Characteristics and outcomes of over a million inflammatory bowel disease subjects in seven countries: a multinational cohort study

Chen Yanover[1], Ramit Magen-Rimon[2], Erica Voss[3,4], Joel Swerdel[3,4], Anna Sheahan[3,4], Nathan Hall[3,4], Jimyung Park[3,5], Rae Woong Park[3,6], Kwang Jae Lee[3,7], Sung Jae Shin[3,7], Seung In Seo[4,8], Kyung-Joo Lee[3,8], Thomas Falconer[3,5], Leonard Haas[3,9], Paul Nagy[3,9], Mary Bowring[3,9], Michael Cook[3,9], Steven Miller[3,9], Tal El-Hay[1], Maytal Bivas-Benita[1], Pinchas Akiva[1], Yehuda Chowers[10], Roni Weisshof[10]

[1]KI Research Institute, Kfar Malal, Israel; [2]Pediatric Gastroenterology & Nutrition institute, Ruth Children's Hospital of Haifa, Rambam Medical Center, Faculty of Medicine, Technion, Haifa, Israel; [3]OHDSI Collaborators, Observational Health Data Sciences and Informatics (OHDSI), New York, NY, USA; [4]Janssen Pharmaceutical Research and Development LLC, Titusville, NJ, USA; [5]Department of Biomedical Informatics, Columbia University Irving Medical Center, New York, New York, USA; [6]Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, Korea; [7]Department of Gastroenterology, Ajou University School of Medicine, Suwon, Korea; [8]Kandong Sacred Heart Hospital; [9]Division of Pediatric Gastroenterology, Hepatology and Nutrition, Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, MD, USA; [10]Department of Gastroenterology, Rambam Health Care Campus and Bruce Rappaport School of Medicine, Technion Israel Institute of Technology, Haifa, Israel

Conflict of interest: EAV, JS, AS, and NH are employees of Janssen Research and Development LLC and shareholders of Johnson & Johnson (J&J) stock; SM consults for Surescripts; YC is employee of Cytoreason.

Background

Crohn's disease (CD) and ulcerative colitis (UC) are chronic inflammatory bowel diseases (IBD) with consistently increasing incidence rates worldwide, especially in newly industrialized and developed countries[1–3]. IBD have significant impact on a patient's and family's quality of life, imposing heavy healthcare financial burden due to chronically administered medications, hospitalizations, and surgical procedures[2].

We describe an open Observational Health Data Sciences and Informatics (OHDSI) network study that characterizes IBD patient cohorts to better understand its natural history, risk factors, symptoms, associated comorbidities (including malignancies and extra-intestinal manifestations), treatment pathways, and outcomes. Specifically, we compute the prevalence and aggregate counts of thousands of attributes, in multiple baseline and follow-up time windows, for IBD, CD, UC, and IBD unclassified (IBDU) patient cohorts as well as various strata (e.g., by age group, sex, race). Results for approximately 1.2 million IBD subjects in sixteen databases (DBs) from seven countries – USA, UK, France, Germany, Japan, Korea, and Australia – are freely available through an interactive website and in downloadable formats and can serve as a resource for further exploration and analysis.

Methods

Study design. A multinational descriptive cohort study using routinely collected healthcare data, standardized to the Observational Medical Outcomes Partnership (OMOP) Common Data Model. All the data partners obtained Institutional Review Board (IRB) or equivalent governance approval.

Data Sources. We report IBD characterization statistics from primary care, outpatient, and inpatient healthcare data in sixteen DBs:

- Nine DBs from the USA:

- Insurance claims data from IBM® MarketScan® Commercial Claims Database (CCAE), IBM® MarketScan® Multi-State Medicaid Database (MDCD), IBM® MarketScan® Medicare Supplemental Database (MDCR), Optum's de-identified Clinformatics® Data Mart Database (CDM), and IQVIA™ Adjudicated Health Plan Claims (PharMetrics Plus)
- Electronic health records (EHRs) from Optum® de-identified Electronic Health Record dataset (Optum-EHR), Columbia University Irving Medical Center (CUIMC), and Johns Hopkins Medicine (JHM)
- Both claims and EHR data from IQVIA Adjudicated Health Plan Claims Data (AMB-EHR)

