of residents with IBD in whom the IBD type could not be classified (based on administrative data algorithms) as either Crohn's or ulcerative colitis.

**Table 2. Descriptive characteristics of the birth cohort of patients diagnosed with IBD with at least 1 year of continuous exposure to the rural/urban environment**

| | Alberta | | | Manitoba | | | Ontario | | | Overall | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Rural (n=65) | Urban (n=233) | P value | Rural (n=47) | Urban (n=149) | P value | Rural (n=219) | Urban (n=1,920) | P value | Rural (n=331) | Urban (n=2,302) | P value |
| Age at diagnosis, mean±s.d. | 7.1±4.6 | 6.7±4.3 | 0.56 | 13.4±4.0 | 13.2±3.8 | 0.07 | 12.2±4.4 | 11.6±4.2 | 0.06 | 11.1±4.3 | 10.8±4.1 | 0.12 |
| Age at diagnosis median (IQR) | 7 (3–11) | 7 (3–11) | 1.0 | 14 (11–16) | 13 (11–16) | 0.66 | 13 (10–15) | 12 (9–15) | | 12.4 (8.7–16.2) | 11.5 (7.5–15.6) | <0.0001 |
| Female, N (%) | 28 (43.1%) | 117 (50.2%) | 0.31 | 22 (46.8%) | 73 (49.0%) | 0.87 | 90 (41.1%) | 839 (43.7%) | 0.62 | 140 (42.3%) | 1,029 (44.7%) | 0.41 |
| Longitudinal follow-up time, mean±s.d. | 7.2±4.6 | 6.1±4.3 | 0.09 | 6.5±3.1 | 7.5±3.1 | 0.08 | 15.6±3.7 | 15.5±3.6 | 0.57 | 10.8±6.4 | 11.3±5.7 | 0.21 |
| *Continuous period in rural/urban (%)* | | | | | | | | | | | | |
| 1 year | 65 (100%) | 233 (100%) | NA | 47 (100%) | 149 (100%) | NA | 219 (100%) | 1,920 (100%) | NA | 331 (100%) | 2,302 (100%) | NA |
| 2 year | 65 (100%) | 233 (100%) | | 47 (100%) | 148 (99.3%) | | 208 (95.0%) | 1,902 (99.0%) | | 320 (96.7%) | 2,283 (99.2%) | |
| 3 year | 65 (100%) | 229 (98.3%) | | 40 (85.1%) | 146 (98.0%) | | 200 (91.3%) | 1,851 (96.4%) | | 305 (92.1%) | 2,226 (96.7%) | |
| 4 year | 62 (95.4%) | 227 (97.4%) | | 36 (76.6%) | 146 (98.0%) | | 185 (84.4%) | 1,806 (94.1%) | | 283 (85.5%) | 2,179 (95.0%) | |
| 5 year | 60 (92.3%) | 218 (93.6%) | | 36 (76.6%) | 145 (97.3%) | | 172 (78.5%) | 1,746 (90.9%) | | 268 (81.0%) | 2,109 (91.6%) | |
| *Diagnosis, n (%)* | | | | | | | | | | | | |
| Crohn's | 31 (47.7%) | 118 (50.6%) | 0.22 | 27 (57.5%) | 81 (54.4%) | 0.74 | 129 (58.9%) | 1,059 (55.2%) | 0.29 | 187 (56.5%) | 1,258 (54.6%) | 0.89 |
| Ulcerative colitis | 19 (29.2%) | 74 (31.8%) | | 20 (42.5%) | 68 (45.6%) | | 79 (36.1%) | 713 (37.1%) | | 118 (35.6%) | 855 (37.1%) | |
| IBD type unclassifiable | 15 (23.1%) | 41 (17.6%) | | NA | NA | | 11 (5.0%) | 148 (7.7%) | | 26 (7.9%) | 189 (8.2%) | |
| *Diagnosis time period, n (%)* | | | | | | | | | | | | |
| 1994–2003[a] | 17 (28.3%) | 61 (27.6%) | 0.62 | 7 (14.9%) | 25 (17.0%) | 0.36 | 22 (10.0%) | 255 (13.3%) | 0.55 | 46 | 341 | NA |
| 2004–2008 | 43 (71.7%) | 160 (72.4%) | | 15 (31.9%) | 61 (41.5%) | | 109 (49.8%) | 948 (49.4%) | | 167 | 1,169 | |
| 2009–2010 | NA | NA | | 25 (53.2%) | 61 (41.5%) | | 88 (40.2%) | 717 (37.3%) | | 113 | 778 | |

[a]Due to small cell sizes, descriptive characteristics of the birth cohort diagnosed 1994–2003 were combined in this table for Manitoba and Ontario. Alberta's birth cohort included only patients diagnosed 1999–2003.

