Emulation of a randomized controlled trial in ulcerative colitis with US and French claims data: Infliximab with thiopurines compared to infliximab monotherapy

Julien Kirchgesner, Rishi Desai, Maria Schneeweiss, Laurent Beaugerie, Seoyoung Kim, Sebastian Schneeweiss

Emulation of a randomized controlled trial in ulcerative colitis with US and French claims data: Infliximab with thiopurines compared to infliximab monotherapy

Short title: Real-world evidence in ulcerative colitis

Julien Kirchgesner, MD, PhD [1,2], Rishi J. Desai, PhD [1], Maria C. Schneeweiss, MD [1] , Laurent Beaugerie, MD, PhD [2], Seoyoung C. Kim, MD, ScD [1,3], Sebastian Schneeweiss, MD, ScD [1]

(1) Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA

(2) Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, AP-HP, Hôpital Saint-Antoine, Department of gastroenterology, Paris, France

(3) Division of Rheumatology, Inflammation, and Immunity, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA

Corresponding author:
Julien Kirchgesner
Division of Pharmacoepidemiology and Pharmacoeconomics,
Department of Medicine, Brigham and Women's Hospital and Harvard Medical School
1620 Tremont Street, Suite 3030, Boston, MA 02120
Phone: 617-278-0930
Fax: 617-232-8602
E-mail: jkirchgesner@bwh.harvard.edu

Body text word count: 2935 words

Key Words: Real world evidence; RCT emulation; Ulcerative colitis, anti-TNF.

Key points:

- Real world evidence (RWE) is increasingly used to assess treatment effectiveness in clinical practice; calibrating RWE findings against randomized controlled trials (RCTs) may support the use of RWE in select circumstances.
- Using two U.S. commercial insurance databases (IBM MarketScan and Optum) and the French nationwide health insurance database (SNDS), we replicated the SUCCESS trial. SUCCESS found a greater effectiveness of the combination of infliximab with thiopurines compared to infliximab monotherapy in patients with ulcerative colitis.
- The primary outcome was a composite endpoint of treatment failure based on hospitalization related to UC or colectomy, treatment switch to another biologic or immunosuppressant, or corticosteroid exposure at week 16.
- Our comparison qualitatively replicated the effectiveness of combination therapy with infliximab and thiopurines observed in SUCCESS. RWE results were consistent across the three databases.
- These findings highlight the opportunity to use healthcare databases to assess treatment effectiveness in UC.

Statement about prior postings and presentations: The findings of this study have not been presented or published previously.

Abstract

Purpose: To understand the validity of real-world evidence (RWE) studies in ulcerative colitis (UC), we emulated the SUCCESS randomized controlled trial (RCT) on the effectiveness of infliximab plus thiopurines, using U.S. and French healthcare insurance claims data.

Methods: The SUCCESS trial showed improved remission with infliximab plus thiopurines combined compared to infliximab monotherapy in patients with UC. Based on two U.S. commercial claims databases (IBM MarketScan and Optum) and the French nationwide health insurance database (SNDS) from 2004 through 2019, all patients with UC who initiated combination therapy or infliximab alone were identified. The primary outcome of treatment failure was emulated by: hospitalization related to UC or colectomy, treatment switch to another biologic or immunosuppressant, or use of corticosteroids 16 weeks after infliximab initiation. We estimated risk ratios (RRs) with 95% confidence intervals (CIs) after 1:1 propensity score (PS) matching.

Results: Among 620 PS-matched pairs of combination therapy and infliximab monotherapy users, treatment failure occurred in 124 (20%) of patients initiating combination therapy and 170 (27%) during monotherapy. Like in SUCCESS, the risk of treatment failure was decreased with combination therapy in the overall cohort (RR=0.73; 95%CI: 0.60–0.90). Findings were consistent across MarketScan, Optum, and SNDS: RR=0.76 (0.57–1.02), 0.82 (0.54–1.24), and 0.61 (0.41–0.90). Similar results were observed for each component endpoint.

Conclusions: RWE results across three large claims databases were consistent with RCT findings. These findings provide support for the use of RWE to assess treatment effectiveness in UC.

Introduction

Real world evidence (RWE) has been mainly used to inform decisions about treatment safety, while it has been less used to inform decisions about treatment effectiveness and complement evidence provided by randomized controlled trials (RCTs).[1–4] This might be related to the challenge of capturing clinically meaningful measures of effectiveness outcomes in electronic health records or medical claims data.[5] The emulation of RCTs by RWE studies may help to calibrate effectiveness outcomes available in real world data (RWD), which could be applied to other treatment comparisons within the same indication. Additionally, the use of multiple or multinational healthcare databases to calibrate RWE against RCTs may help to interpret differences between treatment effect estimates from the two study types, that can be driven by residual bias in RWE but also emulation differences.[6,7] In several immune-mediated inflammatory diseases, notably ulcerative colitis (UC), recent phase III RCTs compared the active treatment with placebo, and the optimal position of each drug remains largely unknown.[8] Using RWE in comparative effectiveness research could help fill evidence gaps left by RCTs.

Infliximab is a tumor necrosis factor antagonist (anti-TNF) approved for the treatment of UC in 2006 in the US and Europe,[9] while thiopurines have been used for the treatment of UC since the nineteen seventies.[10] The SUCCESS trial was an RCT conducted between 2007 and 2010 assessing the effectiveness of infliximab and thiopurines combined compared to either infliximab or thiopurines alone in patients with UC naïve to anti-TNF.[11] It concluded that the combination therapy of infliximab and thiopurines was superior to treatment with its component alone.

The aim of this study was to calibrate RWE against RCTs in UC by emulating the SUCCESS trial, studying the effectiveness of a biologic agent in patients with UC using claims data from the U.S. and France.

Methods

Data source

We conducted a cohort study using two U.S. health care claims databases, IBM MarketScan (MarketScan) 2004-2018 and Optum's Clinformatics® Data Mart Database (Optum) 2005-2019, and the French administrative health database 2009-2018 (*Système National des Données de Santé*, SNDS). Patients enrolled in the MarketScan and Optum databases are representative of a commercially insured population in the U.S.; the SNDS covers 95% of the French population. Comprehensive data on demographics and procedures performed during outpatient visits or inpatient stays, and outpatient filled prescription records are available in these three databases and can be tracked longitudinally. Additionally, the two U.S databases contain diagnoses during outpatient visits and the SNDS contains the patient's status with respect to full reimbursement of care for long-term diseases (LTDs), which includes UC and allows to assess the date of UC diagnosis.[12] The study was approved by the institutional review board of the Brigham and Women's Hospital, and the French Data Protection Authority. Patient informed consent was not required because the databases were deidentified.

