Recent Trends in the Prevalence of Crohn’s Disease and Ulcerative Colitis in a Commercially Insured US Population

Michael D. Kappelman · Kristen R. Moore · Jeffery K. Allen · Suzanne F. Cook

Abstract

Purpose Most US inflammatory bowel disease (IBD) epidemiology studies conducted to date have sampled small, geographically restricted populations and have not examined time trends. The aim of our study was to determine the prevalence of Crohn’s disease (CD) and ulcerative colitis (UC) in a commercially insured US population and compare prevalences across sociodemographic characteristics and time.

Methods Using claims data from approximately 12 million Americans, we performed three consecutive 2-year cross-sectional studies. Cases of CD and UC were identified using a previously described algorithm. Prevalence was estimated by dividing cases by individuals in the source population. Logistic regression was used to compare prevalences by region, age, and sex.

Results In 2009, the prevalences of CD and UC in children were 58 [95 % confidence interval (CI) 55–60] and 34 (95 % CI 32–36), respectively. In adults, the respective prevalences were 241 (95 % CI 238–245) and 263 (95 % CI 260–266). Data analysis revealed that IBD prevalences have slightly increased over time. Based on census data, an estimated 1,171,000 Americans have IBD (565,000 CD and 593,000 UC).

Conclusions Analysis of the epidemiological data revealed an increasing burden of IBD in recent years, which may be used to inform policy.

Keywords Crohn’s disease · Ulcerative colitis · Prevalence · Epidemiology

Introduction

Crohn’s disease (CD) and ulcerative colitis (UC), collectively referred to as the inflammatory bowel diseases (IBDs), are characterized by chronic inflammation of the gastrointestinal tract. The cause(s) of these diseases is unknown; however, genetics/family history, and environmental factors appear to influence disease risk. Epidemiological studies of IBD in the USA are necessary to quantify the public health burden of disease and inform policy regarding the allocation of resources and provision of health services for affected individuals. Because IBD is not a reportable condition in the USA and comprehensive, nationwide registries for IBD surveillance have not been established, published studies on the epidemiology of IBD in the USA [1–6] are limited and primarily include studies which have sampled small, geographically restricted populations. Furthermore, no studies of IBD prevalence have been published using data from the last 5 years and, therefore, current time trends remain unknown.

The aims of this study were (1) to estimate the current prevalence of IBD in the USA as of 2009 using a large, regionally diverse population of commercially insured individuals, and (2) to examine recent trends in disease prevalence. A further aim was to characterize differences
in disease epidemiology by geographic region, age, and sex.

Materials and Methods

Study Design and Data Source

Using the inpatient, outpatient, and pharmaceutical insurance claims data sets contained within the PharMetrics Choice Patient-Centric Database (IMS Health, Watertown, MA), we performed three consecutive cross-sectional studies for the following 2-year time periods: 1 January 2004–31 December 2005, 1 January 2006–31 December 2007, and 1 January 2008–31 December 2009. This database pools together claims data from 20 to 25 commercial health plans covering all four US census regions. The majority of patients in the database have both medical and pharmacy benefits. Mean health plan enrollment is 6.4 [interquartile range (IQR) 5.3–7.2] years for the 2004–2005 cohort, 5.8 (IQR 4.3–7.1) years for the 2006–2007 cohort, and 4.6 (IQR 3.2–6.4) years for the 2008–2009 cohort. Prior studies have reported PharMetrics to be representative of the national commercially insured population on a variety of demographic measures [7].

Source Populations

All patients in the database with continuous health plan enrollment for at least one of the above 2-year time periods comprised the source population for this study. Individuals turning 65 years of age or older (year of period end − year of birth) during each 2-year period were removed from the source population to avoid the possibility of missing data due to Medicare dual-enrollment. The following demographic elements were available for all members of the source population: year of birth, gender, and US census region (Northeast, South, Midwest, and West).