INFLAMMATORY BOWEL DISEASE

INFLAMMATORY BOWEL DISEASE
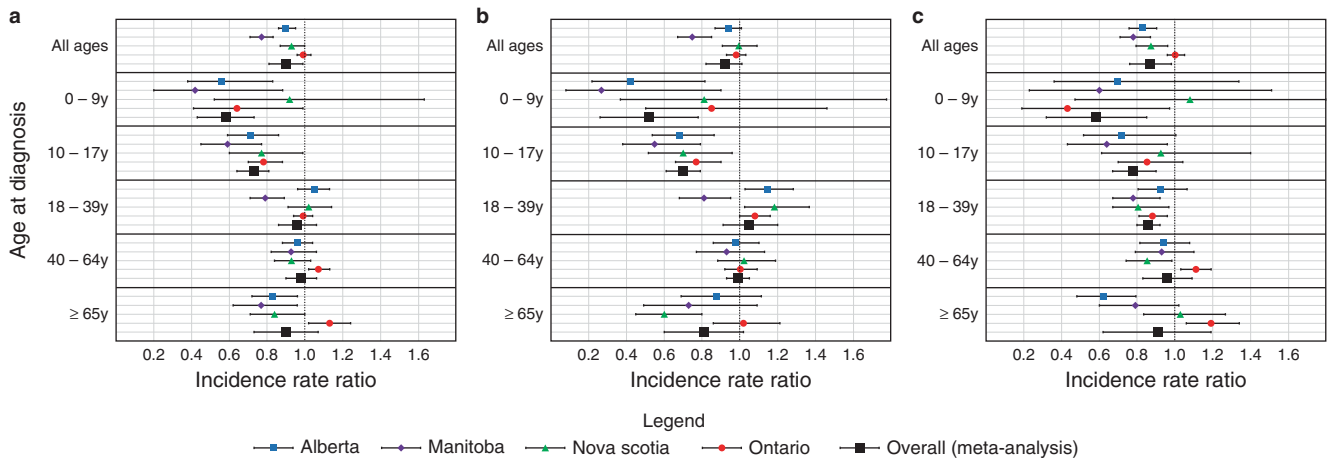


**Figure 1.** Incidence rate ratio for (**a**) IBD, (**b**) Crohn's disease, and (**c**) ulcerative colitis based on rural/urban status at the time of diagnosis.
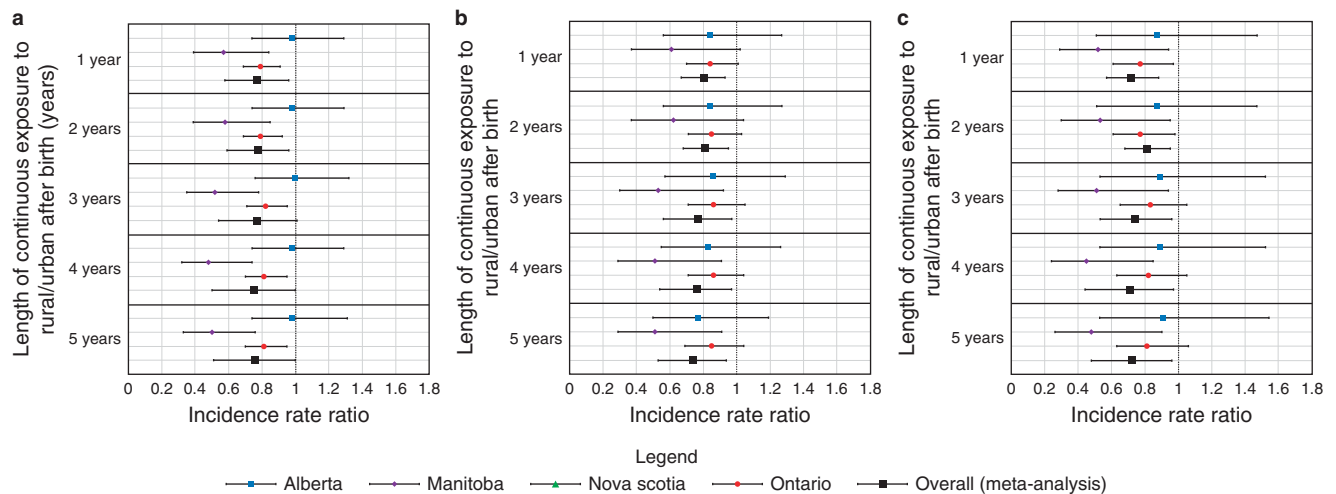