Study population

Inclusion and exclusion criteria were tailored to replicate SUCCESS whenever possible.[11] The SUCCESS trial included adult patients with moderate to severe active UC with a Mayo score of 6 to 12 points, despite use of corticosteroids and/or aminosalicylates. Patients were naïve to anti-TNF and should be either naïve to thiopurines or free from thiopurines for at least 3 months prior to the randomization. Finally, 90% of patients included in SUCCESS were naïve to thiopurines. To emulate this design, we identified all patients aged 18 years or older initiating infliximab after at least 180 days of continuous enrollment for patients included in MarketScan and Optum databases, and with at least one visit for UC using the International Classification of Diseases 9th (ICD-9) or 10th (ICD-10). In the SNDS database, UC diagnosis was based on previous published algorithms,[12–15] and the date of UC diagnosis was defined as the earliest

diagnosis date either from hospital discharge diagnosis or from LTD diagnosis. Cohort entry date was defined as the date of infliximab initiation and patients were followed from the day after cohort entry date to 112 days after cohort entry date (week 16).

In analogy to the SUCCESS trial, we excluded patients previously exposed to any anti-TNF agent any time prior infliximab initiation. We additionally excluded patients exposed to methotrexate, since infliximab can be combined with methotrexate instead of thiopurines as a combination therapy.[8] Thiopurines exposure was only allowed in the month prior infliximab initiation, and patients exposed to thiopurines more than one month before infliximab initiation were excluded. Previous treatment exposure was assessed within all data available in the three databases to minimize the risk of treatment misclassification. Patients with tuberculosis, opportunistic infections, hepatitis B and C within the previous 6 months were excluded. We also excluded patients with systemic lupus erythematosus, heart failure, organ transplant, HIV infection, multiple sclerosis, previous colectomy, ostomy, and previous history of cancer (excluding non-melanoma skin cancer). Since patients hospitalized for extensive severe UC were excluded in SUCCESS, patients hospitalized at infliximab initiation or with a diagnosis of stricture or abscess within the previous 6 months were excluded. The only exclusion criteria included in the SUCCESS trial and not considered in our study was colonic dysplasia, since it is not collected in administrative health databases. Lastly, additional exclusion criteria were included compared to the SUCCESS trial. First, in order to only include patients among whom infliximab was prescribed for UC, patients with a diagnosis code related to Crohn's disease, or diagnosed with rheumatoid arthritis, ankylosing spondylitis, psoriatic arthritis, and psoriasis were excluded. Second, patients previously exposed to biologics and immunosuppressants that were available after the publication of the SUCCESS trial in 2014 (natalizumab, vedolizumab, ustekinumab, and tofacitinib) were excluded.

Treatment groups

Combination therapy was defined as starting thiopurines (either azathioprine or 6-mercaptopurine) within 30 days before infliximab initiation, as thiopurines may be started

before the first infliximab infusion and take up to 3 months to be effective as monotherapy.[8] Infliximab monotherapy was defined as infliximab initiation without thiopurines exposure any time prior. We did not consider patients only exposed to thiopurines monotherapy, since recent guidelines recommend against its use as monotherapy for induction of remission in patients with UC.[8]

Effectiveness measure

In the SUCCESS trial, the primary end point was corticosteroid-free clinical remission at week 16, defined by a total Mayo score of 2 points or less, with no individual subscore exceeding 1 point and without any systemic corticosteroid use. Since the Mayo score is not available in any of the study databases, we developed an effectiveness outcome measure based on three surrogate endpoints: (1) hospitalization related to UC or colectomy; (2) treatment switch to another biologic (i.e., adalimumab, certolizumab pegol, golimumab, natalizumab, vedolizumab, ustekinumab) or immunosuppressant (tofacitinib); (3) exposure to systemic corticosteroids at week 16. The primary outcome, treatment failure, was defined as the occurrence of one of these three endpoints. Related codes are summarized in Supplementary Table 1.

Follow-up started the day after cohort entry and ended on the earliest occurrence of 112 days of follow-up, outcome occurrence, or death, for all subjects started on either one of the treatments.

Patient characteristics

Pre-treatment patient characteristics and markers of UC severity were assessed, including demographics and comorbidities. Comorbidities were assessed in the 180 days before cohort entry in MarketScan and Optum and in all data available in the SNDS, including previous serious infections, *Clostridioides difficile* infection, cardiovascular disease, chronic kidney failure, chronic liver disease, chronic pulmonary disease, venous thromboembolism, and diabetes. UC severity and healthcare use were assessed in the 180 days before cohort entry

in the three databases. UC severity was assessed by corticosteroids and aminosalicylates exposure, occurrence of UC-related hospitalization, abdominal imaging, gastrointestinal endoscopy, fecal pathogen, and the number of C-reactive protein tests ordered. Healthcare use was assessed by hospitalizations not related to UC and the number of gastroenterologist visits.

Statistical analysis

Multivariable logistic regression models estimated patients' propensity score (PS) values, that is, the predicted probability of combination therapy with infliximab and thiopurines versus infliximab monotherapy conditioning on all the confounders listed above and the year of cohort entry.[16] We used PS-matching with a fixed ratio of 1:1 comparing combination therapy and infliximab monotherapy with a matching 'caliper' of 0.02 on the PS scale.[17] After matching, standardized differences were calculated to assess balance between patients treated with combination therapy and their matched controls treated with infliximab monotherapy.[18] Log–binomial regression models were used to estimate adjusted risk ratios (RR) with their 95% confidence intervals (CIs) comparing the risk of treatment failure associated with combination therapy versus infliximab monotherapy.[19] Cohort-specific RRs were combined by an inverse variance-weighted, fixed-effects model.

Additionally, the risk of each individual component of the composite endpoint was assessed in secondary analyses, and subgroup analysis was stratified by exposure to corticosteroids at infliximab initiation. Finally, we performed sensitivity analyses to test the robustness of our results: (1) variables with an absolute standardized difference greater than 0.1 after PS-matching were included as covariate in log–binomial regression models; (2) patients initiating combination therapy with infliximab and thiopurines were PS matched to patients initiating infliximab monotherapy with a variable ratio up to 1:4; (3) the required enrollment period was extent to 365 days in the US databases; (4) treatment switch was excluded from the outcome definition; (5) since some patients may start thiopurines after infliximab initiation, patients starting thiopurines in the first month after infliximab initiation were excluded.

