Research sponsored by the Crohn's & Colitis Foundation.

Opportunistic Infections Are More Prevalent in Crohn's Disease and Ulcerative Colitis: A Large Population-Based Study

Mohammed Zaahid Sheriff, MD, Emad Mansoor, MD,† Jay Luther, MD,‡ Ashwin N. Ananthakrishnan, MD, MPH,‡ Mohannad Abou Saleh, MD,* Edith Ho, MD,† Farren B. S. Briggs, PhD, ScM,§ and Maneesh Dave, MD, MPH†*

Background: Opportunistic infections (OIs) are more common in patients with inflammatory bowel disease (IBD); however, there have been limited large-scale studies of OIs in IBD. We investigated the epidemiological characteristics of OI in Crohn's disease (CD) and ulcerative colitis (UC) using a large population-based database.

Methods: Data were collected from a commercial database (Explorys Inc., Cleveland, OH, USA) that provided electronic health records from 26 major integrated US health care systems from 1999 to March 2018. In this data set, we identified all CD and UC patients, based on Systemized Nomenclature of Medicine–Clinical Terms. Within these cohorts, we identified a variety of OIs and compared the prevalence rate of OI in individuals with IBD with that of controls (patients in the database between March 2013 and March 2018 without the diagnosis of IBD).

Results: Explorys included 153,290 patients with CD and 128,540 patients with UC between March 2013 and March 2018. The prevalence of OIs was 17.8% in CD, 19.2% in UC, and 7% in non-IBD controls. When compared with non-IBD controls, all OIs were more common in CD (prevalence ratio [PR], 2.54; 95% confidence interval [CI], 2.51–2.57) and UC (PR, 2.74; 95% CI, 2.71–2.77). Overall, viral infections were numerically more common, whereas bacterial infections had the highest PRs in CD and UC when compared with controls without IBD.

Conclusions: We found significantly higher rates of OI in IBD. Our study suggests the need for close follow-up of IBD patients to diagnose and provide vaccinations where applicable for prevention of infections.

Key Words: infections, Crohn's disease, ulcerative colitis, database

INTRODUCTION

The inflammatory bowel diseases (IBDs), Crohn's disease (CD) and ulcerative colitis (UC), are chronic inflammatory diseases of the intestinal tract characterized by an exaggerated systemic immune response.[1] For this reason, the cornerstone of therapy for these diseases is immunosuppressive (IS) agents that temper the host inflammatory response.[2, 3] By virtue of this mechanism, these agents have been associated with an increased risk of opportunistic infections (OIs).[4, 5] In parallel, patients with IBD commonly exhibit other risk factors for OIs, including malnutrition, older age, and chronic medical disease such as diabetes.[5] The OIs in IBD encompass bacterial infections (tuberculosis, nocardiosis, *Clostridium difficile* infection, pneumococcal infection, legionellosis, and listeriosis), fungal infections (histoplasmosis, cryptococcosis, *Pneumocystis jirovecii* infection, aspergillosis, and candidiasis), and viral infections (herpes simplex virus, human papilloma virus, influenza virus).[4, 5] Although this propensity for increased risk of infection in IBD patients is widely appreciated, the exact risk has not been comprehensively defined in the United States. In fact, to our knowledge, there have been no large-scale studies that comprehensively define the epidemiology of OIs in IBD in the United States. Given the high burden of IBD in the United States,[6, 7] we sought to determine, using a large population-based commercial database, the overall prevalence of multiple OIs in CD and UC and to further characterize the distribution of these OIs based on certain clinical characteristics.

Received for publications December 11, 2018; Editorial Decision May 31, 2019.

From the *Department of Internal Medicine, University Hospitals Cleveland Medical Center, Case Western Reserve University, Cleveland, Ohio, USA; †Division of Gastroenterology and Liver Disease, University Hospitals Cleveland Medical Center, Case Western Reserve University, Cleveland, Ohio, USA; ‡Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts; §Department of Population and Quantitative Health Sciences, School of Medicine, Case Western Reserve University, Cleveland, Ohio, USA

Supported by: Maneesh Dave is supported by Crohn's and Colitis Foundation Career Development Award 370615 and K08DK110421.

Conflicts of interest: The authors have no potential conflicts (financial, professional, or personal) to disclose.

Address correspondence to: Maneesh Dave, MD, MPH, Division of Gastroenterology and Liver Disease, University Hospitals Cleveland Medical Center, Case Western Reserve University, 11100 Euclid Avenue, Cleveland, OH 44106 (maneeshdave@gmail.com).

doi: 10.1093/ibd/izz147
Published online 17 July 2019

METHODS

Database

We performed a retrospective analysis of a large population-based commercial database (Explorys Inc., Cleveland, OH, USA). The data set contains a collection of electronic health

record (EHR) data from 26 major integrated health care systems spread over 50 states in the United States from 1999 to 2018.[8] The data of more than 50 million patients, approximately 15% of the population, over all 4 United States census regions are included.[9] Explorys contains de-identified patient data from participating institutions and uses a health data gateway (HDG) server behind the firewall of each participating health care organization that collects de-identified data from various health information systems' EHRs using billing inquiries. Data are then standardized and normalized by Explorys. Additionally, the patient matching engine ensures that each patient is represented only once.[10] Diagnoses, findings, and procedures are mapped into the US edition of the Systematized Nomenclature Of Medicine–Clinical Terms (SNOMED-CT) hierarchy. Each participating health care institution has access to Explorys online (password protected), which allows browsing of the data for all participating health care institutions. Explorys data are automatically updated at least once every 24 hours.[8] To prevent the identification of individual patient data through combinations of specific SNOMED-CT attributes, cohort information is statistically de-identified using rounding: All numbers are rounded to the nearest 10. Furthermore, Explorys is a Health Insurance Portability and Accountability Act (HIPAA)–compliant platform and thus is exempt from institutional review board (IRB) review.[8, 11]