**Figure 2.** Incidence rate ratio for (**a**) IBD, (**b**) Crohn's disease, and (**c**) ulcerative colitis based on rural/urban status at birth.

was associated with increased risk of IBD in late childhood or early adulthood, while early life exposure to farm animals was associated with decreased risk (10). Therefore, identification of environmental associations with IBD may be more successful in children.

The mechanism by which rurality protects against IBD is uncertain, and may include dietary and lifestyle factors, environmental exposures, or segregation of individuals with different genetic risk profiles. A similar protective effect of rurality against the development of asthma has been observed in cohort studies (25). In a mouse model of asthma, farm dust or low-dose endotoxin reduced epithelial cell cytokines that activate dendritic cells, thereby suppressing type 2 immunity and house dust mite-induced asthma (26). This suppression results from induction of ubiquitin-modifying enzyme A20. In the gut, the microbiome may be involved in inducing A20, resulting in suppression of inflammatory reactions to commensal bacteria (27). This toler-

ance to a healthy microbiome composition may be missing or inflammation stimulated by the absence of farm dust or another environmental factor in the urban environment, resulting increased risk of IBD.

The protective effect of rural residence on the risk of IBD was strongest for Manitoba (IRR 0.77, 95% CI 0.71–0.83), but was not significant in Ontario (IRR 0.99, 95% CI 0.96–1.03), with Alberta and Nova Scotia falling between. This heterogeneity may have resulted from true differences in risk of IBD based on differential exposures across rural locations. For example, Canadian farms vary by province, with beef farmed on 18.9% of Nova Scotia farms, 19.3% of Ontario farms, 34.6% of Manitoba farms, and 41.5% of Alberta farms (28). By contrast, Ontario has far more soy (10.2%) and hay (10.3%) farming, Alberta and Manitoba have more grain farming (13.1% and 19.0%, respectively), and Nova Scotia farms more fruit and tree-nuts (23.9%) (28). A previous study from Germany reported that early life exposure to farm animals resulted

**Table 3.** Color representation of strength of association between rural/urban household and incidence of IBD by definition of rurality