Analyses were performed using the validated Aetion Evidence Platform (V 4.10) using R[20–22] and SAS (version 9.4) statistical software (SAS Institute).

Results

Patient characteristics

A total of 22 051 patients with UC who initiated infliximab were identified across the three databases. After applying exclusion criteria (Figure 1), 7850 immunosuppressants and anti-TNF naive patients with UC were included, initiating either combination therapy with infliximab and thiopurines (n=644) or infliximab monotherapy (n=7206). Overall, 96.4% of patients treated with combination therapy were matched to patients with infliximab monotherapy and 93.0% of patients treated with infliximab monotherapy were able to match with at least one patient treated with combination therapy.

In the matched population of 620 pairs, mean age at cohort entry was between 39 and 40 years, which was similar to SUCCESS (Table 1).[11] Exposure to corticosteroids at infliximab initiation was higher compared to SUCCESS (69.7% compared to 40.5% in SUCCESS). Four covariates had an absolute standardized difference greater than 0.1 (Supplementary Figure 1). Baseline characteristics before PS-matching are provided in Supplementary Table 2, and year of cohort entry before and after PS-matching are provided in Supplementary Table 3.

Combination therapy versus infliximab monotherapy

In the PS matched cohort, 124 (20.0%; 95% CI 17.0-23.3) and 170 (27.4%; 24.1-31.1) patients were in treatment failure at week 16 after initiating combination therapy and infliximab monotherapy, respectively. Cohorts-specific rates of treatment failure ranged from 17.1% to 21.2% in patients initiating combination therapy and from 25.8% to 28.2% in patients initiating infliximab monotherapy. (Table 2) Exposure to corticosteroids at week 16 was the most frequent outcome, accounting for 54% and 50% of outcomes in patients initiating combination therapy and infliximab monotherapy. Treatment switch only accounted for 15% and 12% of outcomes in patients initiating combination therapy and infliximab monotherapy. The proportion of patients with treatment failure before PS-matching is provided in Supplementary Table 4.

Patients initiating combination therapy had a 27% decreased risk of treatment failure compared to patients initiating infliximab monotherapy (RR, 0.73, 95% CI 0.60-0.90). A similar trend was observed across the three databases: RR, 0.76 (95% CI 0.57–1.02) in MarketScan, RR, 0.82 (95% CI 0.54–1.24) in Optum and RR, 0.61 (95% CI 0.41–0.90) in SNDS. RRs of treatment failure in the emulated cohort and in the SUCCESS trial are provided in Figure 2. The reduced risk of treatment failure with combination therapy was observed for two secondary endpoints, with a 40% (12-58%) reduced risk of hospitalization related to UC or colectomy and a 26% (4-43%) reduced risk of persistent corticosteroids exposure at week 16 (Figure 3). A similar trend was observed for treatment switch, although not statistically significant (RR, 0.85, 95% CI, 0.49–1.47).

Subgroup and sensitivity analyses

Results were similar in subgroups of patients according to corticosteroids exposure at infliximab initiation (RR, 0.74 [95%CI, 0.59-0.93] in patients treated with corticosteroids and 0.75 |95%CI, 0.46-1.22] in patients not treated with corticosteroids) (Supplementary Table 5). Sensitivity analyses revealed similar results (Supplementary Table6). After PS matching with a 1:4 variable ratio, the risk of treatment failure was decreased by 21% (3-36%) with combination therapy compared to infliximab monotherapy.

Discussion

This multi-database RWE cohort study emulating, the SUCCESS trial resulted in consistent findings by assessing effectiveness of combination therapy with infliximab and thiopurines compared to infliximab monotherapy in patients with UC. The real-world effectiveness outcome measure was clinically relevant and based on a composite of treatment failure including hospitalization related to UC or colectomy, treatment switch, or corticosteroid use by week 16. Findings were robust across all three claims databases and multiple sensitivity analyses.

Calibrating clinically meaningful real-world effectiveness outcome measures is of major importance for clinical and regulatory decision making. In UC, all recent phase III RCTs compared the active treatment with placebo in immunosuppressant naïve patients, whereas standard first line treatment is available and the European Medicines Agency recommended a direct comparison with current generally accepted standard first line treatment in this setting.[23] Once a valid real-world effectiveness outcome is agreed upon, real-world evidence may complement existing placebo-controlled trials by providing effectiveness estimates in head-to-head comparisons in addition to the usual safety studies.[15] This would also expand the evidence base to understanding treatment effectiveness in clinical practice and include patients who were excluded from RCTs.[24] Starting from a large source population of patients initiating infliximab, a total of only 644 patients naïve to anti-TNFs and immunosuppressants and initiating infliximab and thiopurines were included. This small study size is a consequence of the restrictive exclusion criteria imposed by emulating SUCCESS, since infliximab is mainly used in acute severe UC during hospitalization and other immunosuppressants and biologics such as adalimumab may be initiated before infliximab in patients with moderate to severe UC.[8] While our study design was tailored to emulate the design of the SUCCESS trial as much as possible, this highlights the opportunities of RWE to complement and expand beyond the existing randomized trial evidence base.[7] Clearly, more RCT evidence will help everybody but in its absence RWE will be able to expand our understanding of treatments in UC.

To our knowledge, this is the first study replicating findings of RCTs in UC based on both U.S and French administrative health databases. Performing the same analysis in RWE databases from different countries allowed us to assess whether different healthcare schemes and potential prescribing patterns may impact the treatment effect observed. While healthcare is guaranteed for all French residents and patients are followed from birth to emigration or death in the French database, the U.S databases only included commercially insured patients and health insurance enrollment changes may reduce the enrollment period to assess covariates. We observed similar results in the stratified analysis according to each database, which suggest that these differences had a minimal impact on the treatment estimate observed.

Although our design was as close as possible to that of the SUCCESS trial, we observed differences in terms of co-treatment exposure and disease severity. Higher rates of corticosteroids exposure reported in our study population may suggest that patients initiating infliximab in real world settings have a more severe disease compared to patients included in the RCTs. In the SUCCESS trial, the exclusion of patients hospitalized for severe extensive UC was based on the investigator judgment that the patient was likely to require colectomy within 12 weeks. Conversely, in real world settings infliximab is more frequently used compared to other biologics in patients experiencing acute severe UC.[8] However, our subgroup analysis stratified according to baseline exposure to corticosteroid provided consistent results. Additionally, as administrative health databases do not provide data on clinical or endoscopic disease activity, we were unable to use the same outcome definition as used in the SUCCESS trial. We used a composite endpoint of treatment failure based on hospitalization related to UC or colectomy, treatment switch to another biologic or immunosuppressants, or corticosteroids continuation at week 16. Our definition may select higher degree of disease activity, as we reported a lower rate of treatment failure compared to SUCCESS. It highlights the fact that the design of an emulated trial based on administrative health databases cannot perfectly emulate the design of a highly-controlled RCT that includes effectiveness outcomes not collected in

administrative health databases. Thus, differences in design between RCTs and their emulations in RWD should be comprehensively reported and considered in the interpretation.

