



Modeling of Treatment Outcomes with Tofacitinib Maintenance Therapy in Patients with Ulcerative Colitis: A Post Hoc Analysis of Data from the OCTAVE Clinical Program

Michael Chiorean · Marco Daperno · Charlie W. Lees ·

Gianluca Bonfanti · Dimitrios Soudis · Irene Modesto ·

J. Jasper Deuring · Roger A. Edwards

Received: March 8, 2023 / Accepted: June 29, 2023 / Published online: July 31, 2023
© The Author(s) 2023

ABSTRACT

Introduction: Tofacitinib is an oral small molecule Janus kinase inhibitor for the treatment of ulcerative colitis (UC). This post hoc analysis assessed whether various statistical techniques could predict outcomes of tofacitinib maintenance therapy in patients with UC.

Methods: Data from patients who participated in a 52-week, phase III maintenance study (OCTAVE Sustain) and an open-label long-term extension study (OCTAVE Open) were included in this analysis. Patients received tofacitinib 5 or 10 mg twice daily (BID) or placebo (OCTAVE Sustain only). Logistic regression analyses were performed to generate models using clinical and laboratory variables to predict loss of responder

J. Jasper Deuring and Roger A. Edwards contributed equally to this work.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12325-023-02603-0>.

M. Chiorean
Swedish Medical Center, Seattle, WA, USA

M. Daperno
SC Gastroenterologia, AO Ordine Mauriziano di Torino, Turin, Italy

C. W. Lees
Institute of Genetics & Molecular Medicine,
University of Edinburgh, & IBD UNIT, Western
General Hospital, NHS Lothian, Edinburgh, UK

G. Bonfanti
Engineering Ingegneria Informatica, Milan, Italy

D. Soudis
Pfizer Hellas S.A., Thessaloniki, Greece

I. Modesto
Pfizer Inc, New York, NY, USA

J. J. Deuring
Pfizer Netherlands GmbH, Rotterdam, The
Netherlands

R. A. Edwards
Health Services Consulting Corporation,
Boxborough, MA, USA

J. J. Deuring (✉)
Pfizer Inc, Rivium Westlaan 142, 2909 LD Capelle
aan den IJssel, The Netherlands
e-mail: jasper.deuring@pfizer.com

status at week 8 of OCTAVE Sustain, steroid-free remission (defined as a partial Mayo score of 0–1 in the absence of corticosteroid use) at week 52 of OCTAVE Sustain, and delayed response at week 8 of OCTAVE Open. Furthermore, differences in loss of response/discontinuation patterns between treatment groups in OCTAVE Sustain were established.

Results: The generated prediction models demonstrated insufficient accuracy for determining loss of response at week 8, steroid-free remission at week 52 in OCTAVE Sustain, or delayed response in OCTAVE Open. Both tofacitinib doses demonstrated comparable response/remission patterns based on visualizations of disease activity over time. The rectal bleeding subscore was the primary determinant

of disease worsening (indicated by an increased total Mayo score), and the endoscopy subscore was the primary determinant of disease improvement (indicated by a decreased total Mayo score).

Conclusion: Visualizations of disease activity subscores revealed distinct patterns among patients with UC that had disease worsening and disease improvement. The statistical models assessed in this analysis could not accurately predict loss of responder status, steroid-free remission, or delayed response to tofacitinib. Possible reasons include the small sample size or missing data related to yet unknown key variables that were not collected during these trials.

Graphical Abstract:

Modeling of Treatment Outcomes with Tofacitinib Maintenance Therapy in Patients with Ulcerative Colitis: A Post Hoc Analysis of Data from the OCTAVE Clinical Program

Michael Chiorean, Marco Daperno, Charlie W. Lees, Gianluca Bonfanti, Dimitrios Soudis, Irene Modesto, J. Jasper Deuring, Roger A. Edwards

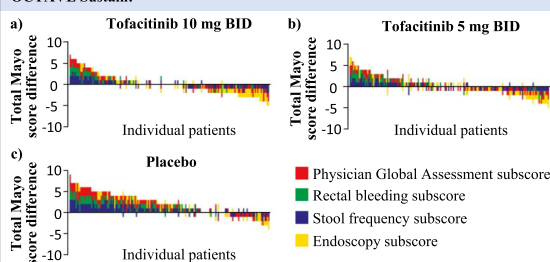
- Tofacitinib is an oral small molecule Janus kinase inhibitor for the treatment of ulcerative colitis (UC)
- This post hoc analysis assessed whether various statistical techniques could predict outcomes of tofacitinib maintenance therapy in patients with UC

Methods:

 Data from patients who participated in a 52-week, phase III maintenance study ($n = 592$; OCTAVE Sustain) and an OLE study ($n = 944$; OCTAVE Open) were included in the analysis

 Logistic regression analyses were performed to generate models using clinical and laboratory variables to predict loss of responder status at week 8 of OCTAVE Sustain, steroid-free remission at week 52 of OCTAVE Sustain, and delayed response at week 8 of OCTAVE Open

