Journal of Clinical Epidemiology 149 (2022) 23–35

Journal of Clinical Epidemiology

# Meta-analysis of multi-jurisdictional health administrative data from distributed networks approximated individual-level multivariable regression

Aman K. Dheri[a,b], M. Ellen Kuenzig[c,d], David R. Mack[a,e,f], Sanjay K. Murthy[b,g,h], Gilaad G. Kaplan[i], Jessy Donelle[g], Glenys Smith[g], Eric I. Benchimol[a,b,c,d,e,f,g,j,*]

[a]*Children's Hospital of Eastern Ontario (CHEO) Inflammatory Bowel Disease Centre, Division of Gastroenterology, Hepatology and Nutrition, CHEO, Ottawa, Ontario, Canada*

[b]*School of Epidemiology and Public Health, University of Ottawa, Ottawa, Ontario, Canada*

[c]*SickKids Inflammatory Bowel Disease Center, Division of Gastroenterology, Hepatology and Nutrition, The Hospital for Sick Children, Toronto, Ontario Canada*

[d]*Child Health Evaluative Sciences, SickKids Research Institute, The Hospital for Sick Children, Toronto, Ontario, Canada*

[e]*CHEO Research Institute, Ottawa, Ontario, Canada*

[f]*Department of Pediatrics, University of Ottawa, Ottawa, Ontario, Canada*

[g]*ICES, Toronto, Ontario, Canada*

[h]*Department of Medicine, University of Ottawa, Ottawa, Ontario, Canada*

[i]*Departments of Medicine and Community Health Sciences, University of Calgary, Calgary, Alberta, Canada*

[j]*Department of Paediatrics, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada*

Accepted 16 May 2022; Published online 20 May 2022

## Abstract

**Background and Objectives:** Compare meta-analysis in a distributed network to individual-level analysis for assessment of time trends of health services utilization with health administrative data.

**Methods:** We used administrative data from Ontario, Canada to analyze temporal trends in pediatric inflammatory bowel disease health services use. Beta coefficients were obtained using negative binomial, logistic, and Cox proportional hazards regression models. We replicated the individual-level analyses in each Ontario Local Health Integration Network (LHIN), then meta-analyzed aggregate trends using both fixed and random effects meta-analysis. We compared the pooled estimates of effect with individual-level analysis.

**Results:** Beta coefficients, summary effect estimates, and 95% confidence intervals (CIs) from the meta-analysis of data from distributed networks were not different than those from individual-level data, regardless of meta-analytic approach used. For example, the 5-year odds ratio of colectomy in ulcerative colitis using individual-level analysis was 0.978 (95% CI 0.950 to 1.007) compared to distributed network fixed effects meta-analysis: 0.982 (95% CI 0.950 to 1.015), and random effects meta-analysis: 0.982 (95% CI 0.950 to 1.015).

Funding: This study was supported by ICES, which is funded by an annual grant from the Ontario Ministry of Health (MOH) and the Ministry of Long-Term Care (MLTC). This study also received funding from a Foundation Grant from the Canadian Institutes of Health Research [grant number 201409FDN-333131-FDN-CECC-164898]. Parts of this material are based on data and/or information compiled and provided by Canadian Institute for Health Information (CIHI). However, the analyses, conclusions, opinions and statements expressed in the material are those of the author(s), and not necessarily those of CIHI.

Conflicts of interest: AKD, MEK, DRK, SKM, GS, and JD have no known conflicts of interest to declare. GGK has received honoraria for speaking or consultancy from Abbvie, Janssen, Pfizer, and Takeda. He has received research support from Janssen, Abbvie, GlaxoSmith Kline, Merck, and Shire. He shares ownership of a patent: TREATMENT OF INFLAMMATORY DISORDERS, AUTOIMMUNE DISEASE, AND PBC. UTI Limited Partnership, assignee. Patent 62/ 555,397. EIB has acted as a legal consultant for Hoffman La-Roche Limited and Peabody & Arnold LLP for matters unrelated to a medication used to treat inflammatory bowel disease.

Data sharing statement: The data from this study is held securely in coded form at ICES. While data sharing agreements prohibit ICES from making the data set publicly available, access may be granted to those who meet prespecified criteria for confidential access, available at www.ices.on.ca/DAS. The full data set creation plan and underlying analytic code are available from the authors upon request, understanding that the programs may rely upon coding templates or macros that are unique to ICES.

* Corresponding author. The Hospital for Sick Children, Division of Gastroenterology, Hepatology and Nutrition, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada. Tel.: +1 416 813 1500; fax: +1 416 813 4972.

*E-mail address:* eric.benchimol@sickkids.ca (E.I. Benchimol).

https://doi.org/10.1016/j.jclinepi.2022.05.006

**Conclusion:** Meta-analysis of multi-jurisdictional estimates were similar to estimates obtained from individual-level analysis. This method is a valid alternative for analysis of multi-jurisdictional data when individual-level data cannot be shared. 

*Keywords:* Routinely collected health data; Inflammatory bowel disease; Pediatrics; Health services research; Health administrative data; Methodology; Multivariable regression analysis

## 1. Introduction

In the ''big data'' era, routinely collected health data is increasingly available for epidemiology and health services research but comes with important privacy concerns, with limits on the ability of researchers to share individual-level data across political borders [1]. For example, in Canada, health administrative data are collected at the provincial level on all residents, however privacy laws prohibit individual-level data from leaving the province. One method for multi-jurisdictional research is the distributed network, which employs identical study methodology on individual-level data to obtain jurisdiction-specific effect estimates, followed by meta-analysis to obtain a pooled estimate [2]. This avoids the need to share individual-level data, often prohibited by governments to preserve privacy. Studies have previously compared the use of this method to the results from individual-level analysis, but under limited conditions [3,4].

The distributed network approach is similar to a two-stage individual participant data (IPD) analysis [5–7]. However, an important distinction between these approaches is data collection. In an IPD meta-analysis, data from previous studies are used which may be heterogenous due to different approaches to study design, data collection, and data analysis. With the meta-analysis in distributed networks, data and methods are standardized across jurisdictions. As a result, IPD meta-analyses have higher expected heterogeneity than in distributed network analyses.

Ontario health administrative data provide an opportunity to explore the validity of meta-analysis in distributed networks because of the province's large population and pre-defined Local Health Integration Networks (LHINs). As LHINs are regional healthcare administrative subunits, they can be used as proxies for multiple jurisdictions.

We aimed to validate meta-analysis in distributed networks using real-world data under a variety of conditions (statistical models and types of data, event rates, sample sizes, number of regions, and heterogeneity) to provide a privacy-preserving tool to analyze aggregate data.

## 2. Methods

This study was approved by the Children's Hospital of Eastern Ontario's Research Ethics Board.

### 2.1. Study design

We previously conducted a retrospective cohort study assessing temporal trends of health services use in pediatric inflammatory bowel disease (IBD) in Ontario [8]. This validation study compared analyses conducted using individual-level data for the entire province (hereafter referred to as the ''individual-level analysis'') and distributed network analyses. To simulate a distributed network within Ontario, we replicated individual-level analyses in each LHIN (administrative units with the ability to plan and regulate some local health care practices within their borders [9], making them suitable proxies for Canadian provinces). We pooled beta coefficients from each LHIN using meta-analysis to obtain the provincial estimate (hereafter referred to as the ''pooled analyses'').