Treatment switch to another biologic or immunosuppressant was not statistically significantly reduced in patients treated with combination therapy compared to infliximab monotherapy. Treatment switch may also include switch related to intolerance or safety. However, point estimates were in the same range of those of the two other secondary outcomes, and treatment switch only accounted for around 15% of the treatment failure reported. Lastly, since endoscopic and histological data were not collected in the three databases, residual confounding by these parameters cannot be entirely ruled out. We adjusted for the number of C-reactive protein tests ordered, while we did not adjust for the number of fecal calprotectin test ordered. Indeed, fecal calprotectin was not used in the early part of our inclusion period and fecal calprotectin is not reimbursed in France.

In conclusion, this study based on three large population-based claims databases of patients with UC in both the US and France provides strong support that clinically-relevant real-world effectiveness outcome measures can be calibrated by comparing RCTs and their emulations in RWD. These measures can be further used to assess effectiveness of other treatments in patients with similar diseases.

Author Contributions:

A list of each author's contributions. The standard contributions include: J.K. and S.S. wrote manuscript, J.K., R.J.D., M.S., L.B., S.C.K., and S.S. designed research, J.K. and S.S. performed research, JK. analyzed data.

Figure Legends:

Figure 1. Study Population Flowchart

Figure 2. Risk ratios for treatment failure associated with combination therapy with infliximab and thiopurines compared to infliximab monotherapy

Figure 3. Risk ratios for each individual component of the composite outcome associated with combination therapy with infliximab and thiopurines compared to infliximab monotherapy (hospitalization related to UC or colectomy [A], treatment switch [B], and corticosteroids continuation [C])

Table Legends:

Table 1. Characteristics of study patients, propensity score-matched with a 1:1 fixed ratio

Table 2. Outcome risk after 1:1 propensity-score matching

References

1. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol* 2005; **58**: 323–337. doi:10.1016/j.jclinepi.2004.10.012.

2. Makady A, Ham R ten, de Boer A, Hillege H, Klungel O, Goettsch W. Policies for Use of Real-World Data in Health Technology Assessment (HTA): A Comparative Study of Six HTA Agencies. *Value Health* 2017; **20**: 520–532. doi:10.1016/j.jval.2016.12.003.

3. Franklin JM, Glynn RJ, Martin D, Schneeweiss S. Evaluating the Use of Nonrandomized Real-World Data Analyses for Regulatory Decision Making. *Clin Pharmacol Ther* 2019; **105**: 867–877. doi:10.1002/cpt.1351.

4. Schneeweiss S. Real-World Evidence of Treatment Effects: The Useful and the Misleading. *Clin Pharmacol Ther* 2019; **106**: 43–44. doi:https://doi.org/10.1002/cpt.1405.

5. US Food and Drug Administration. Framework for FDA's real- world evidence program. 2018. Available at: https://www.fda.gov/media/120060/download. Accessed October 19, 2020.

6. Franklin JM, Glynn RJ, Suissa S, Schneeweiss S. Emulation Differences vs. Biases When Calibrating Real-World Evidence Findings Against Randomized Controlled Trials. *Clin Pharmacol Ther* 2020. doi:10.1002/cpt.1793.

7. Franklin JM, Patorno E, Desai RJ, *et al.* Emulating Randomized Clinical Trials With Nonrandomized Real-World Evidence Studies: First Results From the RCT DUPLICATE Initiative. *Circulation* 2021; **143**: 1002–1013. doi:10.1161/CIRCULATIONAHA.120.051718.

8. Rubin DT, Ananthakrishnan AN, Siegel CA, Sauer BG, Long MD. ACG Clinical Guideline: Ulcerative Colitis in Adults. *Am J Gastroenterol* 2019; **114**: 384–413. doi:10.14309/ajg.0000000000000152.

9. Rutgeerts P, Sandborn WJ, Feagan BG, *et al.* Infliximab for induction and maintenance therapy for ulcerative colitis. *N Engl J Med* 2005; **353**: 2462–2476.

10. Jewell DP, Truelove SC. Azathioprine in Ulcerative Colitis: Final Report on Controlled Therapeutic Trial. *Br Med J* 1974; **4**: 627–630.

11. Panaccione R, Ghosh S, Middleton S, *et al.* Combination Therapy With Infliximab and Azathioprine Is Superior to Monotherapy With Either Agent in Ulcerative Colitis. *Gastroenterology* 2014; **146**: 392-400.e3. doi:10.1053/j.gastro.2013.10.052.

12. Kirchgesner J, Lemaitre M, Rudnichi A, *et al.* Therapeutic management of inflammatory bowel disease in real-life practice in the current era of anti-TNF agents: analysis of the French administrative health databases 2009-2014. *Aliment Pharmacol Ther* 2017; **45**: 37–49. doi:10.1111/apt.13835.

13. Lemaitre M, Kirchgesner J, Rudnichi A, *et al.* Association Between Use of Thiopurines or Tumor Necrosis Factor Antagonists Alone or in Combination and Risk of Lymphoma in Patients With Inflammatory Bowel Disease. *JAMA* 2017; **318**: 1679–1686. doi:10.1001/jama.2017.16071.

14. Kirchgesner J, Nyboe Andersen N, Carrat F, Jess T, Beaugerie L, BERENICE study group. Risk of acute arterial events associated with treatment of inflammatory bowel diseases: nationwide French cohort study. *Gut* 2020; **69**: 852–858. doi:10.1136/gutjnl-2019-318932.

15. Kirchgesner J, Desai RJ, Beaugerie L, Schneeweiss S, Kim SC. Risk of serious infections with vedolizumab versus tumor necrosis factor antagonists in patients with inflammatory bowel disease. *Clin Gastroenterol Hepatol* 2020. doi:10.1016/j.cgh.2020.12.030.

16. Glynn RJ, Schneeweiss S, Stürmer T. Indications for propensity scores and review of their use in pharmacoepidemiology. *Basic Clin Pharmacol Toxicol* 2006; **98**: 253–259. doi:10.1111/j.1742-7843.2006.pto_293.x.