- Three DBs from Europe:
  - EHRs from IQVIA Disease Analyzer – France (France), IQVIA Disease Analyzer – Germany (Germany), and IQVIA Medical Research Data – UK (IMRD-UK)
- Three DBs from Asia:
  - Insurance claims from Japan in JMDC
  - EHRs from Ajou University School of Medicine (AUSOM) and Kangdong Sacred Heart Hospital (KDH)
- An Australian DB:
  - IQVIA Australian Longitudinal Patient Data (Australia)

Study population. IBD cohorts included individuals with at least two diagnoses of IBD or with an IBD diagnosis and a prescription for an IBD medication (for a list of medications, see https://atlas-demo.ohdsi.org/#/conceptset/1873508); CD and UC cohorts further required at least one diagnosis of the corresponding disease and none of the other. IBDU cohorts included IBD subjects, as defined above, with diagnoses of both CD and UC, or neither. Finally, incident cohorts also required that individuals have a minimum observation of 365 days prior to the first IBD-related diagnosis or prescription (index date). Each target cohort was analyzed in full and stratified on follow-up time (1, 3, 5, and 10 years), sex, specific age groups, race, index date year (in 5-year windows), body mass index (BMI) categories, and pregnancy. The definition of all concept-sets and cohorts are available through the OHDSI demo ATLAS (https://atlas-demo.ohdsi.org/) and the project GitHub repository (https://github.com/ohdsi-studies/IbdCharacterization).

Characteristics and outcomes. We extracted baseline characteristics during subjects' entire history and the year and month before the index date. Outcomes and treatments were identified during the 1 month, 1, 3, 5, 10 years, and all-time following each subject's index date. In addition to OHDSI predefined features (demographics, condition groups, drug era groups), we define a large set of IBD-specific attributes; for a full list see study protocol (https://github.com/ohdsi-studies/IbdCharacterization/blob/master/documents/Protocol IBD Characterisation.docx).

Data analysis. The IBD characterization analysis package was run on each data source independently, sharing only aggregated information, i.e., proportion of subjects having each specific attribute out of the corresponding cohort; attributes with less than five individuals have not been shared, to minimize the risk of re-identification.

Results

Collectively, we characterized the disease trajectory of over one million IBD subjects, including 462,502 CD and 589,118 UC subjects. **Tables 1 and 2** provide information on the size, age and sex distribution, and observation time of individuals in the CD and UC cohorts, respectively, in each database; extended tables are available online (https://data.ohdsi.org/IbdTable1/).

Full characteristics of the IBD cohorts and strata are publicly available (https://data.ohdsi.org/IbdCharacterization/). Here, as an example, we briefly discuss two analyses.

Figure 1 shows the cumulative hospitalization rate of CD and UC subjects across DBs. CD – but not UC – any hospitalization rates in Korea and Japan were much higher than in the US (except for MDCD), e.g., within a year post diagnosis, 42.6% and 48.5% in Korea and Japan, respectively, compared to 20.3–29.9% in the US. Similar trends have been previously reported in Korea[4,5], and may, in part, be attributed to differences in subject characteristics (e.g., significantly larger proportion of males), disease characteristics (poorer response to medical treatment), or treatment policy.

Figures 2 and 3 depict the rate of subjects treated with biologic agents during the first year following diagnosis. The highest rates of biologic treatment were in the pediatric age group (10-18y), ranging from 21.1-43.5% in CD and 7.4-18.4% in UC.

Conclusion

To the best of our knowledge, this is the largest unified characterization study of IBD subjects. The characteristics obtained can be used to globally assess demographic and racial differences in disease behavior as well as treatment trends. Importantly, the current effort will set the stage for related research projects, addressing both predictive (e.g., identifying subjects at increased risk of developing complicated disease) and causal questions (e.g., estimating the effectiveness of biologics).

References

1. Benchimol EI, Manuel DG, Guttmann A, et al. Changing Age Demographics of Inflammatory Bowel Disease in Ontario, Canada: A Population-based Cohort Study of Epidemiology Trends. *Inflamm Bowel Dis*. 2014;20(10):1761-1769. doi:10.1097/MIB.0000000000000103

2. Windsor JW, Kaplan GG. Evolving Epidemiology of IBD. *Curr Gastroenterol Rep*. 2019;21(8):40. doi:10.1007/s11894-019-0705-6

3. Freeman K, Ryan R, Parsons N, Taylor-Phillips S, Willis BH, Clarke A. The incidence and prevalence of inflammatory bowel disease in UK primary care: a retrospective cohort study of the IQVIA Medical Research Database. *BMC Gastroenterol*. 2021;21(1):139. doi:10.1186/s12876-021-01716-6