Patient Selection

Using the Explorys search tool, we identified a cohort of UC and CD within the period March 2013 to March 2018. Crohn's disease patients were defined as those having a SNOMED-CT diagnosis of "Crohn's disease," and UC patients were defined as those having a diagnosis of "ulcerative colitis." The specific codes that are represented by these general terms are represented in Supplementary Table 1. The controls used were the remaining patients in the database from March 2013 to March 2018 without a diagnosis of "Crohn's disease" and without a diagnosis of "ulcerative colitis." Within this aggregated cohort, patients with OIs were identified by SNOMED-CT diagnosis codes (Supplementary Table 2). Although international classification of diseases, ninth revision (ICD-9) and SNOMED-CT are both medical terminology systems for recording medical diagnoses and concepts, SNOMED-CT has many more concepts to be coded per clinical document than ICD-9,[11] which makes it more accurate and comprehensive in terms of enlisting pertinent clinical information.[12, 13] Our group has successfully used the Explorys database to study eosinophilic colitis, eosinophilic esophagitis, myocardial infarction in IBD, multiple sclerosis, and the prevalence of colorectal cancer in the elderly.[10, 14–16] Although validation of the SNOMED-CT codes "Crohn's disease" and "ulcerative colitis" has not been performed, prior studies have looked at conditions such as Hidradenitis Suppruitiva, which has a 1:1 mapping of the SNOMED-CT code to the ICD-9 code, and found a positive predictive value of 79.3% and accuracy of 90%.[9]

Statistical Analysis

For patients with CD and UC, demographics were characterized by descriptive statistics. Similarly, for each OI, demographics were described by descriptive statistics. Univariate analysis was performed to assess the differences in prevalence of the OIs between those with CD, UC, and controls (individuals without CD and UC) by calculating the prevalence ratio (PR) and 95% confidence interval (CI). Sex- and race-adjusted PRs and confidence intervals were also calculated using the Open Source Epidemiologic Statistics for Public Health software tool with a 2×2 table to calculate 2-tailed Fisher exact *P* values and 2-tailed Mantel-Haenszel chi-square *P* values.[17] For OIs in which the counts (numbers) were sufficiently large ($n \geq 5$), we report the 2-tailed Mantel Haenszel chi-square *P* values.

To calculate the overall period prevalence, we identified all patients in the database with CD and UC between March 2013 and March 2018. We then divided this number by the total number of patients in the database (from March 2013 to March 2018), thus ensuring that all patients were in the denominator (source population) if they had the disease. We further subdivided the patients into 3 age groups, children (age <18 years), adults (age 18–65 years), and elderly (age >65 years), to characterize the distribution of OIs by age.

To determine the prevalence of an OI in IBD, we identified the total number of OI cases diagnosed between March 2013 and March 2018 in CD and UC subcohorts and divided it by the total number of patients with CD and total number of patients with UC, respectively (from March 2013 to March 2018). Similarly, age- and sex-specific prevalence rates were calculated. In comparing age-specific prevalence, children (age <18 years) and elderly (age >65 years) were compared with adults (age 18–65 years) for each OI by IBD diagnosis. To calculate sex-specific prevalence, males were compared with females for each OI by IBD diagnosis. We also calculated the overall prevalence of fungal infections by adding the cases of histoplasma, cryptococcus, aspergillosis, and candidiasis; bacterial infection by adding the cases of listeriosis, pneumococcal disease, nocardiosis, *Clostridium difficile*, and tuberculosis; and viral diseases by adding the cases of herpes simplex virus (HSV), human papillomavirus (HPV), Epstein-Barr virus (EBV), and cytomegalovirus (CMV) in CD and UC, respectively. In addition, we identified cases of OI in IBD and non-IBD controls in whom there was HIV coinfection with the SNOMED-CT code "human immunodeficiency virus infection." For OIs with coinfection with HIV, we performed an analysis excluding HIV.

RESULTS

Prevalence of Inflammatory Bowel Disease

Among the 35,420,110 individuals in the database between March 2013 and March 2018, we identified 153,290 patients with CD and 128,540 patients with UC. The crude

prevalence (per 100,000) of CD was 432.9, and the crude prevalence (per 100,000) of UC was 362.9. The prevalence of CD was higher in females than in males at 473 per 100,000 vs 384 per 100,000 (PR, 1.23; 95% CI, 1.22–1.24; $P < 0.0001$). Similarly, the prevalence of UC was higher in females than in males, at 388 per 100,000 vs 333 per 100,000 (PR, 1.17; 95% CI, 1.15–1.18; $P < 0.0001$). The prevalence of CD and UC was higher in Caucasians compared with African Americans (PR, 1.64; 95% CI 1.61–1.67; $P < 0.0001$; and PR, 1.98; 95% CI, 1.94–2.03; $P < 0.0001$; respectively). The highest prevalence for CD was in adults (18–65 years), whereas the highest prevalence of UC was in elderly (age >65 years). Interestingly, even though the overall prevalence of CD was higher than that of UC, in the elderly cohort (age >65 years), the prevalence of UC was higher than CD. These baseline demographic data are presented in Table 1. The 5-year age interval–based prevalence of CD and UC further demonstrated this finding and showed that the prevalence of UC was lower than CD until the age of 69 years, and at the age bracket 70–74 years, there was a switch and subsequently UC became more prevalent than CD (Fig. 1).

Overall Prevalence of Opportunistic Infections

We identified 27,300 cases of OI in CD and 24,690 cases of OI in UC between March 2013 and March 2018. The prevalence of OIs was 17.8% in CD, 19.2% in UC, and 7% in non-IBD controls. When compared with the non-IBD controls, all OIs were more common in CD (PR, 2.54; 95% CI, 2.51–2.57) and UC (PR, 2.74; 95% CI, 2.71–2.77). Overall, viral infections were numerically more common, whereas bacterial infections had the highest risk ratios in CD and UC when compared with controls without IBD (Table 2).