17. Rassen JA, Shelat AA, Myers J, Glynn RJ, Rothman KJ, Schneeweiss S. One-to-many propensity score matching in cohort studies. *Pharmacoepidemiol Drug Saf* 2012; **21**: 69–80. doi:10.1002/pds.3263.

18. Franklin JM, Rassen JA, Ackermann D, Bartels DB, Schneeweiss S. Metrics for covariate balance in cohort studies of causal effects. *Stat Med* 2014; **33**: 1685–1699. doi:10.1002/sim.6058.

19. Robbins AS, Chao SY, Fonseca VP. What's the relative risk? A method to directly estimate risk ratios in cohort studies of common outcomes. *Ann Epidemiol* 2002; **12**: 452–454. doi:10.1016/s1047-2797(01)00278-2.

20. Aetion Evidence Platform® (2020). Software for real-world data analysis. Aetion, Inc. *Aetion*. Available at: https://aetion.com/. Accessed January 7, 2021.

21. Wang SV, Verpillat P, Rassen JA, Patrick A, Garry EM, Bartels DB. Transparency and Reproducibility of Observational Cohort Studies Using Large Healthcare Databases. *Clin Pharmacol Ther* 2016; **99**: 325–332. doi:10.1002/cpt.329.

22. Patorno E, Schneeweiss S, Gopalakrishnan C, Martin D, Franklin JM. Using Real-World Data to Predict Findings of an Ongoing Phase IV Cardiovascular Outcome Trial: Cardiovascular Safety of Linagliptin Versus Glimepiride. *Diabetes Care* 2019; **42**: 2204–2210. doi:10.2337/dc19-0069.

23. EMA. Guideline on the development of new medicinal products for the treatment of Ulcerative Colitis. *EMA* 2016.

24. Ha C, Ullman TA, Siegel CA, Kornbluth A. Patients Enrolled in Randomized Controlled Trials Do Not Represent the Inflammatory Bowel Disease Patient Population. *Clin Gastroenterol Hepatol* 2012; **10**: 1002–1007. doi:10.1016/j.cgh.2012.02.004.

Table 1. Characteristics of study patients treated with infliximab and thiopurines or infliximab monotherapy, propensity score-matched with a 1:1 fixed ratio

	MarketScan		Optum		SNDS		SUCCESS Trial	
Variable	Infliximab Mono. (n=288)	Infliximab Combo. (n=288)	Infliximab Mono. (n=151)	Infliximab Combo. (n=151)	Infliximab Mono. (n=181)	Infliximab Combo. (n=181)	Infliximab Mono. (n=78)	Infliximab Combo. (n=80)
Age, mean (SD), year	39.8 (14.0)	40.4 (14.0)	40.6 (15.4)	40.6 (13.7)	39.0 (14.1)	39.8 (15.4)	38.5 (12.7)	38 (12.2)
Sex								
Male	140 (48.6)	142 (49.3)	74 (49.0)	70 (46.4)	91 (50.3)	94 (51.9)	42 (53.9)	48 (60.0)
Female	148 (51.4)	146 (50.7)	77 (51.0)	81 (53.6)	90 (49.7)	87 (48.1)	36 (46.1)	32 (40.0)
UC disease duration, mean (SD), (years)	-	-	-	-	4.4 (5.7)	3.9 (5.3)	6.3 (6.5)	5.2 (5.1)
Aminosalicylates (oral) at cohort entry	135 (46.9)	147 (51.0)	71 (47.0)	81 (53.6)	94 (51.9)	94 (51.9)	-	-
Corticosteroids (oral)								
within 180 days before cohort entry	274 (95.1)	272 (94.4)	134 (88.7)	135 (89.4)	157 (86.7)	153 (84.5)	-	-
at cohort entry	216 (75.0)	217 (75.3)	101 (66.9)	95 (62.9)	119 (65.7)	116 (64.1)	31 (39.7)	38 (47.5)
Opioids	144 (50.0)	140 (48.6)	58 (38.4)	58 (38.4)	17 (9.4)	19 (10.5)	-	-
UC activity assessment [a]								
Hospitalization related to ulcerative colitis	140 (48.6)	136 (47.2)	42 (27.8)	49 (32.5)	48 (26.5)	58 (32.0)	-	-
Abdominal imaging	79 (27.4)	67 (23.3)	49 (32.5)	50 (33.1)	52 (28.7)	57 (31.5)	-	-
Lower GI endoscopy	251 (87.2)	244 (84.7)	132 (87.4)	126 (83.4)	147 (81.2)	147 (81.2)	-	-
Upper GI endoscopy	43 (14.9)	37 (12.8)	14 (9.3)	17 (11.3)	31 (17.1)	43 (23.8)	-	-
CRP tests ordered, mean (SD)	0.99 (1.29)	0.98 (1.13)	1.04 (2.76)	1.09 (1.43)	2.87 (2.49)	3.03 (2.54)	-	-
Fecal pathogen tests ordered	161 (55.9)	157 (54.5)	74 (49.0)	78 (51.7)	57 (31.5)	65 (35.9)	-	-
Comorbidities								
Clostridioides difficile infection	18 (6.3)	13 (4.5)	11 (7.3)	9 (6.0)	3 (1.7)	1 (0.6)	-	-
Serious infection	10 (3.5)	9 (3.1)	1 (0.7)	1 (0.7)	1 (0.6)	1 (0.6)	-	-
Cardiovascular disease	18 (6.3)	16 (5.6)	7 (4.6)	6 (4.0)	13 (7.2)	9 (5.0)	-	-
Chronic kidney failure	0	0	2 (1.3)	3 (2.0)	1 (0.6)	1 (0.6)	-	-
Chronic liver disease	0	0	7 (4.6)	5 (3.3)	1 (0.6)	2 (1.1)	-	-
Chronic pulmonary disease	43 (14.9)	41 (14.2)	18 (11.9)	21 (13.9)	36 (19.9)	31 (17.1)	-	-
Venous thromboembolism	8 (2.8)	8 (2.8)	5 (3.3)	7 (4.6)	3 (1.7)	3 (1.7)	-	-
Diabetes	10 (3.5)	19 (6.6)	12 (7.9)	12 (7.9)	6 (3.3)	8 (4.4)	-	-
Healthcare use characteristics [a]								
Hospitalizations not related to UC	37 (12.8)	31 (10.8)	13 (8.6)	11 (7.3)	18 (9.9)	13 (7.2)	-	-
Gastroenterologist visits, mean (SD)	5.84 (5.78)	5.55 (6.44)	9.08 (6.83)	9.39 (8.11)	7.06 (6.56)	7.40 (7.24)	-	-