| | MIZ Definitions | | | Population Density Definitions | | | | | OECD Definitions | | | Beale Code Definitions | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Main | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| All Ages | * | 1 vs urb: --<br>2 vs urb: --<br>3 vs urb: --<br>4 vs urb: -- | 2 vs urb: --<br>3 vs urb: --<br>4 vs urb: --<br>5 vs urb: --<br>6 vs urb: --<br>7 vs urb: -- | -- | -- | -- | -- | -- | -- | 1 vs urb: -- | 1 vs urb: --<br>2 vs urb: ^ | -- | 1 vs urb: --<br>3 vs urb: --<br>4 vs urb: --<br>5 vs urb: --<br>6 vs urb: ^ | 1 vs urb: --<br>4 vs urb: --<br>6 vs urb: --<br>7 vs urb: --<br>8 vs urb: ^ |
| 0-9y | ** | 1 vs urb: --<br>2 vs urb: ***<br>3 vs urb: **<br>4 vs urb: *** | 2 vs urb: missing<br>3 vs urb: ***<br>4 vs urb: **<br>5 vs urb: ***<br>6 vs urb: --<br>7 vs urb: -- | -- | -- | -- | -- | -- | ** | 1 vs urb: -- | 1 vs urb: **<br>2 vs urb: -- | ** | 1 vs urb: **<br>3 vs urb: --<br>4 vs urb: --<br>5 vs urb: --<br>6 vs urb: -- | 1 vs urb: **<br>4 vs urb: --<br>6 vs urb: **<br>7 vs urb: --<br>8 vs urb: -- |
| 10-17y | ** | 1 vs urb: **<br>2 vs urb: **<br>3 vs urb: **<br>4 vs urb: ** | 2 vs urb: **<br>3 vs urb: **<br>4 vs urb: **<br>5 vs urb: **<br>6 vs urb: --<br>7 vs urb: -- | -- | -- | -- | -- | -- | ** | 1 vs urb: -- | 1 vs urb: --<br>2 vs urb: ^ | ** | 1 vs urb: **<br>3 vs urb: **<br>4 vs urb: --<br>5 vs urb: --<br>6 vs urb: -- | 1 vs urb: **<br>4 vs urb: **<br>6 vs urb: --<br>7 vs urb: --<br>8 vs urb: -- |
| 18-39y | -- | 1 vs urb: --<br>2 vs urb: --<br>3 vs urb: --<br>4 vs urb: -- | 2 vs urb: --<br>3 vs urb: --<br>4 vs urb: --<br>5 vs urb: --<br>6 vs urb: ^<br>7 vs urb: -- | -- | -- | -- | -- | -- | -- | 1 vs urb: -- | 1 vs urb: --<br>2 vs urb: ^^ | -- | 1 vs urb: --<br>3 vs urb: --<br>4 vs urb: --<br>5 vs urb: ^^<br>6 vs urb: ^^ | 1 vs urb: --<br>4 vs urb: --<br>6 vs urb: --<br>7 vs urb: ^^<br>8 vs urb: ^^ |
| 40-64y | -- | 1 vs urb: --<br>2 vs urb: --<br>3 vs urb: --<br>4 vs urb: -- | 2 vs urb: --<br>3 vs urb: --<br>4 vs urb: --<br>5 vs urb: --<br>6 vs urb: ^<br>7 vs urb: -- | -- | -- | -- | -- | -- | -- | 1 vs urb: -- | 1 vs urb: --<br>2 vs urb: ^^ | -- | 1 vs urb: --<br>3 vs urb: --<br>4 vs urb: --<br>5 vs urb: ^^<br>6 vs urb: ^^ | 1 vs urb: --<br>4 vs urb: --<br>6 vs urb: --<br>7 vs urb: ^^<br>8 vs urb: ^^ |
| ≥65y | -- | 1 vs urb: --<br>2 vs urb: --<br>3 vs urb: --<br>4 vs urb: -- | 2 vs urb: --<br>3 vs urb: --<br>4 vs urb: --<br>5 vs urb: --<br>6 vs urb: --<br>7 vs urb: ^ | ^ | -- | -- | -- | -- | -- | 1 vs urb: -- | 1 vs urb: --<br>2 vs urb: ^ | -- | 1 vs urb: --<br>3 vs urb: --<br>4 vs urb: --<br>5 vs urb: --<br>6 vs urb: ^ | 1 vs urb: --<br>4 vs urb: --<br>6 vs urb: --<br>7 vs urb: --<br>8 vs urb: ^ |

**Color coding**

| | |
|---|---|
| -- | P>0.05 – no significant association between rural/urban household and IBD incidence |
| * | IRR 0.80-0.99, P<0.05 – rural residence associated with lower incidence of IBD |
| ** | IRR 0.50-0.79, P<0.05 – rural residence associated with lower incidence of IBD |
| *** | IRR 0.25-0.49, P<0.05 – rural residence associated with lower incidence of IBD |
| ^ | IRR 1.01-1.20, P<0.05 – rural residence associated with higher incidence of IBD |
| ^^ | IRR 1.21-1.49, P<0.05 – rural residence associated with higher incidence of IBD |

IRR, incidence rate ratio; MIZ, metropolitan influenced zones; OECD, Organisation for Economic Co-operation and Development; urb, most urban level of variable.

In definitions with multiple levels of rurality, comparison is presented with the reference group being the most urban group. The lower numbered levels of each variable indicate more rural regions. See **Supplementary Table 2** for a detailed description of each definition.

