Title

Real-world effectiveness of ustekinumab and vedolizumab in TNF-exposed pediatric patients with ulcerative colitis

Journal

Journal of Pediatric Gastroenterology and Nutrition, 78(5)

ISSN

0277-2116

Authors

Patel, Perseus V
Zhang, Amy
Bhasuran, Balu
et al.

Publication Date

2024-05-01

DOI

10.1002/jpn3.12169

Peer reviewed

Real World Effectiveness of Ustekinumab and Vedolizumab in TNF-exposed Pediatric Patients with Ulcerative Colitis

Perseus V. Patel, MD[1,2]; Amy Zhang, MPH[3]; Balu Bhasuran, PhD[4]; Vignesh G. Ravindranath, BS[4]; Melvin B. Heyman, MD[1]; Sofia G. Verstraete, MD[1]; Atul J. Butte, MD, PhD[1,5]; Michael J. Rosen, MD, MSCI[2]; Vivek A. Rudrapatna, MD, PhD[4,5]; ImproveCareNow Pediatric IBD Learning Health System

[1]Department of Pediatrics, University of California San Francisco, San Francisco, CA
[2]Division of Pediatric Gastroenterology, Stanford University School of Medicine, Palo Alto, CA
[3]Quantitative Sciences Unit, Stanford University School of Medicine, Palo Alto, CA
[4]Bakar Computational Health Sciences Institute, University of California San Francisco, San Francisco, CA
[5]Division of Gastroenterology, Department of Medicine, University of California San Francisco, San Francisco, CA

Author Contributions

Perseus Patel (corresponding author): patelpe@stanford.edu

- Study conception and design.
- Acquisition, analysis, and interpretation of the data.
- Primary author of the manuscript.

Amy Zhang:

- Substantial contribution to analysis and interpretation of data.
- Critical revision of the manuscript for important intellectual content.

Balu Bhasuran:

- Substantial contribution to analysis and interpretation of data.
- Critical revision of the manuscript for important intellectual content.

Vignesh Ravindranath:

- Substantial contribution to analysis and interpretation of data.
- Critical revision of the manuscript for important intellectual content.

Melvin B. Heyman:

- Significant contributions to project design and analysis plan.
- Critical revision of the manuscript for important intellectual content.

Sofia G. Verstraete:

- Significant contributions to project design and analysis plan.
- Critical revision of the manuscript for important intellectual content.

Michael J. Rosen:

- Significant contributions to study design.
- Substantial contribution to analysis and interpretation of data.
- Critical revision of the manuscript for important intellectual content.

Atul J. Butte:

- Critical revision of the manuscript for important intellectual content.

Vivek A. Rudrapatna:

- Essential in study conception and design.
- Substantial contribution to analysis and interpretation of data.
- Critical revision of the manuscript for important intellectual content.

Postal Address:
750 Welch Road, Suite 116
Palo Alto, CA 94304

Funding Sources: This work is supported in part by NIH T32 DK007762 (PVP). Research reported in this publication was supported by the National Library of Medicine of the National Institutes of Health under Award Number K99LM014099 (to VAR), and the National Center for Advancing Translational Science (NCATS) UCLA CTSI Grant Number TL1TR001883. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Additional funding provided by the UCSF Division of Gastroenterology, Bakar Computational Health Sciences Institute, and the Stanford Medicine Children's Health Center for IBD and Celiac Disease.
Word Count: 2984
Number of Figures: 1
Number of Tables: 3
Key Words: ~~children~~pediatrics, inflammatory bowel disease, ulcerative colitis, biologic therapy, clinical data science, comparative effectiveness, real-world evidence
Conflicts of Interest:
VAR receives grant funding from Alnylam, Genentech, Merck, Takeda, Janssen, Blueprint Medicines, Stryker, and Mitsubishi Tanabe.
~~**VAR** receives grant support from the following for-profit entities: Janssen Inc, Alnylam Inc, Merck Inc, and Genentech Inc.~~
AJB is a co-founder and consultant to Personalis and NuMedii; consultant to Mango Tree Corporation, and in the recent past, Samsung, 10x Genomics, Helix, Pathway Genomics, and Verinata (Illumina); has served on paid advisory panels or boards for Geisinger Health, Regenstrief Institute, Gerson Lehman Group, AlphaSights, Covance, Novartis, Genentech, and Merck, and Roche; is a shareholder in Personalis and NuMedii; is a minor shareholder in Apple, Meta (Facebook), Alphabet (Google), Microsoft, Amazon, Snap, 10x Genomics, Illumina, Regeneron, Sanofi, Pfizer, Royalty Pharma, Moderna, Sutro, Doximity, BioNtech, Invitae, Pacific Biosciences, Editas Medicine, Nuna Health, Assay Depot, and Vet24seven, and several other non-health related companies and mutual funds; and has received honoraria and travel reimbursement for invited talks from Johnson and Johnson, Roche, Genentech, Pfizer, Merck, Lilly, Takeda, Varian, Mars, Siemens, Optum, Abbott, Celgene, AstraZeneca, AbbVie, Westat, and many academic institutions, medical or disease specific foundations and associations, and health systems. Atul Butte receives royalty payments through Stanford University, for several patents and other disclosures licensed to NuMedii and Personalis. Atul Butte's research has been funded by NIH, Peraton (as the prime on an NIH contract), Genentech, Johnson and Johnson, FDA, Robert Wood Johnson Foundation, Leon Lowenstein Foundation, Intervalien Foundation, Priscilla Chan and Mark Zuckerberg, the Barbara and Gerson Bakar Foundation, and in the recent past, the March of Dimes, Juvenile Diabetes Research Foundation, California Governor's Office of Planning and Research, California Institute for Regenerative Medicine, L'Oreal, and Progenity.
MJR has served on advisory boards for Pfizer, and Fresenius Kabi, and as a consultant for Federation Bio.

Abstract

Background

Vedolizumab (VDZ) and ustekinumab (UST) are second-line treatments in pediatric patients with ulcerative colitis (UC) refractory to anti-tumor necrosis factor (anti-TNF) therapy. Pediatric studies comparing the effectiveness of these medications are lacking.

Aim

Using a registry from ImproveCareNow (ICN), a global research network in pediatric inflammatory bowel disease, we compared the effectiveness of UST and VDZ in anti-TNF refractory UC.