The individual prevalence rates for all OIs are described in Table 3. The OIs with the highest occurrence in IBD patients were *Clostridium difficile* (CD: PR, 11.5; 95% CI, 11.1–11.8; UC: PR, 17.2; 95% CI, 16.8–17.7) and CMV (CD: PR, 10.4; 95% CI, 9.4–11.6; UC: PR, 14.6; 95% CI, 13.2–16.1). The top 3 most prevalent OIs numerically in IBD patients were candidiasis (5858/100,000 in CD and 5702/100,000 in UC), *Clostridium difficile* (2596/100,000 in CD and 3905/100,000 in UC), and human papillomavirus (3901/100,000 in CD and 4092/100,000 in UC). The most prevalent opportunistic bacterial infection was *Clostridium difficile* (2596/100,000 in CD and 3905/100,000 in UC), followed by pneumococcal disease (296.6/100,000 in CD and 303.4/100,000 in UC) and legionella (19.6/100,000 in CD and 31.1/100,000 in UC). The most prevalent viral infections were HPV (3901/100,000 in CD and 4092/100,000 in UC), influenza (2322/100,000 in CD and 2342/100,000 in UC), and HSV (1938/100,000 in CD and 1914/100,000 in UC). The most prevalent fungal infections were candidiasis (5858/100,000 in CD and 5703/100,000 in UC), histoplasmosis (104.4/100,000 in CD and 93.4 /100,000 in UC), and aspergillosis (71.8/100,000 in CD and 85.6/100,000 in UC). As expected from previous published studies, *Clostridium difficile* and cytomegalovirus infections were less common in CD than UC (PR, 0.66; 95% CI, 0.64–0.69; and PR, 0.72; 95% CI, 0.62–0.83; respectively) whereas tuberculosis was more common in CD than UC (PR, 1.26; 95% CI, 1.03–1.54). On exclusion of HIV, the prevalence of OI in IBD continued to be significant, with a similar odds ratio and identical *P* value, as shown in Supplementary Tables 3 and 4. Patients with pneumocystis had the highest rates of coinfection with HIV, with 33% of CD patients and 40% of UC patients with pneumocystis having HIV. Interestingly, exclusion of HIV patients from the cohort resulted in similar albeit slightly higher PRs, though with overlapping confidence intervals of pneumocystis in IBD patients.

TABLE 1. Baseline Demographic Data of CD and UC in Explorys Between March 2013 and March 2018

	Source Population, No. (%)	CD Cases, No. (%)	CD Prevalence[a]	UC Cases, No. (%)	UC Prevalence[a]
Overall	35,420,110	153,290	432.8	128,540	362.9
Male	15,796,220 (45)	60,820 (40)	385.0	52,640 (41)	333.2
Female	19,532,410 (55)	92,440 (60)	473.3	75,880 (59)	388.5
Unknown	91,470 (0)	30 (0)	-	20 (0)	-
Age group[a]					
Children (<18 y)	5,425,580 (15)	2700 (1.8)	49.8	1240 (0.96)	22.9
Adults (18–65 y)	22,440,390 (63)	112,240 (73)	500.2	86,810 (67)	386.8
Elderly (>65 y)	7,498,510 (21)	38,370 (25)	511.7	40,580 (32)	541.2
Race					
Caucasian	22,181,890 (63)	125,660 (82)	566.5	107,000 (83)	482.4
African American	3,923,400 (11)	13,580 (8.9)	346.1	9540 (7.4)	243.2
Asian	691,630 (2.0)	1980 (1.3)	286.3	1840 (1.4)	266.0
Hispanic/Latin American	476,060 (1.3)	1000 (0.65)	652.4	920 (0.72)	715.7

[a]Prevalence reported per 100,000.

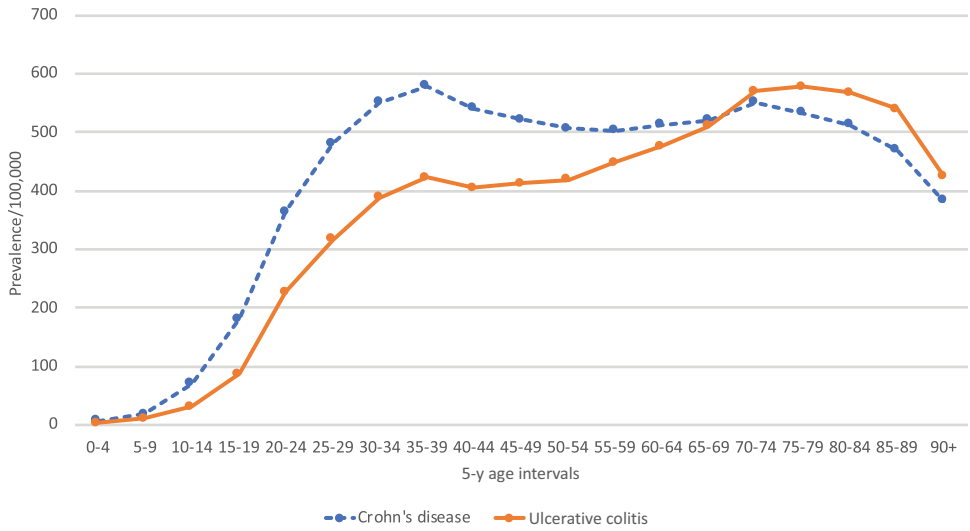

FIGURE 1. Age-specific prevalence of CD and UC between March 2013 and March 2018.

TABLE 2. Overall Prevalence of Fungal, Bacterial, and Viral Infections in CD and UC Compared With Non-IBD Control

	CD	UC	Non-IBD	Prevalence in CD[a]	Prevalence in UC[a]	Prevalence in Non-IBD[a]	PR in CD vs Non-IBD	PR in UC vs Non-IBD	PR in CD vs UC
Overall OI	27,300	24,690	2,465,010	17,809.4	19,208.0	7015.2	2.54; 95% CI, 2.51–2.57; $P < 0.0001$	2.74; 95% CI, 2.71–2.77; $P < 0.0001$	0.93; 95% CI, 0.91–0.94; $P < 0.0001$
All fungal	9350	7630	719,570	6099.5	5935.9	2047.8	2.98; 95% CI, 2.92–3.04; $P < 0.0001$	2.90; 95% CI, 2.84–2.96; $P < 0.0001$	1.03; 95% CI, 1.00–1.06; $P = 0.069$
All bacterial	4720	5630	126,190	3079.1	4390.0	359.1	8.54; 95% CI, 8.33–8.82; $P < 0.0001$	12.20; 95% CI, 11.88–12.52; $P < 0.0001$	0.70; 95% CI, 0.68–0.73; $P < 0.0001$
All viral	13,230	11,430	1,619,250	8630.7	8892.2	4608.2	1.87; 95% CI, 1.84–1.90; $P < 0.0001$	1.93; 95% CI, 1.90–1.96; $P < 0.0001$	0.97; 95% CI, 0.95–0.99; $P = 0.014$

[a]Prevalence reported per 100,000.