### 2.2. Data sources

Ontario's Ministries of Health and Long-Term Care collect data for the administration of the health system, which are available to researchers at ICES (Toronto, Canada). Data include all health care encounters for Ontario residents eligible for the Ontario Health Insurance Plan (>99% of the population). Residents are linked deterministically across databases using an encrypted unique identifier, allowing for longitudinal analysis of health services use. The databases used in this study are reported in detail elsewhere [10] and in Supplementary Material A.

### 2.3. Study setting, population, and outcomes

Ontario children (aged 6 months to <18 years) newly diagnosed with IBD between fiscal years 1994 to 2012 were identified from the Ontario Crohn's and Colitis Cohort (OCCC) using a validated algorithm (Supplementary Material A) [11]. Patients were classified as having Crohn's disease (CD), ulcerative colitis (UC), or IBD type unclassifiable (IBD-U) based on five of their last seven outpatient diagnostic codes. Children missing date of birth, sex, or postal code of residence were excluded (Supplementary Material B). Residents of LHIN 10 (Kingston region) were also excluded due to known incomplete submission of shadow billings. Children were followed from date of diagnosis until either end of follow-up (31 March 2016), migration out of Ontario, end of OHIP eligibility, or death.

**What is new?**

**Key findings**

- The distributed network approach uses meta-analytic techniques to combine effect estimates from multi-jurisdictional studies when individual-level data cannot be shared.
- In various multivariable regression models, meta-analysis of multi-jurisdictional data from a distributed network approximated individual level analysis. Both fixed and random effects meta-analysis resulted in estimates that were similar to individual-level analyses, though random effects meta-analysis tended to have wider confidence intervals in the presence of high heterogeneity.
- Findings were robust under varying conditions (heterogeneity, number of regions, sample size, regression models, event rate).

**What this adds to what is known?**

- Meta-analysis of data from a distributed network of multi-jurisdictional studies resulted in similar estimates to individual-level analyses in a variety of different condition, with different types of regression analyses.

**What is the implication?**

- This method is a valid alternative for analysis of multi-jurisdictional data when individual-level data cannot be shared.

**What should change now?**

- In circumstances where individual-level data cannot be shared, investigators may use this method to produce valid estimates in multivariable regression analysis.

Table 1 summarizes the outcomes evaluated. We assessed IBD-specific and IBD-related (Supplementary Material C) hospitalizations, emergency department (ED) visits, and outpatient visits in IBD patients, overall and stratified by IBD type (CD or UC). Surgeries were identified using validated codes (Supplementary Material D) [13,14]. Only one office-based outpatient visit or ED visit was counted per day. Only hospital admissions longer than 48 hours were included. Hospitalizations were included if the IBD-specific or IBD-related diagnostic code was most responsible for the hospitalization, a pre- or post-admit comorbidity, or responsible for the transfer between hospitals. Transfers were counted as a single hospitalization.

### 2.4. Statistical analysis

#### 2.4.1. Participant characteristics

Descriptive characteristics are described as means (SD), medians (IQR), or counts (proportions) where appropriate.

#### 2.4.2. Multivariable regression models (individual-level analysis)

Regression models used to assess temporal trends in health services and surgical outcomes in IBD children are summarized in Table 1. The fiscal year of IBD diagnosis (1994 to 2012) was the exposure and was modelled as a linear continuous variable. A knot in the trend of outpatient visits was identified in 2005 using the Joinpoint regression software (Version 4.6.0.0, April 2018; National Cancer Institute); separate linear trends were analyzed before and after 2005. Outcomes were assessed at 1, 3, and 5 years after diagnosis.

Negative binomial regression models accounted for follow-up by setting the offset equal to the natural log of follow-up. Patients with incomplete follow-up were excluded from logistic regression models and collinearity among confounding variables was assessed (maximum variance inflation factor of 2.5) [15,16]. The Cox proportional hazards assumption was assessed using a combination of the time-varying covariate method and visual inspection of plots (log of the negative log of the survival probability function against the log of follow-up). Ties were accounted for using the exact method. Patients were censored at end of follow-up or when lost to follow-up. All models were initially conducted using individual-level data for the entire province and adjusted for the confounding effects of age, sex, rural/urban residence, and mean neighbourhood income quintile at diagnosis.

#### 2.4.3. Meta-analysis in a distributed network (pooled analysis)

Identical regression models from the individual-level analysis were applied to each LHIN (Fig. 1) to quantify LHIN-specific changes in pediatric IBD health service use and surgical outcomes over time. Due to the small number of children with IBD in some LHINs, we combined LHIN 9 with 12 and LHIN 13 with 14, assuming similarities in population characteristics given their close geographic proximity. In addition, statistical models in LHINs six and seven were not adjusted for rural/urban status because there were few (LHIN 6) or no (LHIN 7) rural patients in these LHINs. A total of 11 combined LHINs were included. Where maximum-likelihood logistic regression models did not converge, we used exact logistic regression models. However, both the exact and maximum-likelihood logistic regression model estimates from the pooled data meta-analysis were compared to the maximum-likelihood logistic regression model estimates from the individual-level analysis.

**Table 1.** The multivariable regression models employed to quantify temporal trends in health services and surgical outcomes among children with inflammatory bowel disease in Ontario

| Type of data | Regression model[a] | Follow up from diagnosis date (yr) | Effect estimate | Population | Exposure | Outcome |
|---|---|---|---|---|---|---|
| Count (Number of events per person per year) | Negative binomial | 1, 3, 5 | Incidence rate ratio | IBD, CD, UC | Year of diagnosis (linear) | Hospitalizations |
| | | | | | | ED visits |
| | | | | | Year of diagnosis (knot at 2005) | Outpatient visits |
| Binary (Event did or did not occur) | Logistic[b] | 1, 3, 5 | Odds ratio | CD | Year of diagnosis (linear) | Intestinal resection |
| | | | | UC | Year of diagnosis (linear) | Colectomy |
| Time to first event | Cox proportional hazards | 5 | Hazard ratio | IBD, CD, UC | Year of diagnosis (linear) | Hospitalizations |
| | | | | | | ED visits |
| | | | | CD | Year of diagnosis (linear) | Intestinal resection |
| | | | | UC | Year of diagnosis (linear) | Colectomy |

*Abbreviations*: CD, Crohn's disease; ED, emergency department; IBD, inflammatory bowel disease; UC, ulcerative colitis.

[a] All models were adjusted for the confounding effects of sex, age, mean neighbourhood income quintile, and rural/urban residence.

[b] Exact logistic regression using the network Monte Carlo method [12] was used when conventional logistic regression models did not converge due to sparse data within a Local Health Integration Network (LHIN). Otherwise, maximum-likelihood logistic regression models were used.