INFLAMMATORY BOWEL DISEASE

in strong protection against later development of CD. This may explain the stronger protective effect of rural residence in Alberta and Manitoba, where more beef farms exist, compared with Ontario and Nova Scotia. Nevertheless, farming is only one aspect of rural life. Air quality, water sources, and dietary differences may contribute to differences between urban and rural incidence of IBD especially for children, and these factors vary by province. These effects may be stronger in children because their gut microbiome is in evolution and may be vulnerable to changes in the first two years of life (29).

This study is limited by factors that affect all research using health administrative data. Misclassification bias may have resulted from different algorithms used to identify incident IBD in different provinces. However, the algorithms used in Alberta, Manitoba, and Ontario were validated as the most accurate for use in provinces to which they were applied (13–16), and the validation studies did not demonstrate differential accuracy of the algorithms in rural and urban patients. Nova Scotia used the algorithm validated in Manitoba. One of our hypotheses included the idea that results may be different depending on which definition of rural/urban was used. We therefore applied multiple definitions which indeed resulted in different magnitude of risk or protection depending on the definition. This finding may also have resulted from misclassification of rural/urban status. However, we used definitions previously published and used by Statistics Canada and other international organizations (18). In addition, our main analysis used a classification which most closely resembled

Statistics Canada reports of rural to urban residence ratios. There is no perfect, agreed upon definition of rurality, and other investigators should be aware of the variation possible when different definitions are used to test the association between rural/urban residence and IBD.

It is possible that important confounding factors could not be accounted for using our health administrative data. However, we controlled for mean neighborhood income, which is a validated proxy of residence income and social deprivation (21,30). We were not able to determine whether ethnicity, genotype, phenotype, disease severity, or family history of IBD was different in rural and urban residents. For example, there is variation in the proportion of Jewish residents (who are predominantly urban dwellers) in each included province, however this population is relatively small (Alberta: 0.37%, Manitoba: 1.2%, Nova Scotia: 0.24%, and Ontario: 1.7%) (31). We have previously demonstrated that immigrants, who tend to settle in cities, had a lower risk of IBD and therefore immigration status would not explain the higher risk seen in urban residents (32). Similarly, we did not have access to smoking status. However, smoking rates are similar in Canadians living in urban and rural residences (21.1 vs. 22.7%) (33). Another unmeasured variable may have contributed to the rural/urban disparity, such as exposure to pollution, vitamin D and sunlight, or differences in diet. One possible explanation for our findings is that rural patients had poor access to specialist care for diagnosis. Health administrative data can only identify patients who sought medical attention for IBD. If rural children with IBD were less likely to visit a physician, be investigated or receive a diagnosis of IBD compared to urban children, and this gap became less prominent with older age, this might explain the dramatically lower incidence of IBD diagnosis in rural children. A future study will examine disparities in health care access and care in rural/urban Canadians with IBD. In addition, we expect researchers will work to understand exactly what environmental differences exist in rural and urban Canadians to explain the altered risk observed. These differences may contribute to alterations in the intestinal microbiome or epigenetic changes. Nevertheless, this is the largest study to date to examine rural/urban residence association with incidence of IBD, and the first to test various definitions of rurality. In general, the protective effect of rurality on early onset IBD in children <18 years was consistent across most provinces and definitions of rurality.

## CONCLUSION

In summary, we found that people living in rural residences were less likely to be diagnosed with IBD compared to those living in urban residences, particularly in pediatric-onset disease. This effect was significant in both forms of IBD in children. In addition, early life rural residence (in the first five years of life) was strongly protective against the subsequent development of IBD, with a similar strength of association in children with one to five years of continuous exposure. This implies the establishment of risk in the first year of life, a critical period of intestinal

microbiome differentiation. Future research should determine whether urban residence is associated with a microbiome profile of increased risk of IBD, and whether modification of specific risk factors associated with city living (such as diet, pollution, physical inactivity, or exposure to pathogens) may ameliorate the risk.