Prevalence of Opportunistic Infections by Age

Viral infections, that is, influenza (CD: PR, 2.47; 95% CI, 2.11–2.90; UC: PR, 1.72; 95% CI, 1.30–2.26) and EBV (CD: PR, 2.68; 95% CI, 1.71–4.21; UC: RR, 2.80; 95% CI, 1.49–5.26), were more prevalent in children (<18 years of age) when compared with adults (18 to 65 years) for both UC and CD. On the other hand, certain fungal infections, that is, aspergillosis (CD: PR, 2.93; 95% CI, 1.98–4.33; UC: PR, 3.21; 95% CI, 2.15–4.79), histoplasmosis (CD: PR, 1.6; 95% CI, 1.17–2.19; UC: PR, 2.14; 95% CI, 1.50–3.06), and bacterial infections, that is, pneumococcal disease (CD: PR, 3.34; 95% CI, 2.78–4.02; UC: PR, 3.08; 95% CI 2.51–3.76), were more prevalent in the elderly (>65 years) when compared with adults (18–65 years) for both UC and CD. Pneumocystis, however, was not more prevalent in the elderly when compared with adults (CD: PR, 1.46; 95% CI, 0.86–2.50; UC: PR, 1.07; 95% CI, 0.63–1.83; respectively).

The prevalence of OI by age for CD and UC when compared with a control group of adults is described in "Table 4A and B, respectively. In addition, the 5-year age interval prevalence rates for candidiasis, *Clostridium difficile*, pneumococcal disease, influenza, HPV, EBV, and CMV are shown in Supplementary Figure 1. Due to the overall low prevalence of nocardia, legionella, and listeria, the data in Explorys did not provide us the granularity to identify epidemiological trends by age. Similarly, we lacked detailed data in certain age ranges for OI in CD and UC, demonstrated in Table 4 as N/A, that is, missing data.

TABLE 3. Prevalence of OI in UC and CD Compared With Non-IBD Controls

	CD	UC	Non-IBD	Prevalence in CD[a]	Prevalence in UC[a]	Prevalence in Non-IBD[a]	PR in CD vs Non-IBD	PR in UC vs Non-IBD	PR in CD vs UC
Histoplasmosis	160	120	8030	104.4	93.4	22.9	4.57; 95% CI, 3.91–5.34; $P < 0.0001$	4.09; 95% CI, 3.41–4.89; $P < 0.0001$	1.12; 95% CI, 0.88–1.42; $P = 0.387$
Cryptococcosis	40	20	1410	26.1	15.6	4.0	6.50; 95% CI, 4.75–8.90; $P < 0.0001$	3.88; 95% CI, 2.49–6.03; $P < 0.0001$	1.68; 95% CI, 0.98–2.87; $P = 0.073$
Pneumocystis	60	50	3310	39.1	38.9	9.4	4.16; 95% CI, 3.22–5.36; $P < 0.0001$	4.13; 95% CI, 3.12–5.46; $P < 0.0001$	1.01; 95% CI, 0.69–1.46; $P > 0.999$
Aspergillosis	110	110	5820	71.8	85.6	16.6	4.33; 95% CI, 3.59–5.23; $P < 0.0001$	5.17; 95% CI, 4.28–6.24; $P < 0.0001$	0.84; 95% CI, 0.64–1.09; $P = 0.215$
Candidiasis	8980	7330	701,000	5858.2	5702.5	1995.0	2.94; 95% CI, 2.88–3.00; $P < 0.0001$	2.86; 95% CI, 2.80–2.92; $P < 0.0001$	1.03; 95% CI, 1.00–1.06; $P = 0.078$
Listeriosis	10	10	300	6.5	7.8	0.9	7.64; 95% CI, 4.07–14.35; $P < 0.0001$	9.11; 95% CI, 4.85–17.11; $P < 0.0001$	0.84; 95% CI, 0.35–2.01; $P = 0.861$
Legionella	30	40	4000	19.6	31.1	11.4	1.72; 95% CI, 1.20–2.46; $P = 0.008$	2.73; 95% CI, 2.00–3.73; $P < 0.0001$	0.63; 95% CI, 0.39–1.01; $P = 0.070$
Pneumococcal disease	450	390	31,310	293.6	303.4	89.1	3.30; 95% CI, 3.00–3.62; $P < 0.0001$	3.41; 95% CI, 3.08–3.76; $P < 0.0001$	0.97; 95% CI, 0.85–1.11; $P = 0.657$
Nocardiosis	10	10	630	6.5	7.8	1.8	3.64; 95% CI, 1.95–6.80; $P = 0.0001$	4.34; 95% CI, 2.32–8.10; $P = 0.0003$	0.84; 95% CI, 0.35–2.01; $P = 0.861$
Clostridium difficile	3980	5020	79,580	2596.4	3905.4	226.5	11.46; 95% CI, 11.11–11.83; $P < 0.0001$	17.24; 95% CI, 16.77–17.73; $P < 0.0001$	0.66; 95% CI, 0.64–0.69; $P < 0.0001$
Tuberculosis	240	160	10,370	156.6	124.5	29.5	5.31; 95% CI, 4.67–6.03; $P < 0.0001$	4.22; 95% CI, 3.61–4.93; $P < 0.0001$	1.26; 95% CI, 1.03–1.54; $P = 0.0270$
Influenza	3560	3010	515,960	2322.4	2341.7	1468.4	1.55; 95% CI, 1.50–1.60; $P < 0.0001$	1.60; 95% CI, 1.54–1.65; $P < 0.0001$	0.99; 95% CI, 0.95–1.04; $P = 0.735$
HPV	5980	5260	691,020	3901.1	4092.1	1966.6	1.98; 95% CI, 1.94–2.03; $P < 0.0001$	2.08; 95% CI, 2.03–2.14; $P < 0.0001$	0.95; 95% CI, 0.92–0.99; $P = 0.010$
EBV	370	290	46,360	241.4	225.6	131.9	1.83; 95% CI, 1.65–2.03; $P < 0.0001$	1.71; 95% CI, 1.52–1.92; $P < 0.0001$	1.07; 95% CI, 0.92–1.25; $P = 0.411$
CMV	350	410	7690	228.3	319.0	21.9	10.43; 95% CI, 9.37–11.61; $P < 0.0001$	14.57; 95% CI, 13.20–16.09; $P < 0.001$	0.72; 95% CI, 0.62–0.83; $P < 0.0001$
HSV	2970	2460	358,220	1937.5	1913.8	1019.5	1.90; 95% CI, 1.83–1.97; $P < 0.0001$	1.88; 95% CI, 1.81–1.95; $P < 0.0001$	1.01; 95% CI, 0.96–1.07; $P = 0.648$

[a]Prevalence reported per 100,000.

