Supplementary Table 1: Comparison of Baseline Characteristics based on Clinical Response at Week 8 or 16 with Tofacitinib in Patients with Crohn's disease and IBD-U

		Clinical response at 8/16weeks		P values
		No (N=39)	Yes (N=34)	
Age at tofacitinib induction, median (IQR)		29 (23 – 36)	36 (25 - 47)	0.07
Age at diagnosis <16 yrs, n (%)		14 (35.9)	6 (17.7)	0.08
Male, n (%)		14 (35.9)	25 (73.5)	**0.002**
Race, n (%)	Non-Hispanic white	30 (76.9)	31 (91.2)	0.12
	Others	9(23.1)	3 (8.8)	
BMI, median (IQR)		22.7 (20.1 – 26.0)	25.1 (21.9 – 29.7)	**0.04**
CD duration in years, mean (IQR)		9 (5 - 14)	9 (4 - 17)	0.84
Smoking Status[†], n (%)	No smoking history	33 (84.6)	25 (78.1)	0.55
	Current or past smoker	6 (15.4)	7 (21.9)	
Montreal location, n (%)	Ileal or Ileocolonic	23 (59.0)	14 (41.2)	0.16
	Colonic	16 (41.0)	20 (58.8)	
Montreal Phenotype, n (%)	Non-stricturing/Non-penetrating	25 (64.1)	16 (67.7)	0.38
	Stricturing	8 (20.5)	3 (8.8)	
	Penetrating	6 (15.4)	8 (23.5)	
Any perianal disease, n (%)		10 (25.6)	12 (35.3)	0.45
Induction dosing, n (%)	5mg bid/11mg daily	13 (33.3)	6 (17.6)	
	10mg bid	26 (67.7)	28 (82.4)	0.18
Concurrent steroid use at start of tofacitinib, n (%)		19 (48.7)	15 (44.1)	0.81
Concurrent immunomodulator use at start of tofacitinib, n (%)		4 (10.3)	7 (20.6)	0.33
Baseline hemoglobin (g/dL) median (IQR), N=62		12.6 (11.2 -13.2)	13.2 (12.6-14.7)	**0.01**
Baseline albumin, median (IQR), N=52		4 (3.4 – 4.3)	3.9 (3.5 – 4.3)	0.50
Baseline CRP, median (IQR), n=42		7.45 (3 - 43)	6 (2.2 - 37)	0.42
No of biologic class prior to tofacitinib, n (%) *	1	18 (46.2)	18 (54.6)	0.64
	≥2	21 (53.8)	15 (45.4)	
Follow up in months, median (IQR)		7.6 (3.3 - 12.1)		
Total exposure, PYF		67.7		

Footnotes:

[†]Smoking status unknown in 2 patients

* One patient in the Yes group is biologic naïve and excluded.

Abbreviations: bid, Two times a day; BMI, body mass index; CD; Crohn's disease; IBD-U, Inflammatory bowel disease unclassified; IQR, Interquartile range; PYF, Patient years follow up

Supplementary Table 2: Adverse Events and surgical outcomes

Total number of patients with an adverse event, n (%)		**13 (17.1)**
Nature of adverse events	n	IR (/PYF)
Infections	1	1.6
Rash	5	7.8
Venous thrombo-embolic event (Portal vein thrombosis)	1	1.6
Anemia	1	1.6
Major cardiovascular event	-	-
GI perforation	-	-
Others: Thrombocytosis	1	1.6
Cough, headache, SOB/chest pain/ lightheadedness	3	4.7
Nausea	1	1.6
Total number of patients with a serious adverse event, n (%)		**6 (7.9)**
Nature of serious adverse events	n	IR (/PYF)
SAEs leading to discontinuation of tofacitinib, n		
Shortness of breath and chest pain	1	1.6
Cough, chest tightness, lightheadedness	1	1.6
Nausea	1	1.6
SAEs requiring hospitalization and discontinuation, n		
Anemia -hemolytic[†]	1	1.6
Other infection (unspecified)	1	1.6
SAEs leading to a potentially life-threatening event, n		
Deep vein thrombosis	1	1.6
Number of days to developing all AEs, median [IQR]	93 (19 – 329)	
Incidence rate of all AEs, per 100 PYF [95%CI]	20.4 (16.5 – 26.9)	
Incidence rate of SAEs, per 100 PYF [95%CI]	8.9 (7.2 -11.6)	
Number of patients who underwent IBD-related surgery, n (%)		**25 (32.9)**
Experienced as least one Clavien-Dindo grade complication (N=24), n (%)	12 (50)	
Experienced Clavien-Dindo Grade 3 or above complication (N=24), n (%)	2 (8.3)	
Surgical Site Infection within 30 days of surgery (N=24)	1	
Non-surgical/non-abdominal post-op infection (N=24)	2	
Re-operation within 30 days of surgery (N=22)	0	
Re-admission within 30 days of surgery (N=22)	4	