4. Ye BD, Hong SN, Seo SI, et al. Changes in the Long-term Prognosis of Crohn's Disease between 1986 and 2015: The Population-Based Songpa-Kangdong Inflammatory Bowel Disease Cohort Study. *Gut Liver*. 2022;16(2):216-227. doi:10.5009/gnl210044

5. Cha JM, Park SH, Rhee KH, et al. Long-term prognosis of ulcerative colitis and its temporal changes between 1986 and 2015 in a population-based cohort in the Songpa-Kangdong district of Seoul, Korea. *Gut*. 2020;69(8):1432-1440. doi:10.1136/gutjnl-2019-319699

Table 1. Baseline characteristics of Crohn's disease incident cohorts.

Database	N	History (years)*	Follow-up (years)*	Female	Age at index date¶				
					2-6y	6-10y	10-18y	18-65y	≥65y
AUSTRALIA	210	1.81 [1.35, 2.63]	1.64 [0.55, 2.71]	16.2%‡	—	—	4.8%	81.0%	14.3%
AUSOM	431	8.21 [3.48, 12.57]	4.57 [1.45, 9.68]	35.7%	<1.2%	2.3%	17.6%	72.6%	6.7%
KDH	44	6.56 [3.68, 14.84]	4.79 [0.25, 9.26]	38.6%	—	—	<11.4%	88.6%	<11.4%
JMDC	1,837	3.27 [1.96, 5.39]	2.39 [1.03, 4.39]	27.4%	0.6%	0.9%	13.7%	82.9%	1.9%
AMB-EHR	51,328	2.94 [1.79, 4.66]	4.41 [2.32, 7.02]	61.6%	0.2%	0.5%	4.2%	71.4%	23.8%
CCAE	84,959	2.27 [1.48, 3.84]	2.99 [1.3, 5.95]	55.6%	0.5%	1.4%	9.0%	88.6%	0.5%
MDCD	16,538	2.17 [1.44, 3.69]	3.44 [1.54, 5.88]	65.2%	1.4%	3.1%	13.8%	69.9%	11.5%
MDCR	12,809	2.41 [1.55, 4.04]	3.96 [1.82, 6.84]	58.5%	—	—	—	2.7%	97.3%
CDM	61,376	2.28 [1.48, 3.83]	3 [1.26, 5.99]	56.1%	0.3%	0.9%	5.5%	66.0%	27.2%
Optum-EHR	118,610	4.33 [2.45, 6.72]	4.7 [2.18, 7.15]	59.5%	0.2%	0.7%	5.6%	75.5%	18.0%
Pharmetrics	93,217	2.42 [1.5, 3.46]	2.55 [1.12, 4.65]	53.8%	0.3%	1.0%	7.2%	82.6%	8.9%
CUIMC	4,438	8.15 [3.81, 14.29]	5.37 [2.01, 9.37]	56.6%	0.9%	1.6%	10.2%	63.7%	23.6%
JHM	1128	2.24 [0.95, 3.63]	2.27 [1.53, 3.44]	62.5%	0.8%	1.8%	10.1%	63.3%	23.9%
FRANCE	320	1.71 [1.3, 2.42]	2.11 [1.22, 3.04]	62.5%	—	—	5.6%	86.3%	8.1%
GERMANY	8,321	3.53 [2.07, 5.53]	3.6 [1.58, 5.81]	59.3%	0.4%	0.6%	5.3%	75.8%	17.8%
IMRD-UK	6,936	6.09 [3.18, 10.02]	5.07 [2.42, 8.84]	54.4%	0.2%	1.3%	10.9%	73.7%	13.9%

*History and follow-up median; interquartile range is shown in brackets. ¶Age at first IBD-related diagnosis or prescription; †Age groups correspond to very early onset (2-6 years), early onset (6-10 years), pediatrics (10-18 years), adults (18-65 years) and elderly (≥65 years). ‡In the Australian database, 59% of subjects have no designated sex; 13.3% of CD subjects identified as males.