[a] Assessed within 180 days before cohort entry; Abbreviations: Mono, monotherapy; Combo, combination therapy; UC, ulcerative colitis; CRP, C-reactive protein

Table 2. Outcome risk[a] after 1:1 propensity-score matching

	MarketScan, n (%)		Optum, n (%)		SNDS, n (%)		Overall cohort, n (%)	
Variable	Infliximab monotherapy (n=288)	Infliximab and thiopurines (n=288)	Infliximab monotherapy (n=151)	Infliximab and thiopurines (n=151)	Infliximab monotherapy (n=181)	Infliximab and thiopurines (n=181)	Infliximab monotherapy (n=620)	Infliximab and thiopurines (n=620)
Treatment failure	80 (27.8)	61 (21.2)	39 (25.8)	32 (21.2)	51 (28.2)	31 (17.1)	170 (27.4)	124 (20.0)
Hospitalization related to UC or colectomy [b]	32 (11.1)	18 (6.3)	17 (11.3)	12 (7.9)	19 (10.5)	11 (6.1)	68 (11.0)	41 (6.6)
Switch to another biologic or tofacitinib [b]	15 (5.2)	13 (4.5)	3 (2.0)	4 (2.6)	9 (5.0)	6 (3.3)	27(4.4)	23 (3.7)
Corticosteroids exposure at week 16 [b]	49 (17.0)	40 (13.9)	27 (17.9)	19 (12.6)	34 (18.8)	22 (12.2)	110 (17.7)	81 (13.1)

[a] 16-week risk

[b] Secondary outcomes are assessed without censoring at the first outcome occurrence in case of multiple outcome occurrences during follow-up. Abbreviation: UC, ulcerative colitis

Figure 1. Study Population Flowchart

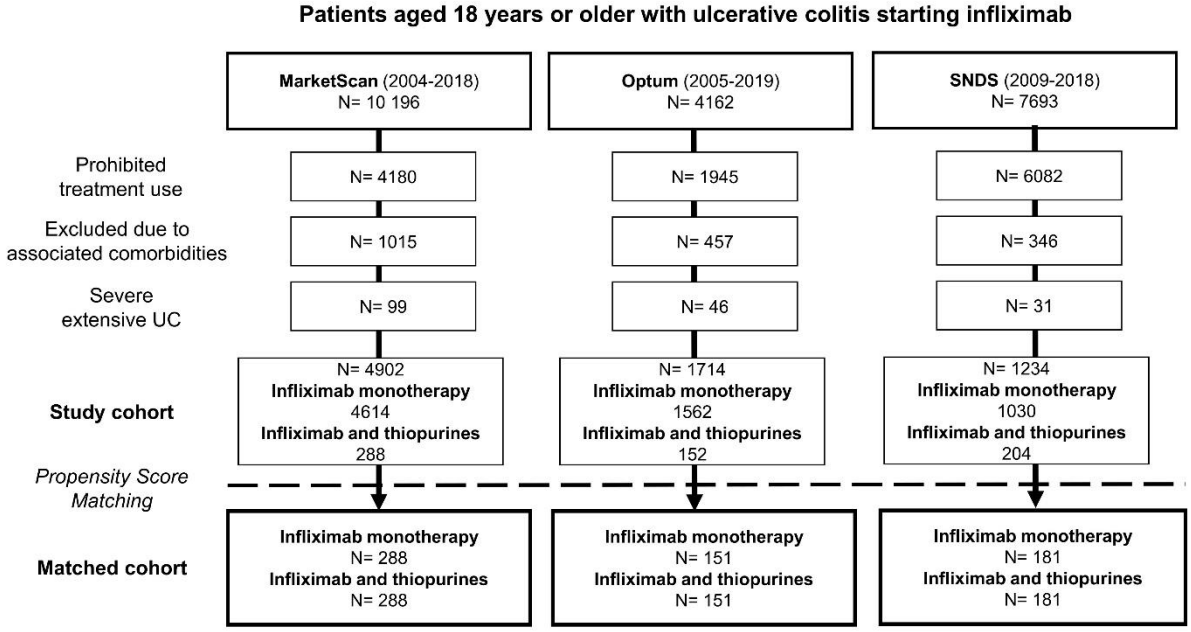


Figure 2. Risk ratios for treatment failure associated with combination therapy with infliximab and thiopurines compared to infliximab monotherapy

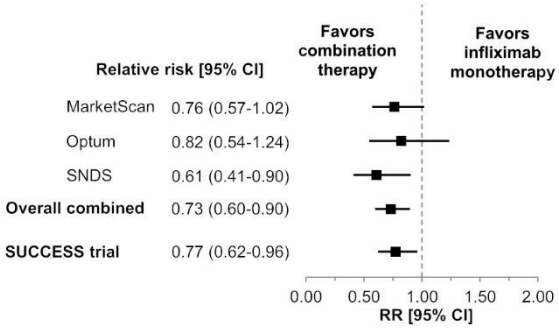
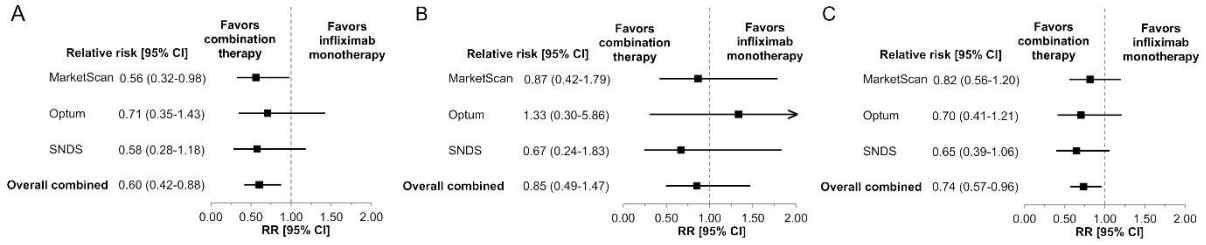


Figure 3. Risk ratios for each individual component of the composite outcome associated with combination therapy with infliximab and thiopurines compared to infliximab monotherapy (hospitalization related to UC or colectomy [A], treatment switch [B], and corticosteroids continuation [C])