### CONFLICT OF INTEREST
**Guarantor of the article:** Eric I. Benchimol, MD, PhD.
**Specific author contributions:** Eric I. Benchimol: study conception and design, analysis and interpretation of data, drafting of manuscript, statistical analysis, and obtained funding; Gilaad G. Kaplan: study conception and design, establishment of Alberta cohort, analysis and interpretation of data, critical revision of the manuscript, statistical analysis, and obtained funding; Anthony R. Otley: establishment of Nova Scotia cohort, analysis and interpretation of data, critical revision of the manuscript, and statistical analysis; Geoffrey C. Nguyen: study conception and design, analysis and interpretation of data, critical revision of the manuscript, statistical analysis, and obtained funding; Fox E. Underwood: health geography expertise, classification of rural/urban postal codes, analysis and interpretation of data, and critical revision of the manuscript; Astrid Guttmann: study conception and design, analysis and interpretation of data, critical revision of the manuscript, statistical analysis, and obtained funding; Jennifer L. Jones: establishment of Nova Scotia cohort, study conception and design, analysis and interpretation of data, critical revision of the manuscript, statistical analysis, and obtained funding; Beth K. Potter: study conception and design, analysis and interpretation of data, critical revision of the manuscript, statistical analysis, and obtained funding; Christina A. Catley: Ontario data methodologist/analyst, analysis and interpretation of results, statistical analysis, technical support, and critical revision of the manuscript; Zoann Nugent: Manitoba data methodologist/analyst, analysis and interpretation of results, statistical analysis, technical support, and critical revision of the manuscript; Yunsong Cui: Nova Scotia data methodologist/analyst, analysis and interpretation of

## Study Highlights

### WHAT IS CURRENT KNOWLEDGE

✓ Incidence of inflammatory bowel disease (IBD) is higher in residents of urban communities, compared to rural communities.

✓ There has been significant heterogeneity in results of studies assessing the association between rural/urban household and IBD.

### WHAT IS NEW HERE

✓ In Canada, the association between rural/urban household and IBD was strongest for children <10 years and adolescents 10–18 years, while the association was not significant for adult- and elderly-onset IBD.

✓ Early life exposure to the rural environment in the first 5 years of life was strongly associated with lower risk of IBD later in life.

✓ The strength of association depended on the definition of rurality used, however the association was consistently significant in childhood-onset IBD using most definitions.

## REFERENCES

1. Benchimol EI, Fortinsky KJ, Gozdyra P et al. Epidemiology of pediatric inflammatory bowel disease: a systematic review of international trends. Inflamm Bowel Dis 2011;17:423–39.
2. Molodecky NA, Soon IS, Rabi DM et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. Gastroenterology 2012;142:46–54.e42.
3. Kaplan GG. The global burden of IBD: from 2015 to 2025. Nat Rev Gastroenterol Hepatol 2015;12:720–7.
4. Aujnarain A, Mack DR, Benchimol EI. The role of the environment in the development of pediatric inflammatory bowel disease. Curr Gastroenterol Rep 2013;15:326.
5. Bernstein CN, Rawsthorne P, Cheang M et al. A population-based case control study of potential risk factors for IBD. Am J Gastroenterol 2006;101:993–1002.
6. Amre DK, Lambrette P, Law L et al. Investigating the hygiene hypothesis as a risk factor in pediatric onset Crohn's disease: a case-control study. Am J Gastroenterol 2006;101:1005–11.
7. Shaw SY, Blanchard JF, Bernstein CN. Association between the use of antibiotics in the first year of life and pediatric inflammatory bowel disease. Am J Gastroenterol 2010;105:2687–92.
8. Shaw SY, Blanchard JF, Bernstein CN. Association between the use of antibiotics and new diagnoses of Crohn's disease and ulcerative colitis. Am J Gastroenterol 2011;106:2133–42.
9. Virta L, Auvinen A, Helenius H et al. Association of repeated exposure to antibiotics with the development of pediatric Crohn's disease--a nationwide, register-based finnish case-control study. Am J Epidemiol 2012;175:775–84.
10. Radon K, Windstetter D, Poluda AL et al. Contact with farm animals in early life and juvenile inflammatory bowel disease: a case-control study. Pediatrics 2007;120:354–61.
11. Soon IS, Molodecky NA, Rabi DM et al. The relationship between urban environment and the inflammatory bowel diseases: a systematic review and meta-analysis. BMC Gastroenterol 2012;12:51.
12. Statistics Canada. Population and dwelling counts, for Canada, provinces and territories, 2011 and 2006 censuses. 2012. Available at: http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/hlt-fst/pd-pl/Table-Tableau.cfm?LANG=Eng&T=101&S=50&O=A. Accessed on 8 January 2016.
13. Rezaie A, Quan H, Fedorak RN et al. Development and validation of an administrative case definition for inflammatory bowel diseases. Can J Gastroenterol 2012;26:711–7.