***Table 2**. Baseline characteristics of ulcerative colitis incident cohorts.*

Database	N	History (years)	Follow-up (years)	Female	Age at index date				
					2-6y	6-10y	10-18y	18-65y	≥65y
AUSTRALIA	210	1.85 [1.34, 2.75]	1.49 [0.64, 2.65]	15.7%‡	—	<2.4%	<2.4%	76.7%	22.4%
AUSOM	660	6.45 [3, 11.64]	5.14 [1.79, 10.45]	36.4%	—	<0.8%	5.5%	85.8%	8.5%
KDH	93	7.28 [2.96, 11.26]	4.45 [1.84, 11.04]	39.8%	<5.4%	—	<5.4%	79.6%	16.1%
JMDC	12,527	3.28 [1.99, 5.29]	2.38 [1.04, 4.25]	35.6%	0.2%	0.2%	4.3%	92.8%	2.4%
AMB-EHR	58,404	2.96 [1.81, 4.66]	4.4 [2.34, 7.03]	58.6%	0.3%	0.4%	2.2%	67.3%	29.8%
CCAE	109,500	2.28 [1.48, 3.86]	3.08 [1.34, 6.12]	55.2%	0.3%	0.5%	3.8%	94.8%	0.6%
MDCD	16,431	2.32 [1.5, 3.94]	3.34 [1.43, 5.87]	65.9%	0.7%	1.6%	6.8%	69.0%	21.8%
MDCR	26,242	2.52 [1.6, 4.18]	3.94 [1.73, 6.89]	58.7%	—	—	—	1.9%	98.1%
CDM	93,009	2.35 [1.51, 3.98]	3.08 [1.28, 6.01]	56.9%	0.2%	0.3%	2.0%	59.3%	38.2%
Optum-EHR	119,188	4.35 [2.49, 6.76]	4.76 [2.25, 7.22]	58.2%	0.1%	0.3%	2.4%	71.9%	25.3%
Pharmetrics	123,149	2.44 [1.52, 3.61]	2.62 [1.16, 4.65]	53.7%	0.2%	0.4%	3.0%	82.9%	13.5%
CUIMC	3,922	7.67 [3.54, 14.27]	4.97 [1.92, 8.91]	59.0%	0.2%	0.7%	3.3%	65.4%	30.3%
JHM	892	2 [0.76, 3.56]	2.3 [1.59, 3.46]	57.5%	1.1%	0.8%	6.7%	60.4%	30.8%
FRANCE	299	1.72 [1.31, 2.4]	2.03 [0.87, 2.95]	51.5%	<1.7%	—	2.7%	79.3%	17.7%
GERMANY	10,668	3.55 [2.04, 5.59]	3.51 [1.4, 5.79]	53.2%	0.3%	0.5%	2.8%	68.1%	28.1%
IMRD-UK	13,924	5.68 [2.96, 9.49]	5.55 [2.63, 9.46]	49.1%	0.2%	0.5%	3.7%	72.8%	22.9%

‡18.1% of the Australian cohort's subjects identified as males. See **Error! Reference source not found.** for m ore information.