Supplementary Material

Supplementary Table 1. Effectiveness outcome measure with related codes

Outcomes	ICD-10	ICD-9	Procedures	Anatomical Therapeutic Chemical (ATC) classification system code
Hospitalization related to ulcerative colitis or colectomy	K51; K56; K60; K62.5; K63.0; K63.1; K63.2; K65.0; K65.1; K92.2; R10	556; 560; 567.21; 567.22; 567.29; 569.3; 569.5; 569.81; 569.83; 578.9; 789.0	Colectomy	-
Switch to another biologic drug or tofacitinib	-	-	-	L04AB04 (adalimumab), L04AB05 (certolizumab pegol), L04AB06 (golimumab), L04AA23 (natalizumab), L04AA33 (vedolizumab), L04AC05 (ustekinumab), or L04AA29 (tofacitinib)
Exposure to corticosteroids at week 16	-	-	-	H02AB04 (methylprednisolone); H02AB06 (prednisolone, only IV or oral intake); H02AB07 (prednisone, only IV or oral intake); H02AB10 (cortisone, only IV or oral intake)

Supplementary Table 2. Characteristics of study patients treated with infliximab and thiopurines or infliximab monotherapy, before propensity score-matching

	MarketScan, n (%)		Optum, n (%)		SNDS, n (%)	
Variable	Infliximab Mono. (n=4614)	Infliximab Combo. (n=288)	Infliximab Mono. (n=1562)	Infliximab Combo. (n=152)	Infliximab Mono. (n=1030)	Infliximab Combo. (n=204)
Age, mean (SD), years	40.6 (14.5)	40.4 (14.0)	42.0 (15.9)	40.4 (13.8)	41.7 (16.1)	39.1 (15.1)
Sex						
Male	2174 (47.1)	142 (49.3)	736 (47.1)	70 (46.1)	522 (50.7)	107 (52.5)
Female	2440 (52.9)	146 (50.7)	826 (52.9)	82 (53.9)	508 (49.3)	97 (47.5)
IBD disease duration, mean (SD), years	-	-	-	-	4.1 (5.6)	3.8 (5.2)
Aminosalicylates (oral) at cohort entry	1450 (31.4)	147 (51.0)	658 (42.1)	82 (53.9)	519 (50.4)	106 (52.0)
Corticosteroids (oral)						
within 180 days before cohort entry	2556 (55.4)	272 (94.4)	1156 (74.0)	136 (89.5)	744 (72.2)	175 (85.8)
at cohort entry	1770 (38.4)	217 (75.3)	737 (47.2)	96 (63.2)	490 (47.6)	139 (68.1)
Opioids [a]	1395 (30.2)	140 (48.6)	470 (30.1)	59 (38.8)	108 (10.5)	24 (11.8)
Ulcerative colitis activity assessment [a]						
Hospitalization related to ulcerative colitis	1114 (24.1)	136 (47.2)	377 (24.1)	49 (32.2)	236 (22.9)	68 (33.3)
Abdominal imaging	833 (18.1)	67 (23.3)	473 (30.3)	51 (33.6)	281 (27.3)	67 (32.8)
Lower GI endoscopy	2983 (64.7)	244 (84.7)	1071 (68.6)	127 (83.6)	693 (67.3)	169 (82.8)
Upper GI endoscopy	417 (9.0)	37 (12.8)	189 (12.1)	17 (11.2)	194 (18.8)	51 (25.0)
CRP tests ordered, mean (SD)	0.65 (1.07)	0.98 (1.13)	1.02 (1.68)	1.09 (1.42)	2.58 (2.41)	3.30 (2.83)
Fecal pathogen tests ordered	1664 (36.1)	157 (54.5)	669 (42.8)	79 (52.0)	251 (24.4)	78 (38.2)
Comorbidities						
Clostridioides difficile infection	198 (4.3)	13 (4.5)	70 (4.5)	9 (5.9)	8 (0.8)	5 (2.5)
Serious infection	96 (2.1)	9 (3.1)	28 (1.8)	1 (0.7)	10 (1.0)	2 (1.0)
Cardiovascular disease	262 (5.7)	16 (5.6)	99 (6.3)	6 (3.9)	54 (5.2)	10 (4.9)
Chronic kidney failure	33 (0.7)	0 (0.0)	28 (1.8)	3 (2.0)	7 (0.7)	1 (0.5)
Chronic liver disease	85 (1.8)	0 (0.0)	56 (3.6)	5 (3.3)	11 (1.1)	2 (1.0)
Chronic pulmonary disease	383 (8.3)	41 (14.2)	167 (10.7)	21 (13.8)	199 (19.3)	36 (17.6)
Venous thromboembolism	124 (2.7)	8 (2.8)	38 (2.4)	7 (4.6)	18 (1.7)	3 (1.5)
Diabetes	307 (6.7)	19 (6.6)	124 (7.9)	12 (7.9)	74 (7.2)	8 (3.9)
Healthcare use characteristics [a]						
Hospitalizations not related to ulcerative colitis	436 (9.4)	31 (10.8)	134 (8.6)	11 (7.2)	90 (8.7)	19 (9.3)
Gastroenterologist visits, mean (SD)	4.78 (5.41)	5.55 (6.44)	7.70 (7.13)	9.40 (8.08)	5.94 (6.36)	7.68 (7.15)

[a] Assessed within 180 days before cohort entry; Abbreviations: Mono, monotherapy; Combo, combination therapy; UC, ulcerative colitis; CRP, C-reactive protein