Methods

We performed a propensity-score weighted regression analysis to compare corticosteroid free clinical remission (CFCR) at 6 months from starting second-line therapy. Sensitivity analyses tested the robustness of our findings to different ways of handling missing outcome data. Secondary analyses evaluated alternative proxies of response and infection risk.

Results

Our cohort included 262 patients on VDZ and 74 patients on UST. At baseline, the two groups differed on their mean pediatric UC activity index (PUCAI) (p=0.03) but were otherwise similar. At month 6, 28.3% of patients on VDZ and 25.8% of those on UST achieved CFCR (p=0.76). Our primary model showed no difference in CFCR [odds ratio 0.81; 95% CI: 0.41-1.59] (p=0.54). The time to biologic discontinuation was similar in both groups [hazard ratio 1.26; 95% CI: 0.76-2.08] (p=0.36), with the reference group being VDZ, and we found no differences in clinical response, growth parameters, hospitalizations, surgeries, infections, or malignancy risk. Sensitivity analyses supported these findings of similar effectiveness.

Conclusions

UST and VDZ are similarly effective for treating anti-TNF refractory UC in pediatric patients. Providers should consider safety, tolerability, cost, and comorbidities when deciding between these therapies.

What is Known?

- Anti-tumor necrosis factor (anti-TNF) medications are first line therapy for pediatric patients with moderate to severe ulcerative colitis.
- A significant proportion of patients will be primary non-responders or experience loss of response to anti-TNFs.
- Vedolizumab (VDZ) and ustekinumab (UST) are common second-line treatments for these patients.

What is New?

- UST and VDZ have similar effectiveness in this cohort, with equivalent rates of corticosteroid-free clinical remission, clinical response, hospitalizations, surgeries, and biologic persistence.
- These results are consistent across multiple methods of correcting for missing data in the ImproveCareNow (ICN) Registry.

Introduction

In contrast to its adult-onset disease, pediatric patients with ulcerative colitis (pUC) commonly present with extensive disease associated with higher rates of hospitalization, corticosteroid failure, colorectal cancer risk, and up to 20% colectomy rate within 5-years of diagnosis[1, 3–6]. Anti-tumor necrosis factor (anti-TNF) medications are first-line treatments for

moderate to severe pUC, with adalimumab and infliximab being the only Food and Drug Administration (FDA) approved biologics for children [6, 7]. However, up to 30% of patients do not respond to anti-TNF induction doses, and as many as 45% of responders subsequently lose response within the first year [1, 7]. Patients who are refractory to or intolerant of these drugs usually require treatments with different mechanisms of action, with providers and families having to decide between off-label therapies with less available information about treatment efficacy.

Typical options for these patients include vedolizumab (VDZ), an inhibitor of leukocyte trafficking to the intestinal tract [8], and ustekinumab (UST), an inhibitor of interleukins 12 and 23 [9]. Despite how commonly these medications are needed and applied, an evidence gap exists regarding how to optimally select between these options. At present, no published studies compare efficacy in children. Furthermore, evidence extrapolated from adult studies is inconsistent. A meta-analysis in adult patients with UC showed that both were effective at inducing clinical remission and endoscopic response, while a systematic review in adults with Crohn's Disease favored sustained corticosteroid-free clinical remission (CFCR) with UST[10, 11]. While some comparative effectiveness studies indicate no difference in adults with Crohn's Disease, one favored VDZ for clinical remission, while another suggested UST was superior for CFCR[12–16]. It is possible that these conflicting findings may be due in part to differences in outcome measures and methods applied for handling missing data.

We sought to address this evidence gap and study the comparative effectiveness of VDZ and UST in anti-TNF-refractory pUC using the ImproveCareNow (ICN) registry. We evaluated multiple measures of effectiveness including CFCR, clinical response, growth and nutritional status, inflammatory bowel disease (IBD)-related hospitalizations, or surgery at 6 months, and

time to treatment discontinuation. We also evaluated safety outcomes at 6 months by comparing rates of severe infections and malignancies. Furthermore, we tested the robustness of our results across different potential explanations for missing outcome measures in the context of routine clinical care.

Methods

Data source

We conducted a prospective cohort study of pUC within ICN initiating vedolizumab or ustekinumab therapy as a second line therapy after failure of an anti-TNF drug between October 1, 2006, and April 5, 2023. ICN is a pediatric IBD research network initiated in 2006 and has captured data on over 32,000 patients treated across >100 centers. It functions as a learning health system, with a continuously updated database that facilitates patient-centered research to optimize outcomes[17]. Institutional review board (IRB) approval is obtained, and participants are enrolled under informed consent at each site. Data are collected at each clinic visit and hospitalization. The captured information includes demographics, disease characteristics, serum and stool tests, medication history, IBD-related surgical history, growth measures, and disease activity scores.

Cohort selection criteria

We queried the ICN database to identify pediatric patients (< 21 years old) meeting the following criteria: 1) a documented, confirmed diagnosis of UC based on the diagnosis at registration and the most recent visit, 2) anti-TNF exposure based on medications documented during clinic visits, and 3) VDZ or UST exposure documented in the clinic visit medication list, without previous exposure documented to either drug (i.e., VDZ or UST as third line therapy were excluded). Concomitant medications such as immunomodulators or antibiotics were

allowed. We excluded patients with a diagnosis of CD or inflammatory bowel disease unclassified, history of colectomy, and those who were started on VDZ or UST prior to anti-TNF therapy. These queries identified 381 pUC who failed anti-TNFs and were subsequently treated with UST or VDZ. Anti-TNF failure was defined as anti-TNF use and cessation prior to treatment with VDZ or UST. Patients must have exhibited active disease at the visit prior to the first recorded visit on VDZ or UST, defined as a Pediatric Ulcerative Colitis Activity Index (PUCAI) $\geq$10[18] or a physician global assessment (PGA) > 1 (if PUCAI not available), or current corticosteroid use, or abnormal inflammatory markers (C-reactive protein $\geq$ 5mg/L). These markers were selected in alignment with guideline-based early treat-to-target goals[19].

Baseline and outcome periods

The timing of clinic visits and data capture in real-world settings commonly differs from those in controlled studies due to a variety of factors, including provider practice patterns, patients' state of illness or health, and social determinants of health impacting the ability to follow-up[20]. ICN is a learning collaborative that disseminates best practice recommendations[21], but there are no formal guidelines on monitoring pediatric patients on UST or VDZ. Therefore, patients in our study were monitored based on individual center protocols and provider recommendations. To emulate a hypothetical trial with planned study visits and to minimize the effects of missing data and selection bias, we performed an exploratory data analysis and retrospectively defined time windows corresponding to the baseline and outcome periods.