Case Identification

The case definition used to identify prevalent cases of IBD was adapted from a previously validated University of Manitoba IBD Epidemiology Database definition [8] and has been used in prior epidemiological studies of IBD [3, 9–12]. For each 2-year period, cases were identified as individuals with at least 3 healthcare contacts, on different days, associated with an International Classification of Diseases 9th revision diagnosis code for CD (555.x) or UC (556.x). Additionally, individuals with at least one claim for CD or UC who also had at least one pharmacy claim for an IBD-specific medication (mesalamine, olsalazine, balsalazide, sulfasalazine, 6-mercaptopurine, azathioprine, infliximab, adalimumab, certolizumab, natalizumab, methotrexate, and enteral-release budesonide) were also included in our case definition. For those patients who had submitted claims for both CD and UC, disease assignment was made according to the majority of the last nine claims. If a distinction between CD and UC could not be made using these criteria, cases were designated as IBD not otherwise specified (NOS). Each 2-year time period had a unique base population depending on the number of patients who met the above-mentioned eligibility criteria. Some individuals were included in more than one time period if they met the inclusion criteria over multiple years.

Statistical Analysis

Analyses were performed for each of the time periods: 2004–2005, 2006–2007, and 2008–2009. Prevalence estimates were made for CD, UC, and total IBD (CD + UC + IBD NOS) as of 31 December 2005, 31 December 2007, and 31 December 2009, respectively, by dividing the number of identified cases by the corresponding number of individuals in the source population. All estimates were reported as cases per 100,000 persons, and also stratified by age, gender, and US census region. Exact 95 % confidence limits (95 % CI) were calculated using a variation on the Wilson Score confidence interval method.

To determine the independent effects of age, gender, and region on the prevalence of IBD in the final time period, multivariable logistic regression was used to compare differences in prevalence by age group, gender, and geographic region while controlling simultaneously for all other data elements. To estimate the projected number of people with IBD in the US population as of 2009 and to determine adjusted prevalence estimates, we multiplied the age, gender, and region-specific prevalences obtained in our sample by the corresponding US population according to census data.

Next, in order to facilitate our analysis of trends in the prevalence of CD and UC across time periods, we standardized the estimates obtained in each period to reflect the age, gender, and regional distribution of the US population according to 2009 census data and conducted pairwise comparisons (2006–2007 vs. 2004–2005; 2008–2009 vs. 2004–2005; 2008–2009 vs. 2006–2007) using t tests. We recognize that t tests are based on the assumption that the samples in the different time periods are independent and, in fact, some individuals in this study contribute to more than one time period. However, such repeated sampling can be expected to increase the precision of the comparison between time periods, making the tests conducted here conservative.

All statistical analyses were performed using SAS ver. 9.1 or above (SAS Institute, Cary, NC).
Ethical Considerations

As this study was an analysis of existing, de-identified data, it is not considered to be human subject research.

Results

Source Population

The source population for each 2-year period comprised 5,621,520 individuals (29% children <20 years of age) in 2004–2005, 8,139,941 individuals (28% children) in 2006–2007, and 12,538,475 individuals (28% children) in 2008–2009. Additional demographics for the source population are shown in Table 1.

IBD Prevalence in 2009

In this commercially insured population, the overall prevalences of CD and UC in the pediatric population were 58 and 34 per 100,000, respectively. In the adult population, the overall prevalences of CD and UC were 241 and 263 per 100,000, respectively. Additional details on the