Prevalence of Opportunistic Infections by Sex

Cytomegalovirus (CD: PR, 1.28; 95% CI, 1.04–1.58; UC: PR, 1.59; 95% CI, 1.31–1.92) was more prevalent in males than in females in both CD and UC (Table 5A, 5B). On the other hand, candidiasis (CD: PR, 0.47; 95% CI, 0.45–0.49; UC: PR, 0.46; 95% CI, 0.44–0.48), influenza (CD: PR, 0.82; 95% CI,

TABLE 4. Prevalence of OI by Age Group in CD and UC as Compared With Non-IBD Controls

4a. Crohn's Disease

	Children, No.	Adults, No.	Elderly, No.	Prevalence, Children[a]	Prevalence, Adults[a]	Prevalence, Elderly[a]	PR in Children vs Adults	PR in Elderly vs Adults
Histoplasmosis	N/A	110	60	N/A	98.0	156.4	N/A	1.60; 95% CI, 1.17–2.19; $P = 0.006$
Cryptococcosis	0	20	20	0.0	17.8	52.1	1.04; 95% CI, 0.06–17.19; $P = 0.772$	2.93; 95% CI, 1.57–5.44; $P = 0.001$
Pneumocystis	N/A	40	20	N/A	35.6	52.1	N/A	1.46; 95% CI, 0.86–2.50; $P = 0.216$
Aspergillosis	0	50	50	0.0	44.5	130.3	0.42; 95% CI, 0.03–6.74; $P > 0.999$	2.93; 95% CI, 1.98–4.33; $P < 0.0001$
Candidiasis	90	6200	2690	3333.3	5523.9	7010.7	0.60; 95% CI, 0.49–0.74; $P < 0.0001$	1.27; 95% CI, 1.22–1.33; $P < 0.0001$
C. difficile	80	2630	1270	2963.0	2343.2	3309.9	1.26; 95% CI, 1.02–1.57; $P = 0.048$	1.41; 95% CI, 1.32–1.51; $P < 0.0001$
Influenza	150	2520	890	5555.6	2245.2	2319.5	2.47; 95% CI, 2.11–2.90; $P < 0.0001$	1.03; 95% CI, 0.96–1.11; $P = 0.398$
HPV	150	4810	1010	5555.6	4285.5	2632.3	1.30; 95% CI, 1.11–1.52; $P = 0.002$	0.61; 95% CI, 0.57–0.66; $P < 0.0001$
EBV	20	310	40	740.7	276.2	104.2	2.68; 95% CI, 1.71–4.21; $P = 0.0003$	0.38; 95% CI, 0.27–0.52; $P < 0.0001$
CMV	N/A	260	90	N/A	231.6	234.6	N/A	1.01; 95% CI, 0.80–1.29; $P = 0.960$
HSV	70	2350	550	2592.6	2093.7	1433.4	1.24; 95% CI, 0.98–1.57; $P = 0.093$	0.68; 95% CI, 0.62–0.75; $P < 0.0001$
Pneumococcal disease	N/A	210	240	N/A	187.1	625.5	N/A	3.34; 95% CI, 2.78–4.02; $P < 0.0001$
Tuberculosis	N/A	170	70	N/A	195.8	172.5	N/A	1.20; 95% CI, 0.91–1.59; $P = 0.218$

4b. Ulcerative Colitis

	Children, No.	Adults, No.	Elderly, No.	Prevalence, Children[a]	Prevalence, Adults[a]	Prevalence, Elderly[a]	PR in Children vs Adults	PR in Elderly vs Adults
Histoplasmosis	N/A	60	60	N/A	69.1	147.9	N/A	2.14; 95% CI, 1.50–3.06; $P < 0.0001$
Cryptococcosis	0	10	N/A	0.0	11.5	N/A	3.50; 95% CI, 0.20–59.88; $P = 0.277$	N/A
Pneumocystis	N/A	40	20	0.0	46.1	49.3	N/A	1.07; 95% CI, 0.63–1.83; $P = 0.902$
Aspergillosis	0	40	60	0.0	46.1	147.9	0.88; 95% CI, 0.05–14.22; $P = 0.874$	3.21; 95% CI, 2.15–4.79; $P < 0.0001$

TABLE 4. Continued

4b. Ulcerative Colitis

	Children, No.	Adults, No.	Elderly, No.	Prevalence, Children[a]	Prevalence, Adults[a]	Prevalence, Elderly[a]	PR in Children vs Adults	PR in Elderly vs Adults
Candidiasis	60	4590	2690	4838.7	5287.4	6628.9	0.92; 95% CI, 0.71–1.17; $P = 0.531$	1.25; 95% CI, 1.20–1.31; $P < 0.0001$
C. difficile	70	2760	2190	5645.2	3179.4	5396.7	1.78; 95% CI, 1.41–2.24; $P < 0.0001$	1.70; 95% CI, 1.61–1.79; $P < 0.0001$
Influenza	50	2040	920	4032.3	2350.0	2267.1	1.72; 95% CI, 1.30–2.24; $P = 0.0005$	0.96; 95% CI, 0.89–1.04; $P = 0.361$
HPV	60	4190	1020	4838.7	4826.6	2513.6	1.00; 95% CI, 0.78–1.29; $P > 0.999$	0.52; 95% CI, 0.49–0.56; $P < 0.0001$
EBV	10	250	40	806.5	288.0	98.6	2.80; 95% CI, 1.49–5.26; $P = 0.008$	0.34; 95% CI, 0.25–0.48; $P < 0.0001$
CMV	N/A	290	120	N/A	334.1	295.7	N/A	0.89; 95% CI, 0.72–1.10; $P = 0.283$
HSV	20	1890	560	1612.9	2177.2	1380.0	0.74; 95% CI, 0.48–1.15; $P = 0.201$	0.63; 95% CI, 0.58–0.70; $P < 0.0001$
Pneumococcal disease	N/A	160	230	N/A	184.3	566.8	N/A	3.08; 95% CI, 2.51–3.76; $P < 0.0001$
Tuberculosis	0	100	60	N/A	115.2	147.9	0.35; 95% CI, 0.02–5.63; $P > 0.999$	1.28; 95% CI, 0.93–1.77; $P = 0.150$

[a]Prevalence reported per 100,000.

0.76–0.87; UC: PR, 0.85; 95% CI, 0.79–0.91), HPV (CD: PR, 0.54; 95% CI, 0.51–0.57; UC: PR, 0.47; 95% CI, 0.44–0.50), and HSV (CD: PR, 0.48; 95% CI, 0.45–0.53; UC: PR, 0.46; 95% CI, 0.42–0.51) were more prevalent in females in both CD and UC (Table 5A, 5B). In CD, *C. difficile* (PR, 0.92; 95% CI, 0.86–0.98) was less prevalent in males; however, in UC, there was a numerical trend toward lower prevalence, but it did not reach statistical significance. Pneumocystis, aspergillosis, and legionellosis were more prevalent in males with CD but not in UC. There were no sex differences in the prevalence rates of histoplasmosis, cryptococcus, pneumococcal disease, EBV, and tuberculosis. Due to the low prevalence of nocardia and listeria, the data in Explorys did not enable us to identify epidemiological trends by sex for these infections.