We defined the baseline as the clinic visit before the first documentation of VDZ or UST use for a patient. Although this corresponds to a date prior to the true date of treatment induction (which is not well-captured across the dataset), we chose this definition because it most accurately reflects the patient's clinical status at the time of deciding between second-line

115 therapies. To address missing baseline data, we used the *sklearn* implementation of Iterative
116 Imputation , a method that iteratively imputes each missing variable as a function of all other
117 variables until convergence is achieved [22].

118 Our outcome period was month 6 ± 2 months after baseline. This minimized missing
119 data, accounted for differences in medication time to onset, and was in-line with previous IBD
120 comparative effectiveness literature [14]. For those who had multiple clinic visits during the
121 outcome period, we selected data from the visit closest to day 183 relative to baseline.

122 *Study Endpoints*

123 Our primary endpoint was corticosteroid-free clinical remission (CFCR), a composite
124 binary variable defined as a PUCAI <10 or PGA=1, in addition to the absence of all of the
125 following during the outcome period: therapy discontinuation, IBD-related hospitalization or
126 surgery (pre-defined variables in ICN), or continued corticosteroid use (including budesonide) at
127 the outcome visit. [1, 23] PGA was used in visits where a complete PUCAI was not documented.
128 As fecal calprotectin was collected in only 10% of visits during the outcome window, we did not
129 incorporate it into our primary endpoint.

130 Our secondary endpoints included corticosteroid-free clinical response, IBD-related
131 hospital admissions and surgeries, growth, nutrition, infections, malignancies, and biologic
132 durability. Clinical response was a binary outcome that included patients in remission, and as per
133 the STRIDE-II guidelines included those with a decrease in PUCAI $\geq$ 20[19], or a PGA indicating
134 mild or quiescent disease. Growth and nutritional status were evaluated based upon pre-defined
135 ordinal variables in the ICN database. Serious infection, a pre-defined ICN variable,
136 encompassed any infection that required hospitalization or intravenous treatment. As an

alternative measure of effectiveness, we evaluated time to treatment discontinuation. This was determined through the current medications documented at each visit after the baseline visit.

Statistical Methods

Primary Outcomes

We used a stabilized inverse probability of treatment weighted (IPTW) logistic regression model to compare the likelihood of CFCR between treatment groups. Baseline differences were assessed using traditional significance testing (Fisher's exact test, Student's t-test, Mann-Whitney U as appropriate) as well as standardized mean differences (SMDs) given the disparity in cohort sizes. ~~As per established measures, w~~We used >0.2, >0.5, and >0.8 cutoffs for SMD to represent small, moderate, and large differences[24, 25]. For this stabilized IPTW model, the numerator for the weights was the probability of treatment selection irrespective of the covariates, and the denominator represented the probability of treatment selection based on all baseline variables in Table 1[26].

We performed a complete case analysis as our primary approach to analyzing the data. This approach assumes that the patients with captured outcomes are representative of the study population. We considered this assumption reasonable with the understanding that many missing outcome measurements at month 6 may reflect idiosyncratic practice styles around the timing of follow-up, patient preferences, and other factors that are independent of a patient's clinical status. However, given the possibility of alternative explanations for missing data in real-world contexts, we performed sensitivity analyses to test the robustness of our primary findings, as detailed in our supplemental methods.

Secondary Outcomes

We used a logistic regression analysis adjusted for baseline differences in our treatment cohorts to compare the odds of patients achieving a clinical response at month-6. We also used Fisher's exact testing to compare growth status, nutritional status, prevalence of hospitalizations, surgeries, malignancies, and serious infections between the treatment groups from baseline until the month-6 follow-up visit.

To compare the rates of medication discontinuation for up to 1 year of follow-up, we conducted a multivariate analysis using a Cox regression that controlled for baseline differences noted in Table 1. Follow-up was censored at the date of medication change. Patients who had been followed for < 1 year after starting their second-line therapy were censored at the time of their last visit.

Ethics

The Human Research Protection Program IRB at the University of California San Francisco (IRB#21-34392) approved this study. At each participating center, patients were consented for their data to be shared with ICN and used for research.

Results

Data Capture

At baseline, disease characteristics such as Paris Classification, extra-intestinal manifestations, growth status, and current medications were captured in >95% of patients. Serum lab values were documented in 80-94% of patients and were similar across the two groups at baseline (Supplemental Table 1).

The primary outcome of CFCR was captured in 89% of the cohort (299/336). This degree of outcome capture was similar in both the VDZ (88.9%; 233/262) and UST (89.1%; 66/74) sub-cohorts (p=0.95).

Baseline Characteristics

Our study included 336 patients started on either VDZ (262) or UST (74). Patients were seen at 76 ICN centers, all in the United States. The treatment groups differed slightly on PUCAI scores [VDZ 26.8 (IQR: 10-40) vs UST 20.0 (IQR: 5-30); SMD=0.29] and concomitant immunomodulator use [VDZ 21.4% vs UST 12.2%; SMD=0.23] (Supplemental Table 1, Supplemental Figure 1). Otherwise, the groups were well-balanced with no differences were found in gender, age at diagnosis, age at initiating second-line biologic therapy, corticosteroid use, baseline serum labs, growth trajectories, nutritional status, PGA, or extra-intestinal manifestations (Supplemental Table 1). Descriptive characteristics of the cohort remained stable after imputation (Table 1), and stabilized IPTW ensured appropriate balance across all baseline variables during modeling.

Primary Endpoint

We performed a complete case analysis as our primary method for analyzing the data. In the study cohort, the proportions of CFCR were similar across treatment arms: 28.3% (VDZ) versus 25.8% (UST). An unadjusted odds ratio (OR) showed no difference between the two medications; OR=0.88 [95% confidence interval (CI) 0.47-1.63] (p=0.68). In the propensity-weighted model, the differences remained non-significant; OR= 0.81 [0.41-1.59] (p=0.54, Table 2), suggesting similar effectiveness.

Sensitivity Analyses

We performed two sensitivity analyses to test whether our findings would remain stable under alternative methods for handling missing outcomes, an issue that affected 11% of our cohort. These included: 1) non-responder imputation (patients missing outcomes were assumed to be non-responders) to model scenarios where patients who are too ill to follow-up, and 2)

responder imputation (patients missing outcomes were assumed to be responders) to model scenarios where patients who are well are less inclined to follow-up. These analyses supported the conclusion of similar effectiveness between UST and VDZ (Table 2).