#### 2.4.4. Meta-analysis

Beta estimates from LHIN-specific regression models were pooled using effects meta-analysis using generic inverse variance weighting. Both random and fixed effects meta-analyses were conducted to allow comparison of pooled estimates using these two approaches. When LHIN-specific regression models did not converge in three or fewer LHINs, these LHINs were excluded from the meta-analysis. Where $>3$ LHINs did not converge, the outcome was not analyzed because it was assumed the exclusion of $\geq 4$ LHINs may compromise the ability of the analysis to appropriately answer the research question. Heterogeneity was quantified using the $I^2$ statistic and tested using Cochran's Q test ($\alpha = 10\%$). Heterogeneity was calculated using $\tau^2$ for random effects meta-analyses using Restricted Estimates Maximum Likelihood [17,18] and Cochran's Q for fixed effects meta-analyses [19].

#### 2.4.5. Comparison of results from pooled analyses and individual-level analyses

Provincial beta estimates from the distributed network meta-analyses (i.e., the pooled analysis) were compared to the individual-level analysis using two methods: (1) $z$-statistic test of the null hypothesis that the beta estimates were the same at $\alpha = 5\%$ (Equation 1) [20] and (2) visual comparison of the provincial effect summaries (odds ratios [OR], hazard ratios [HR], and incidence rate ratios [IRR]) and 95% confidence intervals (CIs).

$$z = \frac{\widehat{\beta_1} - \widehat{\beta_2}}{\sqrt{SE\left(\widehat{\beta_1}\right)^2 + SE\left(\widehat{\beta_2}\right)^2}} \quad \text{Eq.1}$$

#### 2.4.6. Sensitivity analyses

We conducted four sensitivity analyses. Children living in LHINs with pediatric care centers likely receive more specialized care and may be seen more frequently than children not seen at these centers—possibly introducing heterogeneity between these groups. To assess the effect of heterogeneity, we first combined the individual-level data for children living in LHINs with pediatric care centers (LHINs 2, 4, 7, and 11) and children living in LHINs without pediatric care centers (LHINs 1, 3, 5, 6, 8, 9 and 12, and 13 and 14), forming two separate groups of children. We compared the results individual-level analysis and pooled the estimates from each of the two groups using fixed and random effects meta-analysis.

Second, to explore the impact of varying numbers of regions on the meta-analysis, we pooled LHIN-specific estimates for LHINs with pediatric centers and LHINs without pediatric centers separately. These pooled estimates were compared to estimates obtained from the individual-level analyses, stratified by the presence/absence of a pediatric care center.

Third, we wanted to confirm that the exclusion of LHINs in which models did not converge in the pooled data meta-analysis would not compromise the comparison


**Fig. 1.** Map of the Ontario Local Health Integration Network boundaries. (Accessed from: http://www.lhins.on.ca/).
N.B. Pediatric care centres in LHINs 2, 4, 7, and 11.

of the effect estimates between the pooled data meta-analysis and the individual-level analysis. When the three (or fewer) LHINs were excluded from the pooled data meta-analysis due to nonconvergence, we also excluded children from these LHINs from the individual-level analysis.

Fourth, we re-ran a subset of individual-level models evaluating trends in 5-year outcomes (rates of IBD-related hospitalization, time to IBD-related hospitalization, time to intestinal rection and odds of intestinal resection [CD only]) as multilevel models with random intercepts. Variance components were the specified covariance structure for negative binomial and logistic regression models. Results were compared to the individual-level analyses to ensure our individual-level analyses were robust to the clustering effect introduced by the LHINs.

All regression analyses were performed in the SAS Enterprise Guide version 7.1 (SAS Institute, Cary, NC, USA). Meta-analyses were conducted in R Version 3.5.3 using the Metafor package [19,21].

## 3. Results

### 3.1. Participant characteristics

The overall cohort consisted of 5,518 children with new-onset IBD. The LHIN-stratified descriptive characteristics of IBD patients included in the study are presented in Table 2.

### 3.2. Comparison of pooled data meta-analyses and individual-level analyses

In the pooled analysis, the beta estimates for the models describing IRRs, HR, and ORs from both fixed and random effects meta-analyses were comparable both statistically ($z$-statistic $P > 0.05$) and visually to the beta estimates of the individual-level analyses (Fig. 2, Table 3, and Supplementary Material E and F). The estimates from the fixed and random effects meta-analyses were also comparable to the individual-level analysis estimates when there were low event rates and small sample sizes. This was exemplified in the logistic regression analysis for colectomy within 5 years of UC diagnosis (fixed-effects: pooled OR 0.982, 95% CI 0.950 to 1.015; random-effects: pooled OR 0.982, 95% CI 0.950 to 1.015; individual-level analysis: 0.978, 95% CI 0.950 to 1.007) (Fig. 2). The estimates from exact logistic regression models in the pooled data meta-analysis were comparable to the estimates from the individual-level analysis modelled using maximum likelihood logistic regression ($P > 0.05$ in all cases) (Supplementary Material E and F).

### 3.3. Sensitivity analyses

#### 3.3.1. LHINs with and without pediatric inflammatory bowel disease centers

The observed heterogeneity between LHINs with and without pediatric IBD centers varied across outcomes. When comparing LHINs with pediatric care centers to LHINs without pediatric care centers (exemplified in Fig. 3A and Supplementary Material G and H),

**Table 2.** Descriptive characteristics of patients with inflammatory bowel disease, stratified by Local Health Integration Network