DISCUSSION

In this large, geographically diverse study of nonselected patients, we found a higher prevalence of multiple OIs, including fungal, bacterial, and viral infections in both CD and UC in comparison with non-IBD controls. Furthermore, our subgroup analyses highlighted the impact of age and sex on the prevalence of OIs in IBD. To our knowledge, this is the largest study conducted in the United States that estimates the

TABLE 5. Prevalence of OI by Sex in CD and UC

5a. Crohn's Disease

	Males, No.	Females, No.	Prevalence, Males[a]	Prevalence, Females[a]	PR in Males vs Females
Histoplasmosis	70	100	115.1	108.2	1.06; 95% CI, 0.78–1.44; $P = 0.746$
Cryptococcus	20	20	32.9	21.6	1.52; 95% CI, 0.82–2.82; $P = 0.242$
Pneumocystis	40	20	43.3	32.9	3.04; 95% CI, 1.78–5.20; $P < 0.0001$
Aspergillosis	60	50	64.9	82.2	1.82; 95% CI, 1.25–2.65; $P = 0.002$
Candidiasis	2110	6870	3469.3	7431.8	0.47; 95% CI, 0.45–0.49; $P < 0.0001$
Legionellosis	20	10	32.9	10.8	3.04; 95% CI, 1.42–6.49; $P = 0.005$
Pneumococcal disease	180	270	296.0	292.1	1.01; 95% CI, 0.84–1.22; $P = 0.927$
C. difficile	1500	2480	2466.3	2682.8	0.92; 95% CI, 0.86–0.98; $P = 0.009$
Influenza	1240	2310	2038.8	2498.9	0.81; 95% CI, 0.76–0.87; $P < 0.0001$
HPV	1560	4420	2564.9	4781.5	0.54; 95% CI, 0.41–0.57; $P < 0.0001$
EBV	140	230	230.2	248.8	0.93; 95% CI, 0.75–1.14; $P = 0.502$
CMV	160	190	263.1	205.5	1.28; 95% CI, 1.04–1.58; $P = 0.025$
HSV	720	2260	1183.8	2444.8	0.48; 95% CI, 0.45–0.53; $P < 0.0001$
Tuberculosis	110	130	209.0	171.3	1.29; 95% CI, 1.00–1.66; $P = 0.061$

5b. Ulcerative Colitis

	Males, No.	Females, No.	Prevalence, Males[a]	Prevalence, Females[a]	PR in Males vs Females
Histoplasmosis	40	70	76.0	92.3	0.82; 95% CI, 0.56–1.22; $P = 0.378$
Cryptococcus	10	10	19.0	13.2	1.44; 95% CI, 0.60–3.46; $P = 0.547$
Pneumocystis	30	30	39.5	57.0	1.44; 95% CI, 0.87–2.39; $P = 0.197$
Aspergillosis	50	60	95.0	79.1	1.20; 95% CI, 0.83–1.75; $P = 0.388$
Candidiasis	1760	5560	3343.5	7327.4	0.46; 95% CI, 0.43–0.48; $P < 0.0001$
Legionellosis	20	20	38.0	26.4	1.44; 95% CI, 0.78–2.68; $P = 0.316$
Pneumococcal disease	160	230	304.0	303.1	1.00; 95% CI, 0.82–1.23; $P > 0.999$
C. difficile	2000	3020	3799.4	3980.0	0.95; 95% CI, 0.90–1.01; $P = 0.100$
Influenza	1110	1890	2108.7	2490.8	0.85; 95% CI, 0.79–0.91; $P < 0.0001$
HPV	1300	3960	2469.6	5218.8	0.47; 95% CI, 0.44–0.50; $P < 0.0001$
EBV	120	170	228.0	224.0	1.02; 95% CI, 0.81–1.29; $P = 0.929$
CMV	220	200	417.9	263.6	1.59; 95% CI, 1.31–1.92; $P < 0.0001$
HSV	600	1870	1139.8	2464.4	0.46; 95% CI, 0.42–0.51; $P < 0.0001$
Tuberculosis	70	90	133.0	118.6	1.12; 95% CI, 0.82–1.53; $P = 0.522$

[a]Prevalence reported per 100,000.

prevalence of OIs in IBD while also comprehensively defining the effects of age and sex on prevalence.

We report that fungal infections, such as *Candida* and histoplasmosis, were more prevalent fungal infections in IBD. Furthermore, aspergillosis infection in IBD patients tracked with age, whereas *Candida* infections were more common in women with IBD. Similar to our study, a prior systematic review by Stamatiades et al. that included 1524 IBD patients found *Candida* infections, followed by histoplasmsosis, to be the most prevalent fungal infections in IBD patients.[18] One of the largest studies examining histoplasmosis in the general population estimated the overall incidence rate to be 3.3/100,000 among individuals older than 65 years between 1998 and 2008, a number lower than our findings in patients with IBD.[19] The inferred higher numbers of IBD patients in this study with histoplasmosis when compared with the overall estimates of histoplasmosis in the United States provided by Baddley and Benedict et al. suggest an increased risk of histoplasmosis in IBD patients.[19, 20] For pneumocystosis, a prior retrospective cohort study by Long et al. showed a higher crude incidence of PCP of 10.6/100,000 in patients with IBD when compared with 3.0/100,000 in those without IBD,[21] a finding very similar to ours. Before our study, there were limited data on cryptococcus and aspergillosis in IBD, mostly in the form of case series or case reports.[22–26] Histoplasmosis, aspergillosis, and candidiasis were more common in the elderly with IBD when compared with adult patients. There have been no prior studies exploring the relationships of these infections with age (Badley et al. was limited to a population >65 years for histoplasmosis), making this an important study examining these relationships. In addition, *Candida* infections were more prevalent in females in both CD and UC.