Secondary Endpoints

Clinical response was achieved in 39.7% of patients on VDZ and 30% of patients on UST (Fisher's exact test, p=0.20) (Table 3). In a multivariate regression model that controlled for baseline differences in PUCAI and immunomodulator use, the odds of achieving clinical response remained non-significant (OR: 0.65 [0.36-1.15], p=0.14). The only predictor of achieving clinical response (OR: 0.36 [0.21-0.63], p<0.001) or CFCR (OR: 0.33 [0.18-0.62], p<0.001) at month-6 was not needing corticosteroids at baseline.

At their outcome visit, 82.7% of patients on VDZ and 84.2% of patients on UST had a satisfactory nutritional status (p=0.83) while 89.2% of patients on VDZ and 94.0% of patients on UST had a satisfactory growth status (p=0.24) (Table 3). The proportion of patients hospitalized were similar between the treatment arms—12.2% (VDZ) versus 10.8% (UST) (p=0.74). Surgeries and infections were rare (<10 patients), and no malignancies were reported in either group (Table 3).

By 1-year from the start of their second-line biologic, 32% of patients who were started on VDZ and 40% of patients who were started on UST discontinued the medication. We performed a Cox regression that controlled for baseline covariates and found the hazard of medication discontinuation among those on UST was 1.26 times that of patients on VDZ (95% CI: 0.76-2.08, p=0.36) (Figure 1).

Discussion

This is the first comparative effectiveness study of second-line therapies in pediatric patients with UC who failed anti-TNF therapy. Studies on comparative effectiveness in adult populations between these two drugs demonstrated varied results[11–16]. Using the ICN registry, we found that VDZ and UST are similarly effective at inducing CFCR at month-6. This finding of similar efficacy remained robust under a range of sensitivity analyses and secondary effectiveness and safety endpoints. While there are inherent limitations in ~~observational work~~research using observational data, our study helps alleviate the uncertainty that accompanies extrapolation from adult cohorts and serves as hypothesis-generating work for future prospective studies that help define treatment positioning in pUC.

We found a 25-30% remission rate and 30-40% response rate for the two drugs. Clinical trial data for multiple biologic therapies show lower remission rates in TNF-exposed versus TNF-naïve patients [27–30]. Patients who do not respond to anti-TNF are approximately 25% less likely to respond to second-line biologics, and both UST and VDZ are more effective in anti-TNF naïve patients with UC[10, 31]. In a post-hoc analysis of the GEMINI trial, 36.1 % of anti-TNF refractory UC patients achieved remission with VDZ [28]. A prospective trial in TNF-refractory pUC showed that 24% of the pediatric patients achieved CFCR at 6-months [32].

We strove to address limitations in existing real-world studies through the analysis of multiple endpoints, use of propensity scores to address selection bias, and selective measures to address reasons for incomplete data. We utilized both inpatient and outpatient elements in the ICN database, including hospitalizations and surgeries, to define clinical remission. Propensity scoring minimized selection bias while maintaining cohort size and minimizing loss of power [33]. Our results were consistent across multiple models that represented underlying reasons for

missingness in real-world settings. Lastly, we compared alternative proxies of effectiveness, which also showed equivalency. The slightly higher persistence of VDZ was not statistically significant and may be due to its longer time-to-onset or due to lack of other FDA-approved options. Overall, the consistent findings noted throughout our analyses support the hypothesis of similar effectiveness in anti-TNF refractory pUC.

The potential for residual bias exists in ~~observational work~~all retrospective research. Few pediatric resources are as large as ICN, but the nature of collecting data concurrent with clinical care creates heterogeneity in data collection. This is commonly due to practice variation among providers and heterogeneity in patients' disease states, with diagnostics being ordered based on health status rather than a pre-set schedule as seen in prospective trials. Therefore, patients may not have had a clinic visit during our outcome window, or information from the visit may not have been appropriately documented. Additionally, as labs collected outside the treating institution are not routinely uploaded into the database, we were unable to include biomarkers or therapeutic drug monitoring into our study. The database does not accurately capture medication induction dates, dosing changes or reasons for discontinuation.

An additional limitation of our study is the sample size discrepancy between the two groups, which may reduce precision. This is likely because VDZ was approved earlier than UST for IBD, and parents may prefer VDZ given its lower systemic immune suppression[34, 35]. In a recent ICN study, sequential anti-TNF therapy with infliximab then adalimumab or vice-versa were the most common patterns of biologic use in pediatric IBD, with significantly lower rates of drugs with alternative mechanisms of action[36]. This likely speaks to the limited FDA-approved therapies, and possible provider and/or patient comfort with using off-label biologics. As ICN grows, these imbalances may decrease over time.

Future work should focus on improving data quality so that ~~real-world evidence~~ studies using real-world data can more closely ~~mimic~~ emulate randomized clinical trials. Potential drivers of missing data in ICN include the variation in follow-up timing, differences in clinic note templates across centers, and the large number of data points to collect. To combat variability in free-text information and the burden of information extraction, natural language processing techniques have shown promise in extracting extraintestinal manifestations of IBD from clinical notes, and Mayo scores from colonoscopy reports[37, 38]. Future endeavors to automate the extraction of IBD-relevant variables from the electronic health record, including patient symptoms, radiographic, endoscopic, and histologic measures, can improve the completeness of our registries and enhance the quality of downstream research.

In conclusion, we found that VDZ and UST have similar rates of CFCR in anti-TNF refractory pUC. The replicability of our results supports that large disparities do not exist, but physicians should apply these findings cautiously given the inherent limitations of observational work. Larger randomized clinical trials are needed to validate these findings. As treat-to-target goals evolve, the inclusion of biomarkers, endoscopic, and imaging data will improve future comparisons between medications. However, this initial hypothesis-generating study proves that ICN has the potential to provide insight into clinically relevant questions that would otherwise require costly, time-consuming trials. Given the current data, we recommend that clinicians adopt a patient-specific approach to this decision that weights safety, tolerability, cost, route of administration, patient preference, and alternative indications in addition to treatment effectiveness.

Acknowledgements: We thank ICN for providing access to their database, Kelly Olano for her support with data extraction, NIH T32 DK007762 and Stanford Medicine Children's Health Center for IBD and Celiac Disease for PVP's support, and the UCSF Information Commons team for their computing support.