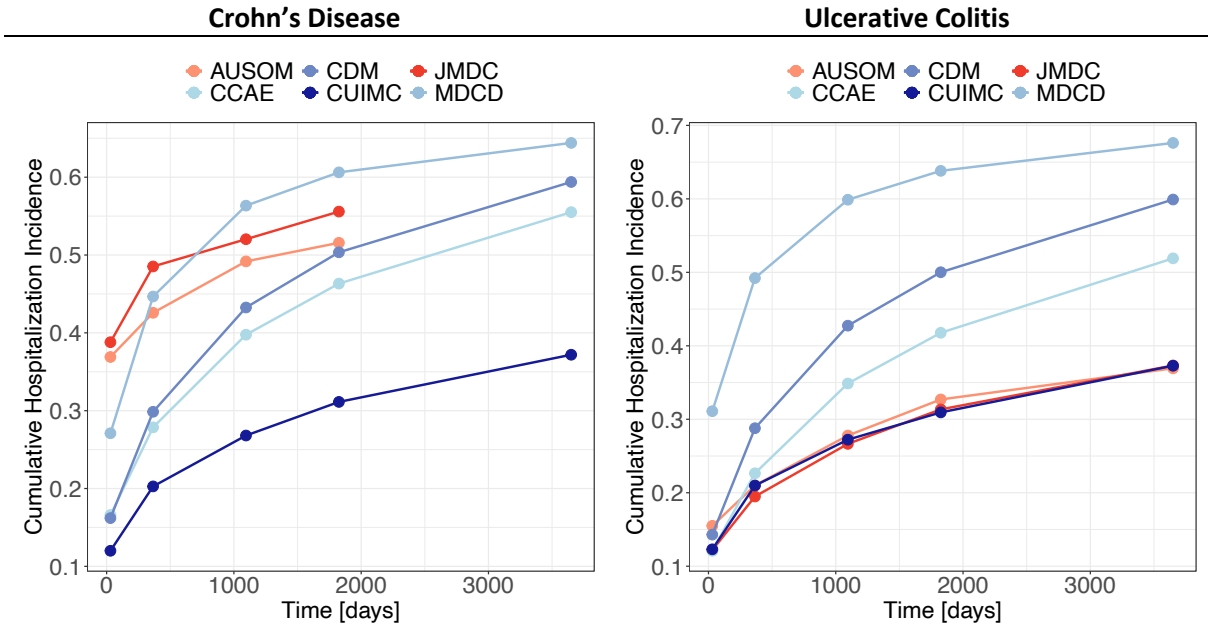

***Figure 1**. Cumulative hospitalization rate, following diagnosis with Crohn's disease (left) and ulcerative colitis (right). Each line represents a different data source, USA DBs are shown in blue shades, Asian ones – in reds.*

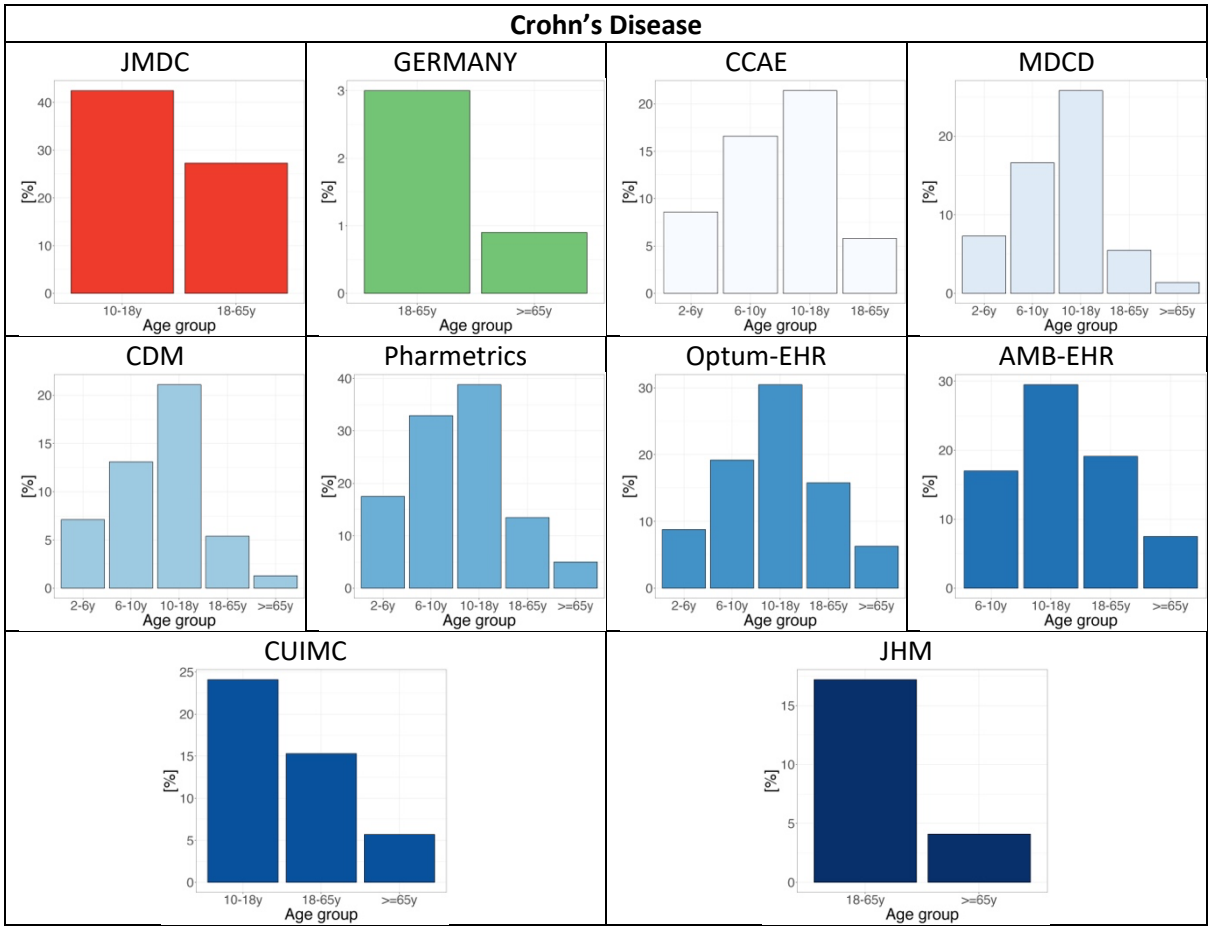

***Figure 2**. Rate of treatment with biologic agents during the first year following index date, by age group, in CD subjects.*