We found bacterial infections to be more prevalent in patients with IBD when compared with those without IBD. Of these infections, *C. difficile* was the most prevalent opportunistic bacterial disease. We found *C. difficile* to be significantly less prevalent in CD than in UC, which is a similar to a recent large Canadian study by Singh et al. that showed lower mean annual incidence of *C. difficile* in CD (377 per 100,000 person-years) compared with UC (512 per 100,000 person-years follow up).[27] Consistent with our results, a prior study by Mir and Kellermeyer et al. found the prevalence of *C. difficile* infection in pediatric IBD to be 8.1%, significantly higher than in the general population,[28] and increased *C. difficile* was also noted in younger IBD patients (age <30 years) in the Canadian study. Our study also showed that pneumococcal infections are highly prevalent bacterial infections in patients with CD and UC. The risks for pneumococcal infection include age, chronic illness, and chronic immunosuppressive therapy, making patients with IBD at increased risk for the disease. A recent study by Kantsø et al. using a nationwide Danish cohort found that patients with IBD were at an increased risk of invasive pneumococcal disease even 4 years before diagnosis, with a UC hazard ratio of 1.51 (95% CI, 1.05–2.17) and a CD hazard ratio of 1.79 (95% CI, 1.05–3.03).[29] Furthermore, we found pneumococcal prevalence to be higher in elderly patients with CD and UC, which is similar to a US claims database study by Long et al., which showed the highest absolute risk of pneumonia in elderly patients with IBD.[21] Although there have been case reports and reviews of legionella, listeria, and nocardiosis in IBD patients,[30–32] there have been no prior large-scale studies of their prevalence in IBD, making ours the first to describe this relationship.

We found that multiple viral infections were more prevalent in IBD when compared with the general population. Influenza was the most prevalent viral infection in IBD, whereas CMV had the highest risk ratios in IBD. Similar to our results, Tinsley et al., using MarketScan data from January 2008 to December 2011, demonstrated increased risk of influenza and influenza-related hospitalization in IBD patients as compared with patients without IBD.[33] We also found influenza to be more prevalent in children when compared with adults, and this relationship with age has not been previously described in patients with IBD. Of the viral infections studied, CMV was significantly less prevalent in CD compared with UC, which is not dissimilar to a case–control study by McCurdy et al. from the Mayo Clinic that showed CMV to be less frequent in CD than in UC.[34] Furthermore, we demonstrated CMV to be more prevalent in males in both CD and UC. HSV was more prevalent in IBD, a relationship that has been described previously[35]; however, we also noted a higher prevalence of HSV in females when compared with males for both CD and UC. Similarly, EBV was also found to be more prevalent in UC and CD than in non-IBD controls; although there are few data on infection, EBV seroprevalence (EBV DNA in blood) has been noted to be more common in IBD patients than in healthy controls.[36]

With regard to tuberculosis (TB), we found a higher prevalence of pulmonary tuberculosis in CD and UC when compared with non-IBD controls. Furthermore, TB was more common in CD than in UC. TNF-alpha inhibitors are a known risk factor for the activation of latent TB, and as such the American College of Gastroenterology and American Gastroenterological Association recommend screening for latent TB before initiating treatment with these agents.[2, 37] These stringent guidelines have likely led to decreased prevalence of TB in IBD.

Our study very importantly highlights that many vaccine-preventable diseases like influenza, HPV, and pneumococcal infection are still common in IBD patients. Even though effective vaccines are available for each of these infections,[38] there is an unmet need to educate providers and patients regarding timely and appropriate vaccination for these common infections.

There are certain limitations that need to be acknowledged with this study. As we used SNOMED-CT codes for the diagnosis of diseases, not all patients with infections may have been captured, and others may have been misclassified. Validation of the diagnosis of these OIs was not possible, as data are de-identified on curation to the Explorys database and definite diagnostic

information such as histology reports and blood tests are not available in the database. Furthermore, a major drawback of the study is that the medication data in Explorys are incompletely curated and medication usage time with respect to IBD diagnosis is not documented. However, by incorporating a very large number of cases and controls, we mitigate the effects of the above confounders. Another limitation of our study is that we were unable to capture information such as sociodemographic factors and geographical locations of these cases with OI. Finally, a limitation of this database is that Explorys rounds to 10, which can have a significant impact on the diseases with very low prevalence, making the prevalence of diseases such as nocardiosis likely to be less reliable when compared with diseases with higher prevalence.

CONCLUSIONS

Overall, this is the largest study to date that has described the prevalence of OI in CD and UC in the United States. We found significantly higher rates of bacterial, fungal, and viral OI in IBD patients when compared with controls without IBD. Our study suggests the need for close follow-up of IBD patients to diagnose these OIs and suggests the importance of ensuring that IBD patients are up to date with vaccinations such as pneumonia, HPV, and influenza.

SUPPLEMENTARY DATA

Supplementary data are available at *Inflammatory Bowel Diseases* online.