### Table 1 Source population and demographics by time period in a commercially insured US population (2004–2009)

| Demographic characteristic | Time period |
|----------------------------|-------------|
|                            | 2004–2005, N (%) | 2006–2007, N (%) | 2008–2009, N (%) |
| **Pediatric**              |             |             |             |
| Age group (years)          |             |             |             |
| <5                         | 218,748 (13) | 305,439 (13) | 461,434 (13) |
| 5–9                        | 403,811 (25) | 567,674 (25) | 886,112 (25) |
| 10–14                      | 477,600 (29) | 660,328 (29) | 1,019,459 (29) |
| 15–19                      | 526,478 (32) | 748,376 (33) | 1,123,040 (32) |
| Sex                        |             |             |             |
| Female                     | 796,394 (49) | 1,116,707 (49) | 1,705,833 (49) |
| Male                       | 830,209 (51) | 1,164,984 (51) | 1,784,000 (51) |
| Unspecified*               | 34          | 126         | 212         |
| Region                     |             |             |             |
| East                       | 519,693 (32) | 828,698 (36) | 1,058,641 (30) |
| Midwest                    | 676,579 (42) | 935,961 (41) | 1,049,693 (30) |
| South                      | 117,605 (7)  | 137,539 (6)  | 802,302 (23)  |
| West                       | 312,760 (19) | 379,619 (17) | 579,409 (17)  |
| Total                      | 1,626,637    | 2,281,817    | 3,490,045     |
| **Adult**                  |             |             |             |
| Age group (years)          |             |             |             |
| 20–29                      | 607,718 (15) | 863,664 (15) | 1,351,846 (15) |
| 30–39                      | 757,898 (19) | 1,088,520 (19) | 1,682,856 (19) |
| 40–49                      | 1,119,239 (28) | 1,576,415 (27) | 2,333,139 (26) |
| 50–59                      | 1,120,411 (28) | 1,674,602 (29) | 2,600,671 (29) |
| 60–64                      | 389,617 (10)  | 654,923 (11)  | 1,079,918 (12) |
| Sex                        |             |             |             |
| Female                     | 2,093,637 (52) | 3,098,409 (53) | 4,784,436 (53) |
| Male                       | 1,901,232 (48) | 2,759,416 (47) | 4,263,838 (52) |
| Unspecified*               | 14          | 299         | 156         |
| Region                     |             |             |             |
| East                       | 1,261,896 (32) | 2,132,166 (36) | 2,825,566 (31) |
| Midwest                    | 1,605,019 (40) | 2,312,937 (39) | 2,697,553 (30) |
| South                      | 324,678 (8)  | 404,490 (7)  | 2,151,521 (24) |
| West                       | 803,290 (20)  | 1,008,531 (17) | 1,373,790 (15) |
| Total                      | 3,994,883    | 5,858,124    | 9,048,430     |
| Overall total              | 5,621,520    | 8,139,941    | 12,538,475    |

* Unspecified numbers on gender are not included in percentages
region-specific prevalence estimates are provided in
Tables 2 and 3.

Demographic Distribution of IBD in 2009

Age and region were significantly associated with the
prevalence of CD and UC in both the pediatric and adult
populations (Table 4). Differences in sex were significant
only for CD.

Effect of Age The prevalence of both CD and UC
increased with increasing age, although for CD the preva-
ience remained fairly stable between 30 and 50 years of
age (Figs. 1, 2).

Effect of geographic region There was significant re-
gional variation for both children and adults with CD and UC.
For both conditions, the prevalence was lower in the South
and West, as compared with the Northeast and Midwest,
except for UC among children where only the Northeast
had a significantly higher prevalence (Table 4).

Effect of sex For children, the prevalence of CD was
lower in girls than in boys [odds ratio (OR) 0.88; 95 % CI
0.81–0.97]. No significant difference in sex was observed
for UC (OR 0.93; 95 % CI 0.83–1.04). In adults, the
prevalence of CD was higher among women (OR 1.17;
95 % CI 1.14–1.20), with no differences observed in UC
(OR 1.00; 95 % CI 0.97–1.03) (Table 4).

Extrapolation to US Population

After standardizing the data from the data set on age,
gender, and region according to 2009 US census data, we
estimate that 1,171,000 Americans have IBD (565,000 CD
and 593,000 UC). Of these, 62,000 are children (38,000
CD and 23,000 UC). Thus, 5 % of all IBD cases in the
USA are of pediatric age (<20 years).

Time Trends in IBD Prevalence

Standardized prevalence estimates based on 2009 census data
for each time period are shown in Table 5. The prevalence of
both CD and UC increased slightly in both the pediatric and
adult populations between 2004–2005 and 2008–2009
(P < 0.001 for all comparisons). Among those individuals
younger than 20 years, standardized CD prevalence esti-
mates increased from 43 to 48 per 100,000, and UC preva-
lence increased from 27 to 29 per 100,000 (Figs. 3, 4). Among
those individuals 20 years of age and older, CD prevalence
increased from 214 to 236 per 100,000 and UC prevalence
increased from 235 to 248 per 100,000 (Figs. 3, 4).