| | | LHIN | | | |
|---|---|---|---|---|---|
| **Characteristic** | | **1 ($n$ = 330)** | **2 ($n$ = 393)** | **3 ($n$ = 326)** | **4 ($n$ = 602)** |
| Sex, *n* (%) | Females | 147 (44.6%) | 175 (44.5%) | 148 (45.4%) | 263 (43.7%) |
| | Males | 183 (55.5%) | 218 (55.5%) | 178 (54.6%) | 339 (56.3%) |
| Age at Diagnosis (yr) | Mean (SD) | 13.7 (3.3) | 13.2 (3.6) | 13.3 (3.5) | 13.2 (3.5) |
| | Median (IQR) | 14.0 (4.0) | 14.0 (5.0) | 14.0 (4.0) | 14.0 (5.0) |
| Diagnosis, *n* (%) | CD | 167 (50.6%) | 226 (57.5%) | 216 (66.3%) | 380 (63.1%) |
| | UC | 141 (42.7%) | 147 (37.4%) | 88 (27.0%) | 185 (30.7%) |
| | IBD-U | 22 (6.7%) | 20 (5.1%) | 22 (6.8%) | 37 (6.2%) |
| Rural residence at diagnosis, *n* (%) | Urban | 298 (90.3%) | 261 (66.4%) | 284 (87.1%) | 565 (93.9%) |
| | Rural | 32 (9.7%) | 120 (33.1%) | 42 (12.9%) | 37 (6.2%) |
| | Unknown | 0 (0.0%) | 2 (0.5%) | 0 (0.0%) | 0 (0.0%) |
| Neighbourhood income quintile at diagnosis, *n* (%) | First (lowest) | 46 (13.9%) | 49 (12.5%) | 29 (8.9%) | 74 (12.3%) |
| | Second | 64 (19.4%) | 73 (18.6%) | 48 (14.7%) | 112 (18.6%) |
| | Third | 59 (17.9%) | 95 (24.2%) | 54 (16.6%) | 112 (18.6%) |
| | Fourth | 71 (21.6%) | 89 (22.7%) | 81 (24.9%) | 155 (25.8%) |
| | Fifth (Highest) | 89 (27.0%) | 85 (21.6%) | 114 (25.0%) | 148 (24.6%) |
| | Unknown | 1 (0.3%) | 2 (0.5%) | 0 (0.0%) | 1 (0.2%) |
| Length of follow up (yr) | Mean (SD) | 10.5 (5.2) | 10.5 (5.3) | 11.0 (5.1) | 11.8 (5.4) |
| | Median (IQR) | 9.8 (8.8) | 9.7 (8.3) | 10.1 (8.3) | 11.0 (9.3) |
| | Min | 1.8 | 1.9 | 1.4 | 1.2 |
| | Max | 21.9 | 22.0 | 21.8 | 22.0 |
| Children missing surgical outcome within 5 yr of diagnosis *n* (%) | CD | 26 (15.6%) | 41 (18.1%) | 30 (13.9%) | 31 (8.2%) |
| | UC | 24 (17.0%) | 28 (19.0%) | 7 (8.0%) | 26 (14.1%) |
| Outcome in IBD children within 5 yr of follow-up[b] | IBD-specific hospitalizations, mean (SD) | 0.250 (0.355) | 0.211 (0.34) | 0.273 (0.404) | 0.306 (0.473) |
| | IBD-related hospitalizations, mean (SD) | 0.263 (0.378) | 0.222 (0.364) | 0.288 (0.424) | 0.317 (0.482) |
| | IBD-specific ED visits, mean (SD) | 0.115 (0.262) | 0.101 (0.217) | 0.116 (0.229) | 0.109 (0.235) |
| | IBD-related ED visits, mean (SD) | 0.286 (0.567) | 0.309 (0.478) | 0.262 (0.424) | 0.272 (0.509) |
| | IBD-specific OP visits, mean (SD) | 4.42 (3.43) | 4.20 (2.96) | 4.32 (2.86) | 4.53 (3.37) |
| | IBD-related OP visits, mean (SD) | 5.46 (3.66) | 4.96 (3.13) | 5.24 (3.60) | 5.33 (4.03) |
| | At least one IBD-specific hospitalization, *n* (%) | 172 (52.1%) | 171 (43.5%) | 166 (50.9%) | 312 (51.8%) |
| | At least one IBD-related hospitalization, *n* (%) | 176 (53.3%) | 174 (44.3%) | 173 (53.1%) | 326 (54.2%) |
| | At least one IBD-specific ED visit, *n* (%) | 101 (30.6%) | 107 (27.2%) | 102 (31.3%) | 180 (29.9%) |
| | At least one IBD-related ED visit, *n* (%) | 173 (52.4%) | 208 (52.9%) | 163 (50.0%) | 304 (50.5%) |
| Outcomes in CD children within 5 yr of follow-up | Intestinal resection, *n* (%) | 45 (31.9%) | 51 (27.6%) | 46 (24.7%) | 85 (24.4%) |
| Outcomes in UC children within 5 yr of follow-up | Colectomy, *n* (%) | 16 (13.7%) | 11 (9.2%) | 9 (11.1%) | 24 (15.1%) |

*Abbreviations*: CD, Crohn's Disease; ED, emergency department; IBD, inflammatory bowel disease; IBD-U, IBD type unclassifiable; IQR, interquartile range; LHIN, Local Health Integration Network; OP, outpatient; SD, standard deviation; UC, ulcerative colitis.

[a] LHINs pooled due to small cells.

[b] Mean outcomes reported as average counts per person per year.

| LHIN | | | | | | |
|---|---|---|---|---|---|---|
| 5 ($n$ = 357) | 6 ($n$ = 514) | 7 ($n$ = 343) | 8 ($n$ = 839) | 11 ($n$ = 591) | 9 & 12[a] ($n$ = 823) | 13 & 14[a] ($n$ = 400) |
| 153 (42.9%) | 220 (42.8%) | 134 (39.1%) | 347 (41.4%) | 269 (45.5%) | 365 (44.4%) | 175 (43.8%) |
| 204 (57.1%) | 294 (57.2%) | 209 (60.9%) | 492 (58.6%) | 322 (54.5%) | 458 (55.7%) | 225 (56.3%) |
| 12.5 (3.9) | 13.1 (3.6) | 12.7 (3.9) | 12.9 (3.5) | 13.0 (3.6) | 13.1 (3.6) | 13.6 (3.3) |
| 13.0 (5.0) | 14.0 (5.0) | 14.0 (6.0) | 14.0 (5.0) | 14.0 (5.0) | 14.0 (5.0) | 15.0 (4.0) |
| 189 (52.9%) | 283 (55.1%) | 191 (55.7%) | 469 (55.9%) | 384 (65.0%) | 481 (58.4%) | 226 (56.5%) |
| 141 (39.5%) | 198 (38.5%) | 124 (36.2%) | 311 (37.1%) | 183 (31.0%) | 289 (35.1%) | 138 (34.5%) |
| 27 (7.6%) | 33 (6.4%) | 28 (8.2%) | 59 (7.0%) | 24 (4.1%) | 53 (6.4%) | 36 (9.0%) |
| 347 (97.2%) | 508-514 (98.8-99.9%) | 343 (100.0%) | 828 (98.7%) | 515 (87.1%) | 701 (85.2%) | 274 (68.5%) |
| 10 (2.8%) | <6 | 0 (0.0%) | 9 (1.1%) | 75 (12.7%) | 121 (14.7%) | 126 (31.5%) |
| 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 2 (0.2%) | 1 (0.2%) | 1 (0.1%) | 0 (0.0%) |
| 36 (10.1%) | 43 (8.4%) | 86 (25.1%) | 91 (10.9%) | 58 (9.8%) | 121 (14.7%) | 81 (20.3%) |
| 88 (24.7%) | 53 (10.3%) | 48 (14.0%) | 111 (13.2%) | 97 (16.4%) | 161 (19.6%) | 73 (18.3%) |
| 134 (37.5%) | 82 (16.0%) | 42 (12.2%) | 152 (18.1%) | 120 (20.3%) | 167 (20.3%) | 75 (18.8%) |
| 62 (17.4%) | 155 (30.2%) | 26 (7.6%) | 223 (26.6%) | 143 (24.2%) | 199 (24.2%) | 81 (20.3%) |
| 37 (10.4%) | 181 (35.2%) | 141 (41.1%) | 260 (31.0%) | 172 (29.1%) | 174 (21.1%) | 86 (21.5%) |
| 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 2 (0.2%) | 1 (0.2%) | 1 (0.1%) | 4 (1.0%) |
| 9.9 (5.1) | 10.6 (5.3) | 11.0 (5.7) | 11.2 (5.5) | 10.7 (5.1) | 11.2 (5.4) | 11.6 (5.5) |
| 8.7 (8.0) | 9.6 (8.2) | 10.1 (9.7) | 10.1 (9.2) | 9.9 (8.4) | 10.6 (9.1) | 10.5 (9.6) |
| 1.7 | 1.1 | 1.6 | 2.3 | 1.4 | 0.4 | 1.2 |
| 21.8 | 22.0 | 22.0 | 22.0 | 21.9 | 22.0 | 22.0 |
| 33 (17.5%) | 46 (16.3%) | 38 (20.0%) | 65 (13.9%) | 66 (17.2%) | 74 (15.4%) | 20 (8.8%) |
| 31 (22.0%) | 36 (18.2%) | 17 (13.7%) | 47 (15.1%) | 18 (9.8%) | 37 (12.8%) | 21 (15.2%) |
| 0.218 (0.317) | 0.248 (0.353) | 0.243 (0.404) | 0.241 (0.362) | 0.272 (0.364) | 0.251 (0.366) | 0.293 (0.427) |
| 0.235 (0.355) | 0.264 (0.377) | 0.259 (0.417) | 0.249 (0.375) | 0.282 (0.378) | 0.262 (0.375) | 0.318 (0.479) |
| 0.096 (0.198) | 0.103 (0.206) | 0.085 (0.221) | 0.09 (0.183) | 0.115 (0.229) | 0.131 (0.251) | 0.175 (0.324) |
| 0.263 (0.376) | 0.258 (0.443) | 0.267 (0.728) | 0.241 (0.382) | 0.285 (0.495) | 0.353 (0.577) | 0.479 (0.803) |
| 4.05 (3.33) | 4.10 (2.93) | 3.98 (2.57) | 4.13 (3.34) | 5.41 (3.61) | 3.94 (2.92) | 4.26 (3.77) |
| 5.52 (5.70) | 5.34 (3.94) | 5.56 (4.76) | 5.50 (4.33) | 6.28 (4.04) | 5.21 (3.76) | 5.70 (5.04) |
| 165 (46.2%) | 258 (50.2%) | 152 (44.3%) | 413 (49.2%) | 317 (53.6%) | 392 (47.6%) | 211 (52.8%) |
| 172 (48.2%) | 263 (51.2%) | 160 (46.6%) | 420 (50.1%) | 321 (54.3%) | 404 (49.1%) | 213 (53.3%) |
| 105 (29.4%) | 147 (28.6%) | 70 (20.4%) | 231 (27.5%) | 177 (29.9%) | 288 (35.0%) | 167 (41.8%) |
| 199 (55.7%) | 241 (46.9%) | 152 (44.3%) | 406 (48.4%) | 298 (50.4%) | 462 (56.1%) | 242 (60.5%) |
| 24 (15.4%) | 51 (21.5%) | 31 (20.3%) | 76 (18.8%) | 68 (21.4%) | 103 (25.3%) | 37 (18.0%) |
| 12 (10.9%) | 20 (12.3%) | 9 (8.4%) | 29 (11.0%) | 35 (21.2%) | 38 (15.1%) | 15 (12.8%) |