References

1. Turner D, Ruemmele FM, Orlanski-Meyer E, et al. Management of Paediatric Ulcerative Colitis, Part 1: Ambulatory Care—An Evidence-based Guideline From European Crohn's and Colitis Organization and European Society of Paediatric Gastroenterology, Hepatology and Nutrition. *Journal of Pediatric Gastroenterology and Nutrition* 2018; 67:257–291. https://doi.org/10.1097/MPG.0000000000002035

2. Turner D, Ruemmele FM, Orlanski-Meyer E, et al. Management of Paediatric Ulcerative Colitis, Part 2: Acute Severe Colitis—An Evidence-based Consensus Guideline From the European Crohn's and Colitis Organization and the European Society of Paediatric Gastroenterology, Hepatology and Nutrition. *Journal of Pediatric Gastroenterology and Nutrition* 2018; 67:292–310. https://doi.org/10.1097/MPG.0000000000002036

3. Turner D, Walsh CM, Benchimol EI, et al. Severe paediatric ulcerative colitis: incidence, outcomes and optimal timing for second-line therapy. *Gut* 2008; 57:331–338. https://doi.org/10.1136/gut.2007.136481

4. Levine A, de Bie CI, Turner D, et al. Atypical disease phenotypes in pediatric ulcerative colitis: 5-year analyses of the EUROKIDS Registry. *Inflamm Bowel Dis* 2013; 19:370–377. https://doi.org/10.1002/ibd.23013

5. Gower-Rousseau C, Dauchet L, Vernier-Massouille G, et al. The natural history of pediatric ulcerative colitis: a population-based cohort study. *Am J Gastroenterol* 2009; 104:2080–2088. https://doi.org/10.1038/ajg.2009.177

6. Croft NM, Faubion WA, Kugathasan S, et al. Efficacy and safety of adalimumab in paediatric patients with moderate-to-severe ulcerative colitis (ENVISION I): a randomised, controlled, phase 3 study. *Lancet Gastroenterol Hepatol* 2021; 6:616–627. https://doi.org/10.1016/S2468-1253(21)00142-4

7. Hyams J, Damaraju L, Blank M, et al. Induction and maintenance therapy with infliximab for children with moderate to severe ulcerative colitis. *Clin Gastroenterol Hepatol* 2012; 10:391-399.e1. https://doi.org/10.1016/j.cgh.2011.11.026

8. Neurath MF. Current and emerging therapeutic targets for IBD. *Nat Rev Gastroenterol Hepatol* 2017; 14:269–278. https://doi.org/10.1038/nrgastro.2016.208

9. Guillo L, D'Amico F, Danese S, et al. Ustekinumab for Extra-intestinal Manifestations of Inflammatory Bowel Disease: A Systematic Literature Review. *Journal of Crohn's and Colitis* 2021; 15:1236–1243. https://doi.org/10.1093/ecco-jcc/jjaa260

10. Burr NE, Gracie DJ, Black CJ, et al. Efficacy of biological therapies and small molecules in moderate to severe ulcerative colitis: systematic review and network meta-analysis. *Gut* 2022; 71:1976–1987. https://doi.org/10.1136/gutjnl-2021-326390

11. Parrot L, Dong C, Carbonnel F, et al. Systematic review with meta-analysis: the effectiveness of either ustekinumab or vedolizumab in patients with Crohn's disease refractory to anti-tumour necrosis factor. *Alimentary Pharmacology & Therapeutics* 2022; 55:380–388. https://doi.org/10.1111/apt.16714

12. Alric H, Amiot A, Kirchgesner J, et al. The effectiveness of either ustekinumab or vedolizumab in 239 patients with Crohn's disease refractory to anti-tumour necrosis factor. *Aliment Pharmacol Ther* 2020; 51:948–957. https://doi.org/10.1111/apt.15706

13. Townsend T, Razanskaite V, Dodd S, et al. Comparative effectiveness of ustekinumab or vedolizumab after one year in 130 patients with anti-TNF-refractory Crohn's disease. *Alimentary Pharmacology & Therapeutics* 2020; 52:1341–1352. https://doi.org/10.1111/apt.16057

14. Biemans VBC, van der Woude CJ, Dijkstra G, et al. Ustekinumab is associated with superior effectiveness outcomes compared to vedolizumab in Crohn's disease patients with prior failure to anti-TNF treatment. *Aliment Pharmacol Ther* 2020; 52:123–134. https://doi.org/10.1111/apt.15745

15. Kappelman MD, Adimadhyam S, Hou L, et al. Real-World Evidence Comparing Vedolizumab and Ustekinumab in Antitumor Necrosis Factor-Experienced Patients With Crohn's Disease. *Official journal of the American College of Gastroenterology | ACG* 2023; 118:674. https://doi.org/10.14309/ajg.0000000000002068

16. Onali S, Pugliese D, Caprioli FA, et al. An Objective Comparison of Vedolizumab and Ustekinumab Effectiveness in Crohn's Disease Patients' Failure to TNF-Alpha Inhibitors. *Official journal of the American College of Gastroenterology | ACG* 2022; 117:1279. https://doi.org/10.14309/ajg.0000000000001773

17. Anon. ImproveCareNow Purpose & Success. ImproveCareNow. https://www.improvecarenow.org/purpose-success. Accessed July 2, 2022

18. Turner D, Otley AR, Mack D, et al. Development, validation, and evaluation of a pediatric ulcerative colitis activity index: a prospective multicenter study. *Gastroenterology* 2007; 133:423–432. https://doi.org/10.1053/j.gastro.2007.05.029

19. Turner D, Ricciuto A, Lewis A, et al. STRIDE-II: An Update on the Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE) Initiative of the International Organization for the Study of IBD (IOIBD): Determining Therapeutic Goals for Treat-to-Target strategies in IBD. *Gastroenterology* 2021; 160:1570–1583. https://doi.org/10.1053/j.gastro.2020.12.031

20. Rudrapatna VA, Cheng Y-W, Feuille C, et al. Creation of a ustekinumab external control arm for Crohn's disease using electronic health records data: a pilot study. 2021:2021.11.12.21266064.

21. ImproveCareNow™. Program Details. ImproveCareNow. https://www.improvecarenow.org/program-details. Accessed October 16, 2023

22. Buuren S van, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software* 2011; 45:1–67. https://doi.org/10.18637/jss.v045.i03

23. Yanagi T, Ushijima K, Koga H, et al. Tacrolimus for ulcerative colitis in children: a multicenter survey in Japan. *Intest Res* 2019; 17:476–485. https://doi.org/10.5217/ir.2019.00027

24. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Academic Press; 2013.