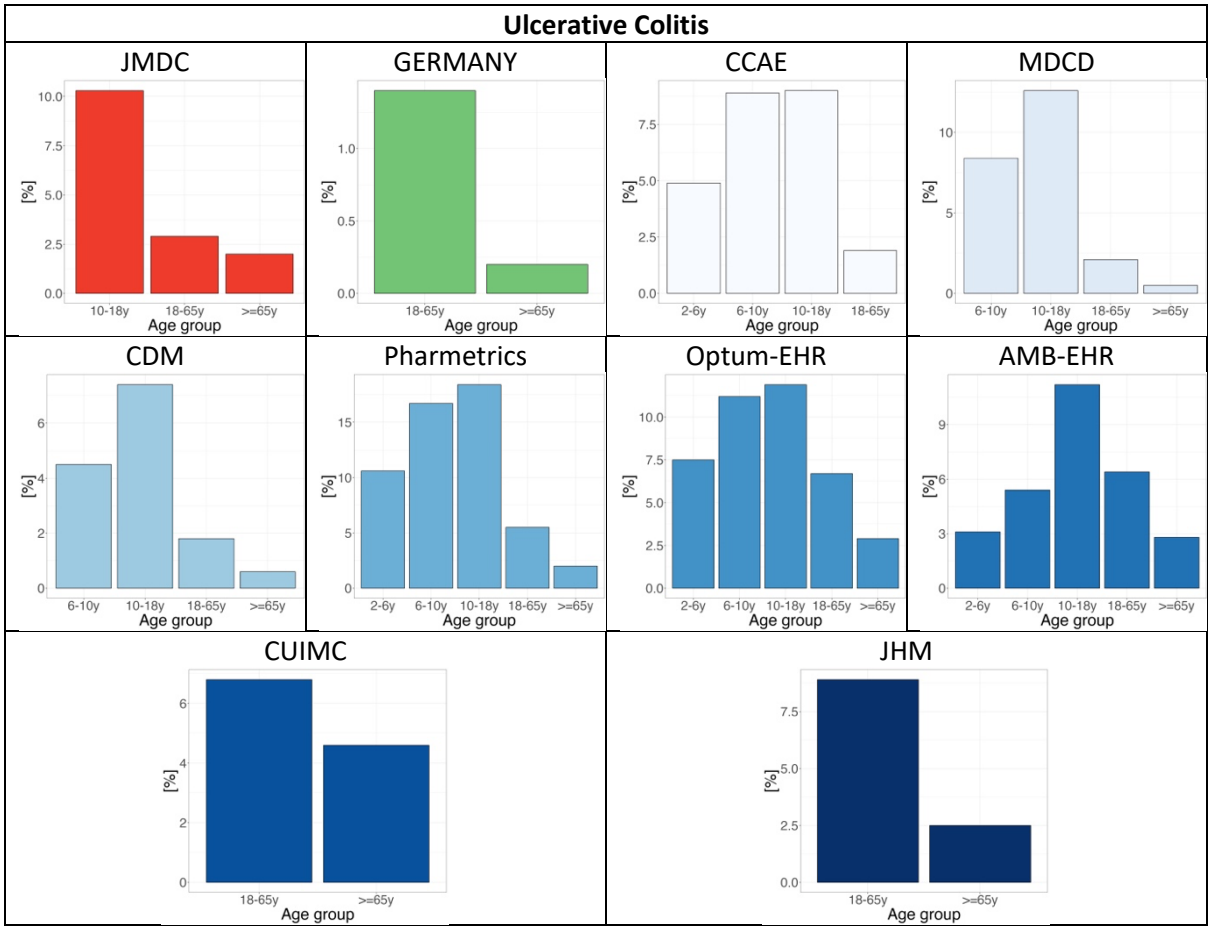

Figure 3. *Rate of treatment with biologic agents during the first year following index date, by age group, in UC subjects.*