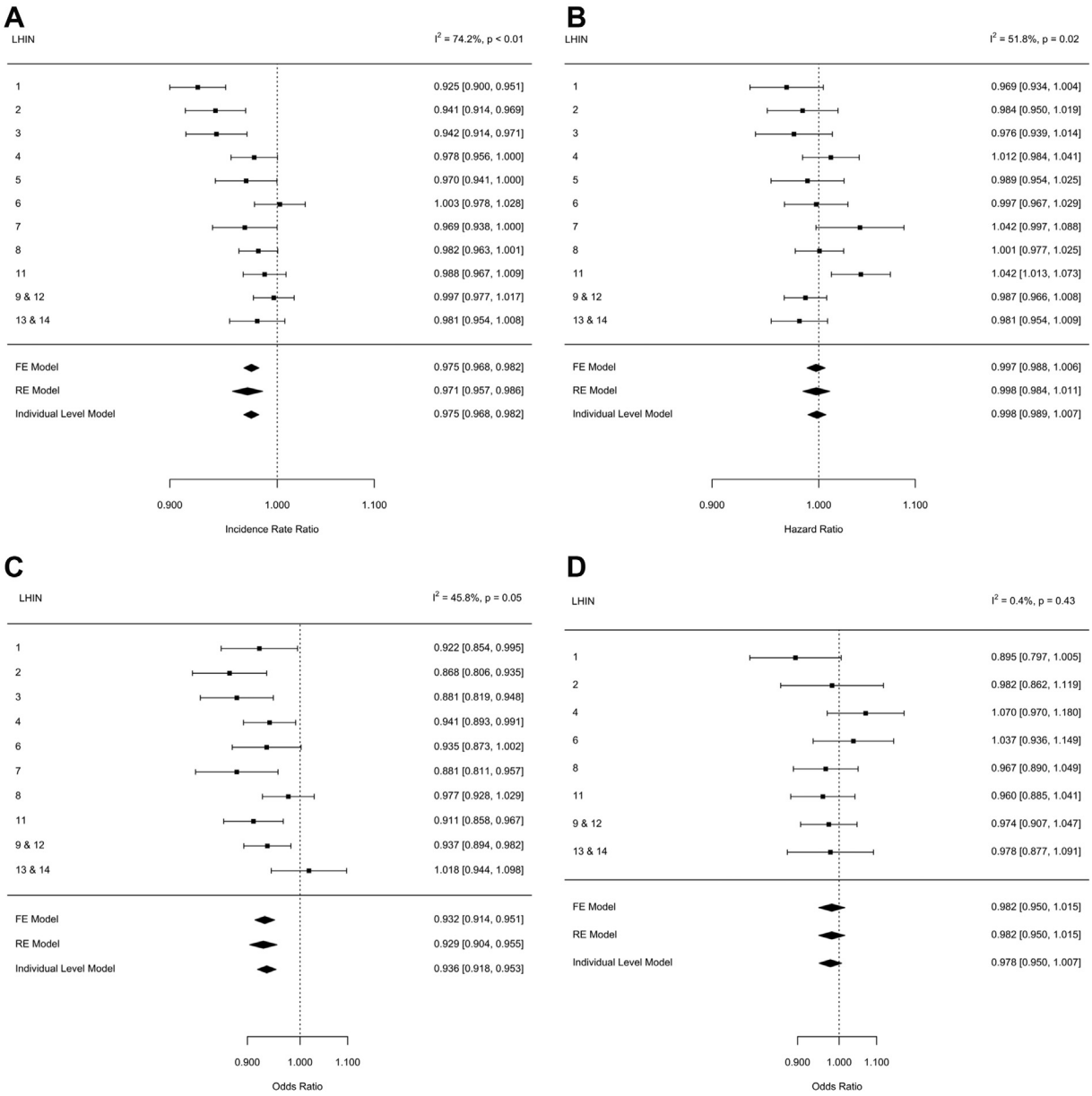


**Fig. 2.** Forest plots for depicting the effect estimates [95% confidence intervals] from fixed and random effects meta-analyses compared with individual-level analyses for (A) incidence rate ratio of IBD-specific hospitalizations, (B) hazard ratio of IBD-specific emergency department visits, (C) odds ratio of intestinal resection (Crohn's disease), and (D) odds ratio of colectomy (ulcerative colitis) within 5 years of diagnosis in children with inflammatory bowel disease. *Abbreviations*: FE, fixed effects; LHIN, Local Health Integration Network; RE, random effects.

considerable heterogeneity ($I^2 > 75\%$) [22] was seen among analyses comparing ED visit and surgical outcomes. Moderate to considerable heterogeneity was seen across hospitalization and outpatient visit outcomes within LHINs with pediatric care centers and within LHINs without the centers. However, in each of the pooled data meta-analyses, the fixed and random effects meta-analyses were similar to the results from individual-level analyses across all levels of heterogeneity observed ($z$-statistic $P > 0.05$). Random effects meta-analyses resulted in slightly wider 95% CIs in the presence of considerable heterogeneity.