Visualization of change in total Mayo score from study baseline to week 52 of OCTAVE Sustain:



Analyzing the OCTAVE Sustain data has not resulted in a prediction model that meets the threshold for definitive prediction for:

Loss of response at week 8 in OCTAVE Sustain Steroid-free remission at week 52 in OCTAVE Sustain Delayed response at week 16 in OCTAVE Open

Disease activity scores:

The change in rectal bleeding subscore was the primary determinant of disease worsening (increased total Mayo score)



The change in endoscopy subscore was the primary determinant of disease improvement (decreased total Mayo score)



Conclusions: Visualizations of disease activity subscores revealed distinct patterns among patients with UC that had disease worsening and disease improvement. The statistical models assessed in this analysis could not accurately predict loss of responder status, steroid-free remission, or delayed response to tofacitinib

The graphical abstract represents the opinions of the authors. For a full list of declarations, including funding and author disclosure statements, and copyright information, please see the full text online.



PEER-REVIEWED
FEATURE

PLAIN LANGUAGE SUMMARY

Doctors use tofacitinib (Xeljanz[®]) to treat people with moderate to severe ulcerative colitis. Patients who respond to (have improved symptoms following) treatment with tofacitinib 10 mg twice a day for 8 weeks, or up to 16 weeks if they do not respond initially (known as induction treatment), can receive tofacitinib treatment at the lowest effective dose to sustain their response (called maintenance treatment). Predicting how patients respond to tofacitinib maintenance treatment may help clinicians work out the lowest effective dose for each patient. In this study, data from the tofacitinib clinical trials were used to assess the ability to predict maintenance therapy response or failure in patients with ulcerative colitis. Differences between patients who received tofacitinib 5 or 10 mg twice a day and who either stopped responding to treatment or stopped taking treatment were looked at. The study could not accurately predict which patients would experience disease worsening, steroid-free remission (very mild or no symptoms, and not taking steroids), or take longer to respond following tofacitinib maintenance treatment. Patterns of patients who had stopped responding to treatment, or stopped taking treatment, were similar between patients who received tofacitinib 5 or 10 mg twice daily. When reviewed using doctor- and patient-reported scores that measure ulcerative colitis disease activity, different factors were important in patients with disease worsening compared with disease improvement. The results suggest that further research is needed to more accurately predict how patients with ulcerative colitis will respond to tofacitinib maintenance treatment.

Keywords: Inflammatory bowel disease; Statistics; Ulcerative colitis

Key Summary Points

Why carry out this study?

Progress in understanding the mechanisms of ulcerative colitis (UC) immunopathogenesis has led to an increase in the number of treatments available, and choice of therapy for patients with UC is currently predicated on disease severity, treatment history, and benefit–risk profiles.

Various statistical techniques could aid the development of models that predict outcomes of treatments and assist decisions on benefits and risks, which is particularly important in the context of dosing because the lowest effective tofacitinib dose should be used for maintenance therapy in patients with UC.

This post hoc analysis assessed whether various statistical techniques could predict the outcomes of tofacitinib 5 or 10 mg twice daily (BID) maintenance therapy, in patients with UC who had previously achieved a response with tofacitinib 10 mg BID induction therapy.

What was learned from the study?

Prediction models demonstrated insufficient accuracy for determining loss of response at week 8 or steroid-free remission at week 52 in OCTAVE Sustain, or delayed response in OCTAVE Open, whereas rectal bleeding and endoscopy subscores were the primary determinants of disease worsening and improvement, respectively.

The prediction models and association analyses were unable to predict differences between tofacitinib 5 or 10 mg BID maintenance therapy; however, it is possible that data related to key variables not collected during the clinical trials included in this analysis could have significantly improved the accuracy of the models, and further studies with larger sample sizes are required.

DIGITAL FEATURES

This article is published with digital features, including a graphical abstract to facilitate understanding of the article. To view digital features for this article, go to <https://doi.org/10.6084/m9.figshare.23294930>.

INTRODUCTION

Ulcerative colitis (UC) is a chronic, immune-mediated disease, which is characterized by recurrent periods of relapsing and remitting symptoms, including diarrhea, fecal urgency, and rectal bleeding [1, 2]. Given the chronic nature of the disease, patients with UC require long-term therapy, with the ultimate goal of inducing and maintaining sustained steroid-free remission.

With progress in understanding the mechanisms behind UC immunopathogenesis, the number of treatment options for patients with UC has expanded [3–5]. The choice of therapy for patients with UC is predicated on disease severity and the benefit–risk profile of medical therapy [6]. The concept of personalized therapy is based on the premise of identifying specific clinical or laboratory variables capable of identifying patients who may respond to certain treatments. While personalized medicine for UC is not yet well developed, this could further streamline future therapeutic choices [7, 8].