REFERENCES

1. Dave M, Papadakis KA, Faubion WA Jr. Immunology of inflammatory bowel disease and molecular targets for biologics. *Gastroenterol Clin North Am.* 2014;43:405–424.
2. Kornbluth A, Sachar DB; Practice Parameters Committee of the American College of Gastroenterology. Ulcerative colitis practice guidelines in adults: American College of Gastroenterology, Practice Parameters Committee. *Am J Gastroenterol.* 2010;105:501–523; quiz 524.
3. Lichtenstein GR, Loftus EV, Isaacs KL, et al. ACG clinical guideline: management of Crohn's disease in adults. *Am J Gastroenterol.* 2018;113:481–517.
4. Dave M, Purohit T, Razonable R, et al. Opportunistic infections due to inflammatory bowel disease therapy. *Inflamm Bowel Dis.* 2014;20:196–212.
5. Rahier JF, Magro F, Abreu C, et al; European Crohn's and Colitis Organisation (ECCO). Second European evidence-based consensus on the prevention, diagnosis and management of opportunistic infections in inflammatory bowel disease. *J Crohns Colitis.* 2014;8:443–468.
6. Kappelman MD, Moore KR, Allen JK, et al. Recent trends in the prevalence of Crohn's disease and ulcerative colitis in a commercially insured US population. *Dig Dis Sci.* 2013;58:519–525.
7. Shivashankar R, Tremaine WJ, Harmsen WS, et al. Incidence and prevalence of Crohn's disease and ulcerative colitis in Olmsted County, Minnesota from 1970 through 2010. *Clin Gastroenterol Hepatol.* 2017;15:857–863.
8. Explorys team. "We unlock the power of BIG DATA to improve healthcare for everyone." 2015. https://www.explorys.com/about-us.html. Accessed August 25, 2016.
9. Garg A, Hundal J, Strunk A. Overall and subgroup prevalence of Crohn disease among patients with hidradenitis suppurativa: a population-based analysis in the united states. *JAMA Dermatol.* 2018;154:814–818.
10. Hill E, Abboud H, Briggs FBS. Prevalence of asthma in multiple sclerosis: a United States population-based study. *Mult Scler Relat Disord.* 2019;28:69–74.
11. Nadkarni PM, Darer JA. Migrating existing clinical content from ICD-9 to SNOMED. *J Am Med Inform Assoc.* 2010;17:602–607.
12. Medicine UNLo. https://www.nlm.nih.gov/research/umls/mapping_projects/icd9cmv3_to_snomedct.html. Accessed August 25, 2018.
13. Kaelber DC, Foster W, Gilder J, et al. Patient characteristics associated with venous thromboembolic events: a cohort study using pooled electronic health record data. *J Am Med Inform Assoc.* 2012;19:965–972.
14. Chouhan V, Mansoor E, Parasa S, et al. Rates of prevalent colorectal cancer occurrence in persons 75 years of age and older: a population-based national study. *Dig Dis Sci.* 2018;63:1929–1936.
15. Panhwar MS, Mansoor E, Al-Kindi SG, et al. Risk of myocardial infarction in inflammatory bowel disease: a population-based national study. *Inflamm Bowel Dis.* 2019;25:1080–1087.
16. Mansoor E, Saleh MA, Cooper GS. Prevalence of eosinophilic gastroenteritis and colitis in a population-based study, from 2012 to 2017. *Clin Gastroenterol Hepatol.* 2017;15:1733–1741.
17. Dean AG, Sullivan KM, Soe MM. *OpenEpi: Open Source Epidemiologic Statistics for Public Health, Version 3.01*. www.OpenEpi.com. Updated April 6, 2013; Accessed April 20, 2019.
18. Stamatiades GA, Ioannou P, Petrikkos G, et al. Fungal infections in patients with inflammatory bowel disease: a systematic review. *Mycoses.* 2018;61:366–376.
19. Baddley JW, Winthrop KL, Patkar NM, et al. Geographic distribution of endemic fungal infections among older persons, United States. *Emerg Infect Dis.* 2011;17:1664–1669.
20. Benedict K, Mody RK. Epidemiology of histoplasmosis outbreaks, United States, 1938-2013. *Emerg Infect Dis.* 2016;22:370–378.
21. Long MD, Martin C, Sandler RS, et al. Increased risk of pneumonia among patients with inflammatory bowel disease. *Am J Gastroenterol.* 2013;108:240–248.
22. Chen LP, Li J, Huang MF, et al. *Cryptococcus neoformans* infection in ulcerative colitis with immunosuppressants. *Inflamm Bowel Dis.* 2011;17:2023–2024.
23. Murai H, Tokunaga H, Kubo I, et al. Myeloradiculitis caused by *Cryptococcus neoformans* infection in a patient with ulcerative colitis: a neuropathological study. *J Neurol Sci.* 2006;247:236–238.
24. Fraison JB, Guilpain P, Schiffmann A, et al. Pulmonary cryptococcosis in a patient with Crohn's disease treated with prednisone, azathioprine and adalimumab: exposure to chicken manure as a source of contamination. *J Crohns Colitis.* 2013;7:e11–e14.
25. Wysocki JD, Said SM, Papadakis KA. An uncommon cause of abdominal pain and fever in a patient with Crohn's disease. *Gastroenterology.* 2015;148:e12–e13.
26. Alonso-Sierra M, Calvo M, González-Lama Y. Nocardia and aspergillus coinfection in a patient with ulcerative colitis during golimumab therapy. *J Crohns Colitis.* 2016;10:1127–1128.
27. Singh H, Nugent Z, Yu BN, et al. Higher incidence of *Clostridium difficile* infection among individuals with inflammatory bowel disease. *Gastroenterology.* 2017;153:430–438.e2.
28. Mir SA, Kellermayer R. *Clostridium difficile* infection in newly diagnosed pediatric inflammatory bowel disease in the mid-Southern United States. *J Pediatr Gastroenterol Nutr.* 2013;57:487–488.
29. Kantsø B, Simonsen J, Hoffmann S, et al. Inflammatory bowel disease patients are at increased risk of invasive pneumococcal disease: a Nationwide Danish Cohort Study 1977-2013. *Am J Gastroenterol.* 2015;110:1582–1587.
30. Miranda-Bautista J, Padilla-Suárez C, Bouza E, et al. Listeria monocytogenes infection in inflammatory bowel disease patients: case series and review of the literature. *Eur J Gastroenterol Hepatol.* 2014;26:1247–1252.
31. Beigel F, Jürgens M, Filik L, et al. Severe legionella pneumophila pneumonia following infliximab therapy in a patient with Crohn's disease. *Inflamm Bowel Dis.* 2009;15:1240–1244.
32. Abreu C, Rocha-Pereira N, Sarmento A, et al. Nocardia infections among immunomodulated inflammatory bowel disease patients: a review. *World J Gastroenterol.* 2015;21:6491–6498.
33. Tinsley A, Navabi S, Williams ED, et al. Increased risk of influenza and influenza-related complications among 140,480 patients with inflammatory bowel disease. *Inflamm Bowel Dis.* 2019;25:369–376.
34. McCurdy JD, Jones A, Enders FT, et al. A model for identifying cytomegalovirus in patients with inflammatory bowel disease. *Clin Gastroenterol Hepatol.* 2015;13:131–137; quiz e7.
35. Seksik P, Cosnes J, Sokol H, et al. Incidence of benign upper respiratory tract infections, HSV and HPV cutaneous infections in inflammatory bowel disease patients treated with azathioprine. *Aliment Pharmacol Ther.* 2009;29:1106–1113.
36. Magro F, Santos-Antunes J, Albuquerque A, et al. Epstein-Barr virus in inflammatory bowel disease-correlation with different therapeutic regimens. *Inflamm Bowel Dis.* 2013;19:1710–1716.
37. Vaughn BP, Doherty GA, Gautam S, et al. Screening for tuberculosis and hepatitis B prior to the initiation of anti-tumor necrosis therapy. *Inflamm Bowel Dis.* 2012;18:1057–1063.
38. Farraye FA, Melmed GY, Lichtenstein GR, et al. ACG clinical guideline: preventive care in inflammatory bowel disease. *Am J Gastroenterol.* 2017;112:241–258.