### 3.3.2. *Pooling of LHINs with presence or absence of pediatric inflammatory bowel disease care centers separately*

In all scenarios, the fixed and random effects models were similar to the results from individual-level analysis across all numbers of regions included in the meta-analysis ($z$-statistic $P > 0.05$) (Fig. 3B and C, and Supplementary Material I and J).

### 3.3.3. *The impact of model non-convergence on effect estimates*

In the sensitivity analysis exploring whether effect estimates from the individual-level analysis were robust to

the exclusion of LHINs that did not converge, the estimates were similar in the individual-level analysis (Supplementary Material K).

#### *3.3.4. Comparison of individual-level models with and without random intercepts*

The inclusion of random intercepts to account for clustering in LHINs resulted in similar effect estimates and confidence intervals to the main analyses (Supplementary Material L).

## 4. Discussion

We demonstrated that results of multivariable regression of individual-level health administrative data were similar to results from the meta-analysis of regional data using a distributed network approach in a study assessing time trends in health services utilization of children with IBD. This was consistent for fixed and random effects meta-analyses, all outcomes, levels of heterogeneity, regression approaches, event rates, number of regions, and sample sizes. Findings were robust to substituting maximum likelihood-based with exact logistic regression models when models failed to converge due to rare events. Fixed and random effects meta-analyses were comparable to the individual-level analysis in situations with high and low heterogeneity; random effects meta-analyses had wider confidence intervals when heterogeneity was high.

This study validates a privacy-preserving method when conducting multi-jurisdictional multi-database studies using real-world health administrative data. This is increasingly important as multi-jurisdictional research becomes more common in parallel with rising availability of routinely collected health data. Our findings are consistent with those of previous validation studies. Because the creation of the Sentinel network and the Canadian Network for Observational Drug Effect Studies (CNODES), studies have compared and validated alternative methods of combining aggregate data. Weighted regression on stratum-specific means is one example but requires a continuous outcome and cannot accommodate a continuous predictor [23]. A simulation study compared HRs and corresponding CIs from case-centered logistic regression analysis to those from meta-analysis [24]. The meta-analysis was susceptible to bias and lower-powered than case-centered regression in some circumstances, such as low sample size or event rate [24]. Similar to our study, fixed and random effects meta-analyses gave comparable results. However, this study lacked an ideal reference standard [24]. A more recent simulation studies addressed this limitation [25], demonstrating site-specific Cox regression summary table data, risk-set data, and site-specific estimate meta-analysis were all comparable to results obtained from individual-level data. Inconsistencies were again noted for small sample sizes ($<1000$ patients per site) and rare events, particularly in the meta-analysis method [25]. Other studies, which evaluated confounding adjustment methods in distributed networks [26] and conducted analysis of real-world data in distributed networks [3,4,27−29], have also looked at the validity of meta-analysis, in addition to other methods of combining aggregate data. However, the present validation study differed from previous work because it explored a number of conditions (regression approaches, event rates, sample sizes, number of pooled regions, and heterogeneity), compared the use of fixed effects to random effects meta-analyses for all analyses, and tested the beta coefficients of the statistical models from the different methods in addition to visual assessment.

This study used population-based data of all children with IBD living in Ontario, Canada, thereby providing sufficient regional sample sizes. Study limitations stemmed from an inability to introduce more clinical and methodological heterogeneity that would be seen when combining regions with differing healthcare systems, population characteristics, or differences in the collection and use of health administrative data (e.g., differing accuracy of algorithms used to identify exposures or outcomes). In addition, to our knowledge, there is no definitive method available for comparing the results of meta-analysis in the distributed network and individual-level data analysis. Although we used the $z$-statistic test, and the results were in agreement with the visual assessments, it is possible not all assumptions were met [20].

We specifically explored the validity of the distributed network analytic approach as described and used by Canadian networks such as CNODES and the Canadian Gastro-Intestinal Epidemiology Consortium (CanGIEC) [2,30]. The meta-analysis of regional data resulted in similar effect estimates to individual-level analysis. Therefore, this approach is a valid alternative to analyzing individual-level data when privacy regulations prevent sharing of individual-level records. However, future studies should explore the validity of the meta-analysis technique in distributed networks on both a larger scale (e.g., international multijurisdictional studies where there is greater heterogeneity across jurisdictions) and smaller scale (e.g., multi-center data from electronic medical records) to further understand the impact of heterogeneity that this study was unable to capture. For example, although we expected populations to differ across LHINs we expected similarities because all data was from Ontario, emulating a common data model. This degree of homogeneity in data collection may not be

**Table 3.** Results of the 5-year outcomes comparing individual-level analysis and meta analysis of data from models in a distributed network[a]

| Effect estimate type | Outcome | Outcome Type | Population | Sample Size | Number of LHINs |
|---|---|---|---|---|---|
| IRR | Hospitalizations | IBD-related | IBD | 5,506 | - |
| | | | | | 11 |
| | Emergency Department visits | IBD-related | IBD | 5,506 | - |
| | | | | | 11 |
| | Outpatient visits (Before 2005) | IBD-related | IBD | 5,506 | - |
| | | | | | 11 |
| | Outpatient visits (After 2005) | IBD-related | IBD | 5,506 | - |
| | | | | | 11 |
| Hazard Ratio | Hospitalizations | IBD-related | IBD | 5,506 | - |
| | | | | | 11 |
| | Emergency Department Visits | IBD-related | IBD | 5,506 | - |
| | | | | | 11 |
| | Surgery | Intestinal Resection | CD | 3,205 | - |
| | | | | | 11 |
| | Surgery | Colectomy | UC | 1940 | - |
| | | | | | 11 |
| Odds Ratio | Surgery | Intestinal Resection | CD | 2742 | - |
| | | | | | 10 |
| | Surgery | Colectomy | UC | 1653 | - |
| | | | | | 8 |

*Abbreviations*: CD, Crohn's disease; CI, confidence interval; FE, fixed effects meta-analysis in LHIN-based analysis; IBD, inflammatory bowel disease; HR, hazard ratio for 1 year increase in diagnosis/index date; LHIN, Local Health Integration Network; RE, random effects meta-analysis in LHIN-based analysis; UC, ulcerative colitis.

[a] These are a subset of all the results. The full results tables can be found in Supplementary Material E.

generalizable to all contexts in which these methods may be applied. In addition, studies should further assess the use of fixed vs random effects meta-analytic approaches with large effect estimates and continue exploring the use of meta-analysis in distributed networks under differing conditions, such as different analytic methods. Although we compared the difference between fixed and random effects meta-analysis models, we believe the choice in model should consider the clinical and research contexts and the underlying data to satisfy model assumptions.