25. Andrade C. Mean Difference, Standardized Mean Difference (SMD), and Their Use in Meta-Analysis: As Simple as It Gets. *J Clin Psychiatry* 2020; 81:20f13681. https://doi.org/10.4088/JCP.20f13681

26. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med* 2015; 34:3661–3679. https://doi.org/10.1002/sim.6607

27. Feagan BG, Sandborn WJ, Gasink C, et al. Ustekinumab as Induction and Maintenance Therapy for Crohn's Disease. *New England Journal of Medicine* 2016; 375:1946–1960. https://doi.org/10.1056/NEJMoa1602773

28. Feagan BG, Rubin DT, Danese S, et al. Efficacy of Vedolizumab Induction and Maintenance Therapy in Patients With Ulcerative Colitis, Regardless of Prior Exposure to Tumor Necrosis Factor Antagonists. *Clin Gastroenterol Hepatol* 2017; 15:229-239.e5. https://doi.org/10.1016/j.cgh.2016.08.044

29. Sands BE, Sandborn WJ, Van Assche G, et al. Vedolizumab as Induction and Maintenance Therapy for Crohn's Disease in Patients Naïve to or Who Have Failed Tumor Necrosis Factor Antagonist Therapy. *Inflamm Bowel Dis* 2017; 23:97–106. https://doi.org/10.1097/MIB.0000000000000979

30. Vermeire S, O'Byrne S, Keir M, et al. Etrolizumab as induction therapy for ulcerative colitis: a randomised, controlled, phase 2 trial. *Lancet* 2014; 384:309–318. https://doi.org/10.1016/S0140-6736(14)60661-9

31. Singh S, George J, Boland BS, et al. Primary Non-Response to Tumor Necrosis Factor Antagonists is Associated with Inferior Response to Second-line Biologics in Patients with Inflammatory Bowel Diseases: A Systematic Review and Meta-analysis. *J Crohns Colitis* 2018; 12:635–643. https://doi.org/10.1093/ecco-jcc/jjy004

32. Dhaliwal J, McKay HE, Deslandres C, et al. One-year outcomes with ustekinumab therapy in infliximab-refractory paediatric ulcerative colitis: a multicentre prospective study. *Aliment Pharmacol Ther* 2021; 53:1300–1308. https://doi.org/10.1111/apt.16388

33. Xu S, Ross C, Raebel MA, et al. Use of stabilized inverse propensity scores as weights to directly estimate relative risk and its confidence intervals. *Value Health* 2010; 13:273–277. https://doi.org/10.1111/j.1524-4733.2009.00671.x

34. Cohen RD, Bhayat F, Blake A, et al. The Safety Profile of Vedolizumab in Ulcerative Colitis and Crohn's Disease: 4 Years of Global Post-marketing Data. *J Crohns Colitis* 2020; 14:192–204. https://doi.org/10.1093/ecco-jcc/jjz137

35. Hazlewood GS, Pokharel G, Deardon R, et al. Patient preferences for maintenance therapy in Crohn's disease: A discrete-choice experiment. *PLoS One* 2020; 15:e0227635. https://doi.org/10.1371/journal.pone.0227635

36. Kaplan JL, Liu C, King EC, et al. Use, Durability, and Risks for Discontinuation of Initial and Subsequent Biologics in a Large Pediatric-Onset IBD Cohort. *J Pediatr Gastroenterol Nutr* 2023; 76:566–575. https://doi.org/10.1097/MPG.0000000000003734

37. Cai T, Lin T-C, Bond A, et al. The Association Between Arthralgia and Vedolizumab Using Natural Language Processing. *Inflamm Bowel Dis* 2018; 24:2242–2246. https://doi.org/10.1093/ibd/izy127

38. Silverman AL, Bhasuran B, Mosenia A, et al. Accurate, Robust, and Scalable Machine Abstraction of Mayo Endoscopic Subscores from Colonoscopy Reports. 2022:2022.06.19.22276606.

39. Gisbert JP, Chaparro M. Predictors of Primary Response to Biologic Treatment [Anti-TNF, Vedolizumab, and Ustekinumab] in Patients With Inflammatory Bowel Disease: From Basic Science to Clinical Practice. *J Crohns Colitis* 2020; 14:694–709. https://doi.org/10.1093/ecco-jcc/jjz195

40. Anyane-Yeboa A, Quezada S, Rubin DT, et al. The Impact of the Social Determinants of Health on Disparities in Inflammatory Bowel Disease. *Clinical Gastroenterology and Hepatology* 2022. https://doi.org/10.1016/j.cgh.2022.03.011

41. Bernstein CN, Walld R, Marrie RA. Social Determinants of Outcomes in Inflammatory Bowel Disease. *Am J Gastroenterol* 2020; 115:2036–2046. https://doi.org/10.14309/ajg.0000000000000794

42. Sands BE, Peyrin-Biroulet L, Loftus EV, et al. Vedolizumab versus Adalimumab for Moderate-to-Severe Ulcerative Colitis. *New England Journal of Medicine* 2019; 381:1215–1226. https://doi.org/10.1056/NEJMoa1905725

43. Sandborn WJ, Su C, Sands BE, et al. Tofacitinib as Induction and Maintenance Therapy for Ulcerative Colitis. *New England Journal of Medicine* 2017; 376:1723–1736. https://doi.org/10.1056/NEJMoa1606910

Supplemental Methods

Sensitivity Analyses:

We designed sensitivity analyses to address potential biases related to missing outcome data. To test the robustness of our findings to our choice of the complete case analysis method, we designed two sensitivity analyses, each corresponding to an alternative explanation for missing outcome data in real-world settings. (1) Non-responder imputation. This is a method commonly used in the analysis of IBD clinical trials [42, 43] and accounts for patients who may have been too ill to present to clinic. (2) Responder imputation. This corresponds to a situation where responders and symptomatically quiescent patients are less inclined to follow-up due to lack of clinical necessity.