Footnotes: Abbreviations: AE, Adverse events; IR, Incidence rate; PYF, Patient years follow up; SAE, serious adverse events

Supplementary Methods:

This study is reported according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement for cohort studies. IRB approval for data collection and data sharing was obtained at each site to create the **T**ofacitinib **R**eal-world **O**utcomes in **P**atients with ulcerat**I**ve colitis and **C**rohn's disease (TROPIC) consortium.

Patients were included if they were receiving care at one of the sites and prescribed tofacitinib for a diagnosis of CD or IBD-U (defined by International Organization of Inflammatory Bowel Diseases). Data were collected retrospectively at each site between May 2018 and November 2019.

Data collection was performed using a standardized data collection form on REDCap (Research Electronic Data Capture, version 7.3.5: REDCap Consortium, Vanderbilt University, Nashville, TN, U.S.A.), using pre-specified definitions and criteria for coding. Assessment was done prior to tofacitinib initiation and during follow-up (+/- 4 weeks) at weeks 8, 16, 26, 39 and 52. A Charlson comorbidity score at start of therapy was calculated for each patient. At each time point, clinical response (based on PGA), corticosteroid usage, and need for dose escalation or de-escalation was recorded. Cardiac and thromboembolic events and their risk factors were noted. Laboratory parameters were recorded if obtained at baseline and at follow up timepoints. Information on CD-related hospitalizations, CD-related complications, and adverse events were recorded.

Adverse events (AEs) were defined as serious if life-threatening or resulting in a hospitalization, disability or discontinuation of therapy.

Effectiveness was determined using the PGA, with the primary effectiveness outcome being clinical response defined as >50% reduction in symptoms at weeks 8 and/or 16 during follow-up. For this analysis, patients who had achieved clinical response at week 8, but then lost response at week 16 were classified as having lost response. Additionally, if a patient achieved a response at week 8 and had no further follow-up at week 16 they were classified as having clinical response for the primary outcome analysis. Secondary effectiveness outcomes included clinical remission (resolution of symptoms), corticosteroid-free clinical remission, corticosteroid-free clinical response, endoscopic remission (resolution of ulceration), and change in CRP.

Descriptive statistics are presented as medians with interquartile range (IQR) for continuous variables and frequencies and percentages for categorical variables. Non-parametric continuous variables were compared using Mann-Whitney U or Kruskall-Wallis or Friedman test as appropriate, while categorical variables were compared using Pearson's chi-squared or Fisher's exact tests. All data were analyzed based on observed values, with no imputation for missing values.

A univariate logistic regression was used to assess demographic and clinical variables that were associated with clinical response. We then constructed a multivariate logistic regression models to identify the independent predictors of response as Odds Ratios (ORs). The threshold for statistical significance in the final model was set at 0.05 for all statistical tests and all p-values were two-sided. Incidence rates (IRs) were calculated based on the unique number of patients with events per 100 patient-years of exposure. Statistical analysis was performed using Stata version 16.0 (StataCorp, College Station, TX) and GraphPad Prism version 8.3.0 (GraphPad Software; La Jolla, California, USA).