In conclusion, the distributed network approach, in which identical analytic code is deployed across regions then meta-analyzed, allows for multi-jurisdictional research sharing of individual-level data is prohibited. This approach is robust to events rates, sample size, heterogeneity, type of regression model, and number of regions used in this study and may be a valid approach for multi-jurisdictional research.

| $\tau^2$ | $I^2$, *P*-value | Analysis | Beta Estimate | Standard Error | Effect estimate (95% CI) |
|---|---|---|---|---|---|
| - | - | Individual | −0.0260 | 0.0037 | 0.974 (0.967, 0.981) |
| N/A | 73.4%, $P < 0.01$ | FE | −0.0259 | 0.0038 | 0.974 (0.967, 0.982) |
| 0.0005 | 75.1%, $P < 0.01$ | RE | −0.0296 | 0.0077 | 0.971 (0.956, 0.986) |
| - | - | Individual | 0.0147 | 0.0039 | 1.015 (1.007, 1.023) |
| N/A | 66.0%, $P < 0.01$ | FE | 0.0161 | 0.0039 | 1.016 (1.008, 1.024) |
| 0.0004 | 68.9%, $P < 0.01$ | RE | 0.0149 | 0.0071 | 1.015 (1.001, 1.029) |
| - | - | Individual | −0.0092 | 0.0030 | 0.991 (0.985, 0.997) |
| N/A | 64.7%, $P < 0.01$ | FE | −0.0085 | 0.0030 | 0.992 (0.986, 0.997) |
| 0.0002 | 62.8%, $P < 0.01$ | RE | −0.0104 | 0.0050 | 0.990 (0.980, 0.999) |
| - | - | Individual | 0.0278 | 0.0044 | 1.028 (1.019, 1.037) |
| N/A | 60.2%, $P < 0.01$ | FE | 0.0279 | 0.0043 | 1.028 (1.020, 1.037) |
| 0.0003 | 59.6%, $P < 0.01$ | RE | 0.0297 | 0.0069 | 1.030 (1.016, 1.044) |
| Individual | - | - | −0.0170 | 0.0035 | 0.983 (0.976, 0.990) |
| FE | N/A | 65.9%, $P < 0.01$ | −0.0171 | 0.0035 | 0.983 (0.976, 0.990) |
| RE | 0.0003 | 69.3%, $P < 0.01$ | −0.0199 | 0.0065 | 0.980 (0.968, 0.993) |
| Individual | - | - | 0.0094 | 0.0035 | 1.009 (1.003, 1.016) |
| FE | N/A | 60.6%, $P < 0.01$ | 0.0091 | 0.0035 | 1.009 (1.002, 1.016) |
| RE | 0.0002 | 63.1%, $P < 0.01$ | 0.0076 | 0.0060 | 1.008 (0.996, 1.019) |
| Individual | - | - | −0.0616 | 0.0072 | 0.940 (0.927, 0.954) |
| FE | N/A | 44.2%, $P = 0.06$ | −0.0644 | 0.0074 | 0.938 (0.924, 0.951) |
| RE | 0.0005 | 43.8%, $P = 0.06$ | −0.0668 | 0.0101 | 0.935 (0.917, 0.954) |
| Individual | - | - | −0.0304 | 0.0119 | 0.970 (0.948, 0.993) |
| FE | N/A | 0.0%, $P = 0.70$ | −0.0335 | 0.0123 | 0.967 (0.944, 0.991) |
| RE | <0.0001 | 0.0%, $P = 0.70$ | −0.0335 | 0.0123 | 0.967 (0.944, 0.991) |
| Individual | - | - | −0.0666 | 0.0095 | 0.936 (0.918, 0.953) |
| FE | N/A | 46.3%, $P = 0.05$ | −0.0705 | 0.0101 | 0.932 (0.914, 0.951) |
| RE | 0.0009 | 45.8%, $P = 0.05$ | −0.0735 | 0.0141 | 0.929 (0.904, 0.955) |
| Individual | - | - | −0.0221 | 0.0147 | 0.978 (0.950, 1.007) |
| FE | N/A | 0.0%, $P = 0.43$ | −0.0185 | 0.0169 | 0.982 (0.950, 1.015) |
| RE | <0.0001 | 0.4%, $P = 0.43$ | −0.0185 | 0.0169 | 0.982 (0.950, 1.015) |

## Acknowledgments

M. Ellen Kuenzig was supported by a Post-Doctoral Fellowship Award from the Canadian Institutes of Health Research (CIHR), Canadian Association of Gastroenterology (CAG), Crohn's and Colitis Canada, and Mitacs Canada. David Mack was supported by a University of Ottawa Distinguished Clinical Research Chair. David Mack is supported in part through a University of Ottawa, Faculty of Medicine, Distinguished Clinical Research Chair in Pediatric IBD. Sanjay Murthy was supported by the Department of Medicine and Division of Gastroenterology and by a Tier 2 Clinical Research Chair in Inflammatory Bowel Diseases from the Faculty of Medicine at the University of Ottawa. Eric I. Benchimol was supported by a New Investigator Award from the CIHR, CAG and Crohn's and Colitis Canada. Eric Benchimol was also supported by the Career Enhancement Program of the Canadian Child Health Clinician Scientist Program.

## CRediT authorship contribution statement

**Aman K. Dheri:** Conceptualization, Methodology, Formal analysis, Writing – original draft. **M. Ellen Kuenzig:** Conceptualization, Methodology, Formal analysis, Writing – original draft, Supervision. **David R. Mack:** Conceptualization,

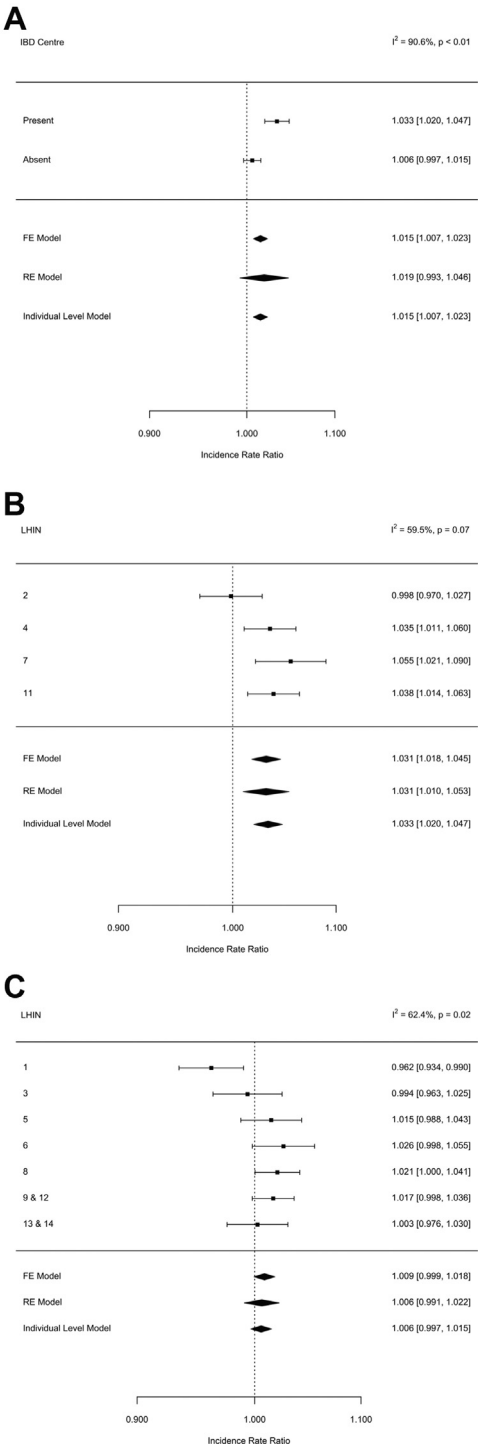


**Fig. 3.** Forest plots for depicting the effect estimates [95% confidence intervals] from fixed and random effects meta-analyses compared with individual-level analyses for incidence rate ratios of IBD-related emergency department visits within 5 years of diagnosis by pooling children with IBD in (A) LHINs with and without pediatric care centres, (B) LHINs with pediatric care centres, (C) LHINs without pediatric care centres. *Abbreviations*: FE, fixed effects; LHIN, Local Health Integration Network; RE, random effects.