INFLAMMATORY BOWEL DISEASE

INFLAMMATORY BOWEL DISEASE

14. Bernstein CN, Blanchard JF, Rawsthorne P *et al.* Epidemiology of Crohn's disease and ulcerative colitis in a central Canadian province: a population-based study. Am J Epidemiol 1999;149:916–24.

15. Benchimol EI, Guttmann A, Griffiths AM *et al.* Increasing incidence of paediatric inflammatory bowel disease in Ontario, Canada: evidence from health administrative data. Gut 2009;58:1490–7.

16. Benchimol EI, Guttmann A, Mack DR *et al.* Validation of international algorithms to identify adults with inflammatory bowel disease in health administrative data from Ontario, Canada. J Clin Epidemiol 2014;67:887–96.

17. Postal CodeOM Conversion File (PCCF), Reference Guide, 2013 Statistics Canada: Ottawa, 2013.

18. du Plessis V, Beshiri R, Bollman RD *et al.* Definitions of "Rural" Statistics Canada: Ottawa, 2002, December 1. Report No.: 21-601-MIE—No. 061

19. Statistics Canada. Population, urban and rural, by province and territory. (Webpage) 4 Feb 2011. Available at: http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/demo62a-eng.htm. Accessed on 8 January 2016.

20. Wilkins R. PCCF+ Version 4E User's Guide: Automated Geographic Coding Based on the Statistics Canada Postal Code Convesion Files, Including Postal Codes to July 2004 Health Analysis and Measurement Group, Statistics Canada: Ottawa, Canada, Catalogue No. 93-387-XIE 2005.

21. Glazier RH, Creatore MI, Agha MM *et al.* Socioeconomic misclassification in Ontario's Health Care Registry. Can J Public Health 2003;94:140–3.

22. Guevara JP, Berlin JA, Wolf FM. Meta-analytic methods for pooling rates when follow-up duration varies: a case study. BMC Med Res Methodol 2004;4:17.

23. Openshaw S. The Modifiable Areal Unit Problem Geo Books: Newcastle, UK, 1983.

24. Kaplan GG, Hubbard J, Korzenik J *et al.* The inflammatory bowel diseases and ambient air pollution: a novel association. Am J Gastroenterol 2010;105:2412–9.

25. Lawson JA, Chu LM, Rennie DC *et al.* Prevalence, risk factors, and clinical outcomes of atopic and nonatopic asthma among rural children. Ann Allergy Asthma Immunol 2017;118:304–10.

26. Schuijs MJ, Willart MA, Vergote K *et al.* Farm dust and endotoxin protect against allergy through A20 induction in lung epithelial cells. Science 2015;349:1106–10.

27. Wang J, Ouyang Y, Guner Y *et al.* Ubiquitin-editing enzyme A20 promotes tolerance to lipopolysaccharide in enterocytes. J Immunol 2009;183:1384–92.

28. Statistics Canada. Snapshot of Canadian agriculture (2006). [Webpage] 24 March 2014. Available at: http://www.statcan.gc.ca/ca-ra2006/articles/snapshot-portrait-eng.htm. Accessed on 6 March 2017.

29. Smith MI, Yatsunenko T, Manary MJ *et al.* Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. Science 2013;339:548–54.

30. Smith GD, Hart C, Watt G *et al.* Individual social class, area-based deprivation, cardiovascular disease risk factors, and mortality: the Renfrew and Paisley Study. J Epidemiol Community Health 1998;52:399–405.

31. Population by religion, by province and territory (2001 Census). [Webpage] 25 Jan 2005. Available at: http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/demo30a-eng.htm. Accessed on 28 May 2017.

32. Benchimol EI, Mack DR, Guttmann A *et al.* Inflammatory bowel disease in immigrants to Canada and their children: a population-based cohort study. Am J Gastroenterol 2015;110:553–63.

33. Statistics Canada. Smoking 2008. [Webpage] 27 November 2015. Available at: http://www.statcan.gc.ca/pub/82-625-x/2010001/article/11110-eng.htm. Accessed on 2 December 2016.