Supplementary Table 3. Year of cohort entry before and after PS-matching

	MarketScan				Optum				SNDS			
	Before PS-matching		After PS-matching		Before PS-matching		After PS-matching		Before PS-matching		After PS-matching	
Variable	Infliximab Mono. (n=4614)	Infliximab Combo. (n=288)	Infliximab Mono. (n=288)	Infliximab Combo. (n=288)	Infliximab Mono. (n=1562)	Infliximab Combo. (n=152)	Infliximab Mono. (n=151)	Infliximab Combo. (n=151)	Infliximab Mono. (n=1030)	Infliximab Combo. (n=204)	Infliximab Mono. (n=181)	Infliximab Combo. (n=181)
Year of Cohort Entry												
2004	15 (0.3)	2 (0.7)	3 (1.0)	2 (0.7)	-	-	-	-	-	-	-	-
2005	50 (1.1)	9 (3.1)	7 (2.4)	9 (3.1)	26 (1.7)	8 (5.3)	10 (6.6)	8 (5.3)	-	-	-	-
2006	153 (3.3)	23 (8.0)	15 (5.2)	23 (8.0)	60 (3.8)	21 (13.8)	19 (12.6)	20 (13.2)	-	-	-	-
2007	205 (4.4)	17 (5.9)	24 (8.3)	17 (5.9)	66 (4.2)	10 (6.6)	7 (4.6)	10 (6.6)	-	-	-	-
2008	288 (6.2)	9 (3.1)	8 (2.8)	9 (3.1)	81 (5.2)	8 (5.3)	8 (5.3)	8 (5.3)	-	-	-	-
2009	420 (9.1)	19 (6.6)	15 (5.2)	19 (6.6)	105 (6.7)	14 (9.2)	14 (9.3)	14 (9.3)	104 (10.1)	25 (12.3)	23 (12.7)	21 (11.6)
2010	409 (8.9)	21 (7.3)	21 (7.3)	21 (7.3)	102 (6.5)	12 (7.9)	12 (7.9)	12 (7.9)	114 (11.1)	11 (5.4)	10 (5.5)	11 (6.1)
2011	614 (13.3)	30 (10.4)	30 (10.4)	30 (10.4)	95 (6.1)	7 (4.6)	5 (3.3)	7 (4.6)	102 (9.9)	21 (10.3)	19 (10.5)	19 (10.5)
2012	602 (13.0)	28 (9.7)	26 (9.0)	28 (9.7)	122 (7.8)	10 (6.6)	12 (7.9)	10 (6.6)	127 (12.3)	20 (9.8)	24 (13.3)	20 (11.0)
2013	377 (8.2)	23 (8.0)	20 (6.9)	23 (8.0)	97 (6.2)	10 (6.6)	10 (6.6)	10 (6.6)	110 (10.7)	26 (12.7)	18 (9.9)	24 (13.3)
2014	460 (10.0)	28 (9.7)	33 (11.5)	28 (9.7)	81 (5.2)	7 (4.6)	8 (5.3)	7 (4.6)	125 (12.1)	24 (11.8)	27 (14.9)	23 (12.7)
2015	253 (5.5)	24 (8.3)	26 (9.0)	24 (8.3)	105 (6.7)	10 (6.6)	13 (8.6)	10 (6.6)	124 (12.0)	22 (10.8)	18 (9.9)	21 (11.6)
2016	321 (7.0)	25 (8.7)	30 (10.4)	25 (8.7)	161 (10.3)	11 (7.2)	13 (8.6)	11 (7.3)	134 (13.0)	18 (8.8)	16 (8.8)	18 (9.9)
2017	258 (5.6)	18 (6.3)	15 (5.2)	18 (6.3)	146 (9.3)	12 (7.9)	10 (6.6)	12 (7.9)	63 (6.1)	27 (13.2)	16 (8.8)	15 (8.3)
2018	189 (4.1)	12 (4.2)	15 (5.2)	12 (4.2)	164 (10.5)	5 (3.3)	2 (1.3)	5 (3.3)	27 (2.6)	10 (4.9)	10 (5.5)	9 (5.0)
2019	-	-	-	-	151 (9.7)	7 (4.6)	8 (5.3)	7 (4.6)	-	-	-	-

Abbreviations: Mono, monotherapy; Combo, combination therapy

Supplementary Table 4. Outcome risk[a] , before propensity-score matching

Variable	MarketScan, n (%)		Optum, n (%)		SNDS, n (%)		Overall cohort, n (%)	
	Infliximab monotherapy (n=4614)	Infliximab and thiopurines (n=288)	Infliximab monotherapy (n=1562)	Infliximab and thiopurines (n=152)	Infliximab monotherapy (n=1030)	Infliximab and thiopurines (n=204)	Infliximab monotherapy (n=7206)	Infliximab and thiopurines (n=644)
Treatment failure	755 (16.4)	61 (21.2)	287 (18.4)	32 (21.1)	263 (25.5)	34 (16.7)	1305 (18.1)	127 (19.7)
Hospitalization related to UC or colectomy [b]	297 (6.4)	18 (6.3)	99 (6.3)	12 (7.9)	118 (11.5)	12 (5.9)	514 (7.1)	42 (6.5)
Switch to another biologic or tofacitinib [b]	103 (2.2)	13 (4.5)	41 (2.6)	4 (2.6)	39 (3.8)	8 (3.9)	183 (2.5)	25 (3.9)
Corticosteroids exposure at week 16 [b]	461 (10.0)	40 (13.9)	196 (12.5)	19 (12.5)	171 (16.6)	23 (11.3)	828 (11.5)	82 (12.7)

[a] 16-week risk

[b] Secondary outcomes are assessed without censoring at the first outcome occurrence in case of multiple outcome occurrences during follow-up. Abbreviation: UC, ulcerative colitis

Supplementary Table 5. Relative risk of treatment failure in patients treated with combination therapy versus infliximab monotherapy, subgroups analyses

	MarketScan	Optum	SNDS	Overall combined
Subgroup				
Corticosteroids at infliximab initiation				
Yes	0.85 (0.62-1.17)	0.72 (0.45-1.17)	0.60 (0.40-0.91)	0.74 (0.59-0.93)
No	0.71 (0.34-1.49)	1.11 (0.49-2.50)	0.40 (0.13-1.20)	0.75 (0.46-1.22)

Supplementary Table 6. Relative risk of treatment failure in patients treated with combination therapy versus infliximab monotherapy, sensitivity analyses

Analyses	MarketScan	Optum	SNDS	Overall combined
Variables with an absolute standardized difference greater than 0.1 after PS-matching included as covariate in log–binomial regression models	0.76 (0.57-1.02)	0.82 (0.55-1.24)	0.61 (0.41-0.90)	0.73 (0.6-0.90)
Propensity score-matching with a 1:4 variable ratio	0.86 (0.63-1.16)	0.90 (0.59-1.37)	0.61 (0.41-0.91)	0.79 (0.64-0.97)
Analysis excluding treatment switch as outcome	0.71 (0.51-0.97)	0.81 (0.53-1.24)	0.61 (0.40-0.93)	0.71 (0.57-0.88)
Analysis extending the enrollment period to 365 days in the US databases	0.79 (0.56-1.12)	1.00 (0.63-1.59)	0.61 (0.41-0.90)	0.77 (0.61-0.96)
Analysis excluding patients starting thiopurines in the first month after infliximab initiation	0.79 (0.59-1.06)	0.78 (0.51-1.20)	0.61 (0.40-0.92)	0.74 (0.60-0.91)

Supplementary Figure 1. Standardized differences before and after propensity score matching (MarketScan [A], Optum[B], and SNDS [C])