Methodology, Formal analysis, Writing – review & editing, Supervision. **Sanjay K. Murthy:** Conceptualization, Methodology, Formal analysis, Writing – review & editing, Supervision. **Gilaad G. Kaplan:** Conceptualization, Methodology, Writing – Review & editing. **Jessy Donelle:** Methodology, Data curation. **Glenys Smith:** Methodology, Data curation. **Eric I. Benchimol:** Conceptualization, Methodology, Formal analysis, Writing – original draft, Supervision, Funding acquisition, Resources.

## Appendix B

## Supplementary Data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2022.05.006.

## References

[1] Nicholls SG, Langan SM, Benchimol EI. Reporting and transparency in big data: the nexus of ethics and methodology. In: Mittelstadt B, Floridi L, editors. The Ethics of Biomedical Big Data. Law, Governance and Technology Series, 29. Cham: Springer; 2016:339–65.

[2] Suissa S, Henry D, Caetano P, Dormuth CR, Ernst P, Hemmelgarn B, et al. CNODES: the Canadian network for observational drug effect studies. Open Med 2012;6:134–40.

[3] Selmer R, Haglund B, Furu K, Andersen M, Nørgaard M, Zoëga H, et al. Individual-based versus aggregate meta-analysis in multidatabase studies of pregnancy outcomes: the Nordic example of selective serotonin reuptake inhibitors and venlafaxine in pregnancy. Pharmacoepidemiol Drug Saf 2016;25:1160–9.

[4] Yang D, Zolfaghari S, Postuma R, Saha-Chaudhuri P. A comparison of privacy-preserving statistical methods for analyzing distributed data with binary outcome variables. Canadian Society for Epidemiology and Biostatistics, Ottawa, Ontario: n.d., p. Abstract 174.

[5] Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. Stat Med 2017;36:855–75.

[6] Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. BMJ 2010;340:521–5.

[7] Yamaguchi Y, Sakamoto W, Goto M, Staessen JA, Wang J, Gueyffier F, et al. Meta-analysis of a continuous outcome combining individual patient data and aggregate data: a method based on simulated individual patient data. Res Synth Methods 2014;5:322–51.

[8] Dheri AK, Kuenzig ME, Mack DR, Murthy SK, Kaplan GG, Donelle J, et al. Shifting health care use from hospitalisations and surgeries to outpatient visits in children with inflammatory bowel disease: a population-based cohort study from Ontario, Canada. J Crohns Colitis 2021;15:1991–2000.

[9] Ministry of Health and Long Term Care. Local health integration network history. History 2009:2. Available at http://www.lhins.on.ca/. Accessed February 2, 2020.

[10] Dheri A. Distributed network meta-analysis estimates results from individual-level analysis using Ontario health administrative data on pediatric inflammatory bowel disease health services use: a population-based cohort study Aman. Ottawa, Ontario: University of Ottawa; 2020.

[11] Benchimol EI, Guttmann A, Griffiths AM, Rabeneck L, Mack DR, Brill H, et al. Increasing incidence of paediatric inflammatory bowel disease in Ontario, Canada: evidence from health administrative data. Gut 2009;58:1490−7.
[12] Mehta CR, Patel NR, Senchaudhuri P. Efficient Monte Carlo methods for conditional logistic regression. J Am Stat Assoc 2000;95:99−108.
[13] Ma C, Crespin M, Proulx M-C, DeSilva S, Hubbard J, Prusinkiewicz M, et al. Postoperative complications following colectomy for ulcerative colitis: a validation study. BMC Gastroenterol 2012;12:39.
[14] Ma C, Moran GW, Benchimol EI, Targownik LE, Heitman SJ, Hubbard JN, et al. Surgical rates for Crohn's disease are decreasing: a population-based time trend analysis and validation study. Am J Gastroenterol 2017;112:1840−8.
[15] Allison PD. Logistic regression using the SAS® system. Technometrics 2000;42:323−4.
[16] Allison PD. Logistic regression using SAS: theory and application, second edition ITPro collection. Cary, NC: SAS Institute; 2012.
[17] Veroniki AA, Jackson D, Viechtbauer W, Bender R, Bowden J, Knapp G, et al. Methods to estimate the between-study variance and its uncertainty in meta-analysis. Res Synth Methods 2016;7:55−79.
[18] Viechtbauer W. Bias and efficiency of meta-analytic variance estimators in the random-effects model. J Educ Behav Stat 2005;30:261−93.
[19] Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw 2010;36:1−48.
[20] Clogg CC, Petkova E, Haritou A. Statistical methods for comparing regression coefficients between models. Am J Sociol 1995;100:1261−93.
[21] R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation; 2019.
[22] Higgins J, Green S. Cochrane handbook for systematic reviews of interventions version 5.3.0. Chichester: Cochrane Collab; 2015.
[23] Moineddin R, Urquia ML. Regression analysis of aggregate continuous data. Epidemiology 2014;25:929−30.
[24] Toh S, Reichman ME, Houstoun M, Ding X, Fireman BH, Gravel E, et al. Multivariable confounding adjustment in distributed data networks without sharing of patient-level data. Pharmacoepidemiol Drug Saf 2013;22:1171−7.
[25] Yoshida K, Gruber S, Fireman BH, Toh S. Comparison of privacy-protecting analytic and data-sharing methods: a simulation study. Pharmacoepidemiol Drug Saf 2018;27:1034−41.
[26] Li X, Fireman BH, Curtis JR, Arterburn DE, Fisher DP, Moyneur É, et al. Validity of privacy-protecting analytical methods that use only aggregate-level information to conduct multivariable-adjusted analysis in distributed data networks. Am J Epidemiol 2019;188:709−23.
[27] Toh S, Reichman ME, Houstoun M, Southworth MR, Ding X, Hernandez AF, et al. Comparative risk for angioedema associated with the use of drugs that target the renin-angiotensin-aldosterone system. Arch Intern Med 2012;172:1582−9.
[28] Rassen JA, Avorn J, Schneeweiss S. Multivariate-adjusted pharmacoepidemiologic analyses of confidential information pooled from multiple health care utilization databases. Pharmacoepidemiol Drug Saf 2010;19:848−57.
[29] Toh S, Shetterly S, Powers JD, Arterburn D. Privacy-preserving analytic methods for multisite comparative effectiveness and patient-centered outcomes research. Med Care 2014;52:664−8.
[30] Benchimol EI, Bernstein CN, Bitton A, Carroll MW, Singh H, Otley AR, et al. Trends in epidemiology of pediatric inflammatory bowel disease in Canada: distributed network analysis of multiple population-based provincial health administrative databases. Am J Gastroenterol 2017;112:1120−34.