Discussion

In our source population of 12,538,475 commercially
insured Americans during the period 2008–2009, the

### Table 2 Prevalence of Crohn’s disease in a commercially insured US population (2008–2009)

| Age group | Region | Prevalence per 100,000 (95 % CI) |
|-----------|--------|---------------------------------|
| Pediatric (age <20 years) | Northeast | 65.1 (60.4–70.2) |
| | South | 54.1 (49.2–59.5) |
| | Midwest | 61.4 (56.8–66.3) |
| | West | 43.2 (38.0–48.9) |
| | Overall | 57.8 (55.3–60.4) |

### Table 3 Prevalence of ulcerative colitis in a commercially insured US population (2008–2009)

| Age group | Region | Prevalence per 100,000 (95 % CI) |
|-----------|--------|---------------------------------|
| Pediatric (age <20) | Northeast | 38.3 (34.7–42.2) |
| | South | 33.4 (29.6–37.7) |
| | Midwest | 31.5 (28.3–35.2) |
| | West | 31.1 (26.8–36.0) |
| | Overall | 33.9 (32.0–35.9) |

### CI Confidence interval

* Exact 95 % confidence limits were calculated using a variation on the Wilson Score confidence interval method
prevalence of IBD among individuals aged 20 years and older was estimated to be 241 per 100,000 for CD and 263 per 100,000 for UC. We also estimated the prevalences for the pediatric population (<20 years of age) to be 58 per 100,000 for CD and 34 per 100,000 for UC. When extrapolating to the US population in 2009 based on census data, we estimated that approximately 1,171,000 Americans have IBD (565,000 CD and 593,000 UC). Furthermore, the prevalences of both CD and UC have slightly increased, as is to be expected for conditions without cure and low mortality.

Our overall CD and UC prevalence estimates in 2009 were somewhat higher than most of those reported for

| Demographic characteristics | CD OR (95% CI) | P value | UC OR (95% CI) | P value |
|----------------------------|---------------|---------|---------------|---------|
| Pediatric (age <20 years)  |               |         |               |         |
| Female                     | 0.88 (0.81–0.97) | 0.0059  | 0.93 (0.83–1.04) | 0.1891  |
| Age, years (reference = 15–19 years) |      |         |               |         |
| <5                         | 0.05 (0.03–0.07) | <0.0001 | 0.08 (0.05–0.12) | <0.0001 |
| 5–10                       | 0.10 (0.08–0.12) | <0.0001 | 0.16 (0.13–0.20) | <0.0001 |
| 10–14                      | 0.37 (0.34–0.41) | <0.0001 | 0.43 (0.38–0.50) | <0.0001 |
| Region (reference South)   |               |         |               |         |
| Northeast                  | 1.19 (1.06–1.34) | 0.0002  | 1.14 (0.97–1.33) | 0.0177  |
| Midwest                    | 1.16 (1.03–1.31) | <0.0028 | 0.96 (0.82–1.13) | 0.2745  |
| West                       | 0.85 (0.72–0.99) | <0.0001 | 0.98 (0.81–1.18) | 0.5258  |
| Adult (age ≥20 years)      |               |         |               |         |
| Female                     | 1.17 (1.14–1.20) | <0.0001 | 1.00 (0.97–1.03) | 0.9682  |
| Age, years (reference = 60–64 years) |      |         |               |         |
| 20–29                      | 0.73 (0.70–0.77) | <0.0001 | 0.45 (0.43–0.48) | <0.0001 |
| 30–39                      | 0.84 (0.80–0.88) | 0.1172  | 0.66 (0.63–0.69) | <0.0001 |
| 40–49                      | 0.83 (0.80–0.87) | 0.0648  | 0.71 (0.68–0.74) | 0.5788  |
| 50–59                      | 0.88 (0.84–0.92) | 0.0148  | 0.83 (0.80–0.87) | <0.0001 |
| Region (reference = South) |               |         |               |         |
| Northeast                  | 1.13 (1.10–1.17) | <0.0001 | 1.25 (1.20–1.29) | <0.0001 |
| Midwest                    | 1.21 (1.16–1.25) | 0.0001  | 1.22 (1.18–1.26) | <0.0001 |
| West                       | 1.01 (0.96–1.05) | <0.0001 | 1.19 (1.14–1.24) | <0.0653 |

OR Odds ratio, CD Crohn’s disease, UC ulcerative colitis

*a ORs and 95% CIs were derived from multivariate logistic regression, after controlling for each of the other variables listed