	Vedolizumab	Ustekinumab	p-value	SMD[Ω]
N	262	74	-	-
Female	128 (48.9%)	33 (44.6%)	0.60	0.08
Age at diagnosis (years), median [IQR*]	13.5 [10.3-15.6]	13.1 [11.6-16.5]	0.25	0.15
Age at baseline visit (years), median [IQR]	16.2 [13.5-18.3]	16.6 [14.0-18.5]	0.38	0.12
Disease duration at baseline visit (years), median [IQR]	2.0 [1.0-3.7]	1.5 [0.8-4.1]	0.74	0.04
Paris Classification				
Extent (pancolitis)	186 (71.0%)	53 (71.6%)	0.41	0.06
Severity (S1)	124 (47.3%)	32 (43.2%)	0.24	0.18
Immunomodulators	56 (21.4%)	9 (12.2%)	0.13	0.23
Corticosteroids	112 (42.7%)	25 (33.8%)	0.20	0.17
Labs, median [IQR]				
Hematocrit (%)	37.9 [34.6-40.9]	37.3 [33.4-40.6]	0.66	0.06
C-reactive protein (mg/L)	1.8 [0.5-5.5]	2.2 [0.5-6.2]	0.79	0.03
Sedimentation rate (mm/h)	17.0 [9.0-29.0]	16.0 [9.0-26.0]	0.58	0.07
Albumin (g/dL)	4.0 [3.7-4.3]	4.1 [3.7-4.4]	0.78	0.04
[#]Growth (satisfactory)	225 (85.6%)	65 (87.8%)	0.32	0.13
[#]Nutrition (satisfactory)	192 (73.3%)	58 (78.4%)	0.48	0.14
^PUCAI, median [IQR]	26.8 [10-40]	20 [5-30]	0.03	0.29
[o]Extra-intestinal manifestations	< 10	<10	0.76	0.06

Table 1: Baseline characteristics after imputation

[Ω] *Standardized mean difference.*

**IQR: interquartile range.*

[#]*Pre-defined ImproveCareNow variables.*

^*PUCAI: Pediatric Ulcerative Colitis Activity Index.*

[o]*Extra-intestinal manifestations include arthritis, fevers, pyoderma gangrenosum, erythema nodosum and uveitis.*

Table 2: Results for corticosteroid-free clinical remission

Analysis Method	**Odds Ratio**	**95% confidence interval**	**p-value**
Unweighted cohort			
Complete case analysis	0.78	0.42-1.48	0.45
Impute missing as non-responder	0.80	0.43-1.48	0.48
Impute missing as responder	0.80	0.46-1.39	0.43
Propensity-score weighted cohort			
Complete case analysis	0.81	0.41-1.59	0.54
Impute missing as non-responder	0.94	0.50-1.78	0.85
Impute missing as responder	0.90	0.51-1.60	0.73

Table 3: Secondary Outcomes

Outcome Measure	Vedolizumab	Ustekinumab	p-value
Clinical Response	39.7% (n=232)	30.3% (n=66)	0.20
IBD-related hospitalizations	12.2% (n=262)	<10 (n=74)	0.74
IBD-related surgeries	<10 (n=262)	<10 (n=74)	0.43
Nutritional Status (satisfactory)	82.7% (n=168)	84.2% (n=38)	0.83
Growth status (satisfactory)	89.2% (n=251)	94.0% (n=67)	0.24
Infections	<10 (n=167)	<10 (n=38)	0.89
Malignancies	0% (n=232)	0% (n=66)	-

Table 3: The secondary endpoints of clinical remission, nutritional status, and growth status are assessed during the outcome time window of month-6 + 2 months. The secondary endpoints of hospitalizations, surgeries, malignancies, and infections were evaluated from baseline until the month-6 follow-up visit.

Figure 1: Time-to-event analyses for biologic durability

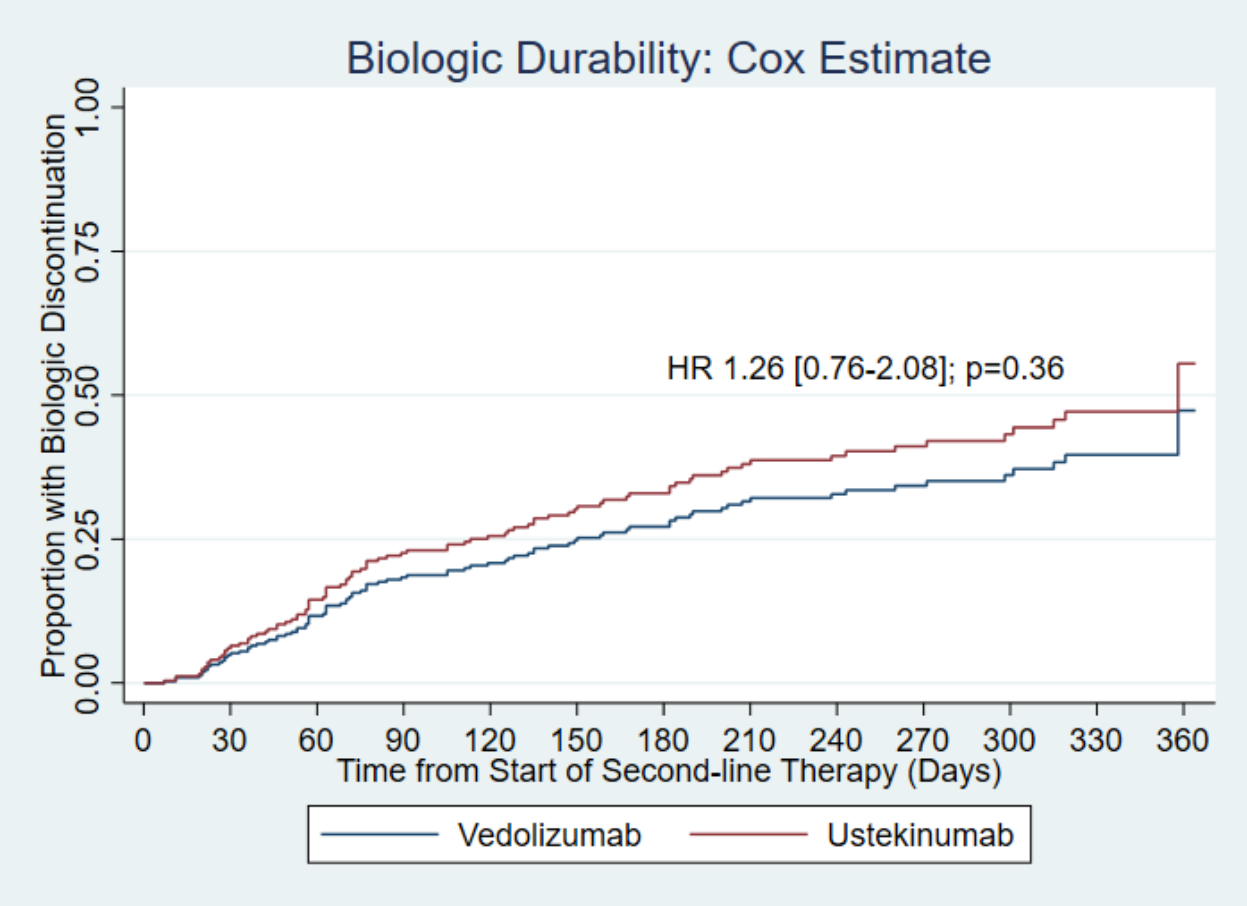


Figure 1: Cox estimate for time to discontinuation of the second-line biologic with a hazard ratio that accounts for control of all baseline covariates.

Supplemental Table 1: Baseline Characteristics prior to imputation

[Ω] *Standardized mean difference.*

**IQR: interquartile range.*

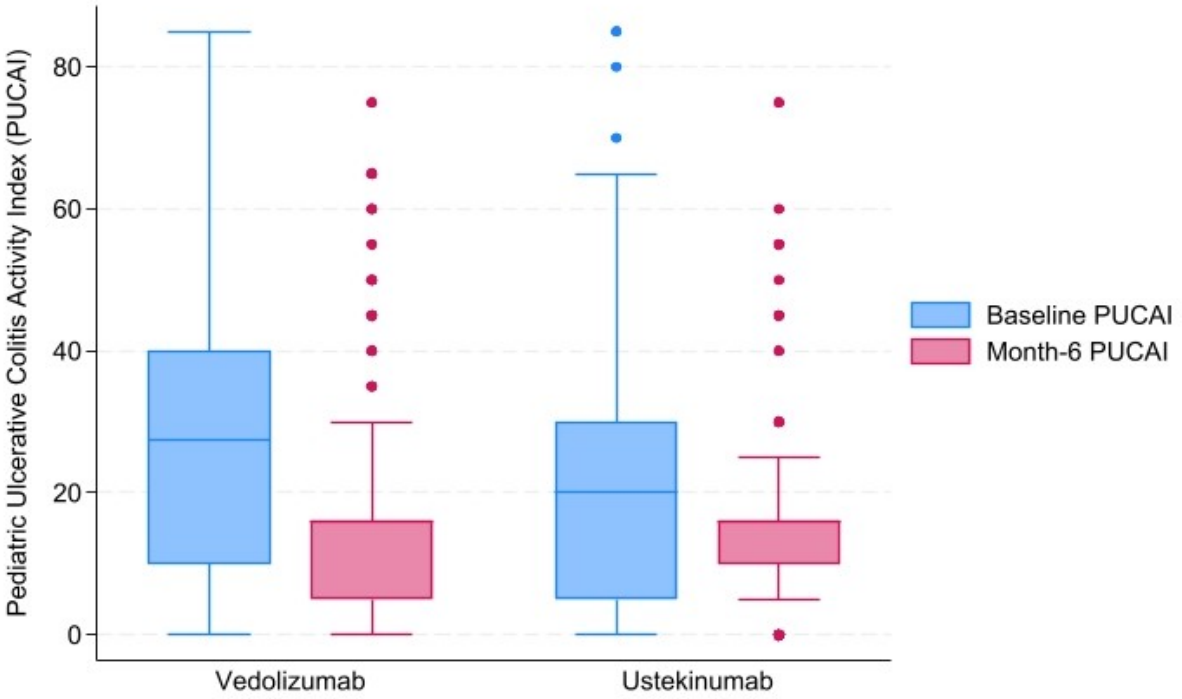
[#]*Pre-defined ImproveCareNow variables.*

^PUCAI: Pediatric Ulcerative Colitis Activity Index.

[o]*Extra-intestinal manifestations include arthritis, fevers, pyoderma gangrenosum, erythema nodosum and uveitis.*

	Vedolizumab	**Ustekinumab**	**p-value**	**SMD[Ω]**
N	262	74	-	-
Female	128 (48.9%) (n=262)	33 (44.6%) (n=74)	0.60	0.09
Age at diagnosis (years), median [IQR*]	13.9 [9.9-15.6] (n=243)	13.3 [11.3-16.8] (n=67)	0.20	0.17
Age at baseline visit (years), median [IQR]	16.4 [13.3-18.4] (n=262)	16.9 [14.2-18.5] (n=74)	0.33	0.13
Disease duration at baseline visit (years), median [IQR]	1.9 [1.0-3.7] (n=243)	1.5 [0.8-4.6] (n=67)	0.91	0.01
Paris Classification				
Extent (pancolitis)	184 (70.2%) (n=253)	50 (67.6%) (n=70)	0.75	0.07
Severity (S1)	124 (47.3%) (n=253)	31 (41.9%) (n=70)	0.38	0.18
Immunomodulators	56 (21.7%) (n=258)	9 (12.5%) (n=72)	0.11	0.23
Corticosteroids	112 (42.7%) (n=258)	25 (33.8%) (n=70)	0.22	0.17
Labs, median [IQR]				
Hematocrit (%)	38.0 [34.5-41] (n=247)	37.5 [33.4-40.7] (n=69)	0.73	0.05
C-reactive protein (mg/L)	1.1 [0.3-4.7] (n=219)	1.0 [0.3-4.1] (n=59)	0.85	0.03
Sedimentation rate (mm/h)	16.0 [8.0-29.0] (n=238)	14.5 [8.0-27.0] (n=68)	0.57	0.08
Albumin (g/dL)	4.1 [3.7-4.3] (n=246)	4.1 [3.7-4.4] (n=69)	0.67	0.06
[#]Growth (satisfactory)	221 (84.4%) (n=249)	63 (85.1%) (n=67)	0.44	0.12
[#]Nutrition (satisfactory)	192 (73.3%) (n=258)	58 (78.4%) (n=72)	0.84	0.13
^PUCAI, median [IQR]	25 [10-40] (n=246)	20 [5-30] (n=69)	0.03	0.30
[o]Extra-intestinal manifestations	<10 (n=254)	<10 (n=72)	0.76	0.04

Supplemental Figure 1



*Supplemental Figure 1: Box and whisker plot showing the changes in pediatric ulcerative colitis activity index (PUCAI) scores across time for each treatment group. The box spans the interquartile range (IQR), while the whiskers include values that differ from the upper and lower box bounds by less than 1.5*IQR.*