Fig. 1 Age-specific prevalence of Crohn’s disease (CD) per 100,000 persons in a commercially insured US population, 2008–2009

Fig. 2 Age-specific prevalence of ulcerative colitis (UC) per 100,000 persons in a commercially insured US population, 2008–2009
other recent US and Canadian studies [1–3, 5, 13], which may reflect the increasing prevalence of IBD over time (especially CD), given that these other studies did not use data subsequent to 2004. The higher prevalences in our study compared to previous studies could also have resulted from our exclusion of adults aged over 65 years; the other studies included all adults 18 years and over. Our prevalence estimates are consistent with recent data from Olmsted County, MN which report estimates of 273 per 100,000 for UC and 222 per 100,000 for CD in 2005. As in our study, data from Olmsted County also demonstrate increasing IBD prevalence in recent years. The prevalence of CD has been reported to have increased between 1991 and 2001 [5] and that of both CD and UC to have increased between 2001 and 2005 [14].

With regards to pediatric-specific prevalence data, our estimates are between the prevalence estimated by Abramson et al. in northern California (12 and 20 per 100,000 for CD and UC, respectively) and by Loftus et al. in Olmsted County, MN (115 and 107 per 100,000 for CD and UC, respectively) [5, 6]. Our estimates are similar to those reported in a Canadian study, where Bernstein et al. reported pediatric-specific prevalence estimates ranging from 31 to 71 per 100,000 for CD and 18–31 per 100,000 for UC [13].

In our study we observed increasing prevalence with age for children and adults with CD and UC; however, the prevalence of CD leveled off somewhat for adults over 30. These age effects are consistent with previous studies [1–3, 6]. We also observed a higher prevalence among males than females in pediatric CD, with the opposite occurring in the adult population. We did not find any significant differences in the prevalence of UC among the sexes for both adults and children. These results are similar to those reported in other US and Canadian studies [3, 4, 6, 13].

The major strengths of this study are the large sample size and the geographic diversity of the sample. These strengths allowed us to provide up-to-date estimates of the prevalence of IBD in a US population and describe differences by geography, age, sex, and time period. Because claims data lack clinical details, the results of this study are subject to some degree of diagnostic misclassification. To reduce the effects of misclassification, we used strict case definitions which included the need of several healthcare contacts associated with IBD specific ICD 9 codes as well as pharmacy claims for IBD medications. Although similar

### Table 5

| Age group     | 2004–2005 | 2006–2007 | 2008–2009 |
|---------------|-----------|-----------|-----------|
|               | CD (95% CI) | UC (95% CI) | CD (95% CI) | UC (95% CI) | CD (95% CI) | UC (95% CI) |
| Pediatric (age <20 years) | | | | |
| Prevalence per 100,000 | 43.0 (42.6–43.5) | 27.3 (27.0–27.7) | 45.2 (44.7–45.6) | 27.5 (27.2–27.9) | 48.0 (47.5–48.5) | 28.7 (28.3–29.0) |
| Adult (age ≥20 years) | | | | |
| Prevalence per 100,000 | 213.6 (213.0–214.3) | 234.9 (234.2–235.6) | 230.3 (229.5–230.9) | 246.5 (245.8–247.2) | 235.6 (234.9–236.3) | 248.1 (247.4–248.9) |

CI confidence interval, CD Crohn’s disease, UC ulcerative colitis

* Exact 95% confidence limits were calculated using a variation on the Wilson Score confidence interval method

**Fig. 3** Standardized prevalence estimates of CD per 100,000 persons in a commercially insured US population by age group, 2004–2009

**Fig. 4** Standardized prevalence estimates of UC per 100,000 persons in a commercially insured US population by age group, 2004–2009
to administrative case definitions that have been previously validated in the USA and Canada [1, 2, 13], the specific case definition has not been validated in this database.

Additionally, because individuals in the USA change health plans approximately every 2 years, we divided our analysis into distinct 2-year intervals where the source population for each period was the population of plan enrollees with continuous insurance coverage. This may have led to an underestimation of prevalent cases with low healthcare utilization. However, any misclassification would likely have been stable over the time periods studied, thus not affecting the time trend analyses. Another limitation to this study is the lack of data on race/ethnicity, socioeconomic status, or precise geographic location of the study population; thus, we were not able to evaluate the impact of these factors on IBD epidemiology. A final limitation is that this study included only individuals with commercial insurance who were younger than 65 years; thus, Medicaid and self-insured patients were not included. As such, our findings may not be completely generalizable to the broader US population and may lack comparability to other studies.

In conclusion, the burden of IBD in the USA continues to increase, owing largely to the continued steady increase in prevalence for these chronic conditions. The results of our study, using insurance claims data, suggest that nearly 1.2 million Americans are living with IBD. This up-to-date epidemiological data may be used to support disease surveillance and inform policy, including anticipating the need for clinical services by this patient population, establishing research priorities, and supporting IBD advocacy efforts.

Acknowledgments We thank Charlotte Carroll for her contributions to the study. Support for this study was provided by GlaxoSmithKline.

Conflict of interest None.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

1. Herrinton LJ, Liu L, Lafata JE, et al. Estimation of the period prevalence of inflammatory bowel disease among nine health plans using computerized diagnoses and outpatient pharmacy dispensings. *Inflamm Bowel Dis*. 2007;13:451–461.

2. Herrinton LJ, Liu L, Lewis ID, Griffin PM, Allison J. Incidence and prevalence of inflammatory bowel disease in a northern California managed care organization, 1996–2002. *Am J Gastroenterol*. 2008;103:1998–2006.

3. Kappelman MD, Rifas-Shiman SL, Kleinman K, et al. The prevalence and geographic distribution of Crohn’s disease and ulcerative colitis in the United States. *Clin Gastroenterol Hepatol*. 2007;5:1424–1429.

4. Kugathasan S, Judd RH, Hoffmann RG, et al. Epidemiologic and clinical characteristics of children with newly diagnosed inflammatory bowel disease in Wisconsin: a statewide population-based study. *J Pediatr*. 2003;143:525–531.

5. Loftus CG, Loftus EV Jr, Harmsen WS, et al. Update on the incidence and prevalence of Crohn’s disease and ulcerative colitis in Olmsted county, Minnesota, 1940–2000. *Inflamm Bowel Dis*. 2007;13:254–261.

6. Abramson O, Durant M, Mow W, et al. Incidence, prevalence, and time trends of pediatric inflammatory bowel disease in northern California, 1996 to 2006. *J Pediatr*. 2010;157:233.e1–239.e1.

7. Stempel DA, Mauskopf J, McLaughlin T, Yazdani C, Stanford RH. Comparison of asthma costs in patients starting fluticasone propionate compared to patients starting montelukast. *Respir Med*. 2001;95:227–234.

8. Bernstein CN, Blanchard JF, Rawsthorne P, Wajda A. Epidemiology of Crohn’s disease and ulcerative colitis in a central canadian province: a population-based study. *Am J Epidemiol*. 1999;149:916–924.

9. Kappelman MD, Rifas-Shiman SL, Porter CQ, et al. Direct health care costs of Crohn’s disease and ulcerative colitis in US children and adults. *Gastroenterology*. 2008;135:1907–1913.

10. Kappelman MD, Porter CQ, Galanko JA, et al. Utilization of healthcare resources by US children and adults with inflammatory bowel disease. *Inflamm Bowel Dis*. 2011;17:62–68.

11. Ollendorf DA, Lidsky L. Infliximab drug and infusion costs among patients with Crohn’s disease in a commercially-insured setting. *Am J Ther*. 2006;13:502–506.

12. Cohen RD, Waters HC, Tang B, Rahman MI. Effects of fistula on healthcare costs and utilization for patients with Crohn’s disease treated in a managed care environment. *Inflamm Bowel Dis*. 2008;14:1707–1714.

13. Bernstein CN, Wajda A, Svenson LW, et al. The epidemiology of inflammatory bowel disease in Canada: a population-based study. *Am J Gastroenterol*. 2006;101:1559–1568.

14. Ingle SB, Loftus EV, Tremaine WJ, et al. Increasing incidence and prevalence of inflammatory bowel disease in Olmsted county, Minnesota, during 2001–2004. *Gastroenterology*. 2007;132:A19–A20.