Tofacitinib is an oral small molecule Janus kinase inhibitor for the treatment of UC. The efficacy and safety of tofacitinib 10 mg twice daily (BID) in patients with moderately to severely active UC were demonstrated in an 8-week phase II induction study (NCT00787202) [9], two 8-week phase III induction studies (OCTAVE Induction 1 and 2; NCT01465763 and NCT01458951) [10], and a 52-week phase III maintenance study (OCTAVE Sustain; NCT01458574) [10], and were evaluated further in an open-label long-term extension study (OCTAVE Open; NCT01470612) [11] and a phase IIIb/IV study (RIVETING; NCT03281304) [12].

The use of various statistical techniques can aid in the development of models to predict the outcomes of patients with UC treated with tofacitinib, which could assist healthcare professionals in decisions on benefits and risks including in the context of dosing. This is particularly important because the lowest effective tofacitinib dose should be used for maintenance therapy in patients with UC [13]. A previous analysis of data from patients with UC in OCTAVE Induction 1 and 2 who received 8 weeks of induction therapy with tofacitinib 10 mg BID demonstrated that early and clinically meaningful predictions of responder status were possible using these techniques [14].

The aim of this post hoc analysis was to assess whether similar techniques could be used to predict early and long-term outcomes of tofacitinib 5 or 10 mg BID as maintenance therapy in patients with UC who had previously responded to tofacitinib 10 mg BID induction therapy. Outcomes included loss of response and steroid-free remission at week 8 and week 52 of OCTAVE Sustain, respectively, as well as differences in loss of response/discontinuation patterns between treatment groups. This analysis also assessed delayed response in patients who did not respond to the initial 8-week induction therapy.

METHODS

Patients

Full details of the OCTAVE clinical program study design, and inclusion and exclusion criteria, have been described previously [10, 11]. Briefly, patients with UC who completed OCTAVE Induction 1 and 2 with a clinical response (defined as a decrease from induction study baseline total Mayo score of ≥ 3 points and $\geq 30\%$, plus a decrease in rectal bleeding subscore of ≥ 1 point or an absolute rectal bleeding subscore of 0 or 1) were eligible to participate in OCTAVE Sustain and were re-randomized to receive tofacitinib 5 mg BID, tofacitinib 10 mg BID, or placebo. Patients who did not respond in OCTAVE Induction 1 and 2 were eligible to participate in OCTAVE Open

and received tofacitinib 10 mg BID for at least another 8 weeks.

Compliance with Ethics Guidelines

The study protocols were approved by the Institutional Review Board or Independent Ethics Committee for each participating center. This post hoc analysis is based on the NCT01465763, NCT01458951, NCT01458574, and NCT01470612 studies. Written informed consent was obtained from all patients, per the ethics committee-approved protocols. All studies were conducted in compliance with the ethical principles derived from the Declaration of Helsinki and in compliance with all International Council for Harmonisation Good Clinical Practice Guidelines.

Outcomes and Models

Prediction Analyses

Two modeling techniques (logistic regression and Least Absolute Shrinkage and Selection Operator [LASSO] regression analyses) were used to generate a model that could predict the following outcomes: (1) loss of response (based on change in partial Mayo score [PMS]; Table S1 in the supplementary material) at week 8 of OCTAVE Sustain (week 16 overall) with tofacitinib 5 or 10 mg BID, (2) steroid-free remission (based on PMS) at week 52 of OCTAVE Sustain following treatment with tofacitinib 5 or 10 mg BID or placebo, and (3) delayed induction response in OCTAVE Open following non-response in OCTAVE Induction 1 and 2 and extended treatment with tofacitinib 10 mg BID for an additional 8 weeks (Fig. 1). Response (based on change in PMS) and steroid-free remission (based on PMS) have been defined previously (Table S1) [14, 16].

This analysis only included those patients who received active treatment (tofacitinib 10 mg BID) in the induction studies and were re-randomized in OCTAVE Sustain or proceeded to OCTAVE Open. A total of 593 patients from OCTAVE Induction 1 and 2 entered OCTAVE Sustain. Of these patients, 70 entered OCTAVE Sustain after responding to placebo

and were subsequently excluded from this analysis. Of the patients who received active treatment in the induction studies, 176 received tofacitinib 5 mg BID, 173 received tofacitinib 10 mg BID, and 174 received placebo in OCTAVE Sustain.

Two modeling strategies were utilized in these analyses, an empirical model and a hypothesis-based model. The empirical model included separate analyses that used variables from baseline to week 8 of OCTAVE Induction 1 and 2 and variables that were only available at baseline of OCTAVE Induction 1 and 2. The hypothesis-based model included variables that were selected on the basis of information extracted from the current literature and the expert opinions of authors. A hypothesis-based model based on total Mayo score and PMS, individual subscores, medication use, laboratory variables, and medical history was included among other analyses in order to evaluate the contribution of each score towards the precision of the predictive models. The full lists of variables that were entered into the models are shown in Tables S2–S4 in the supplementary material and included PMS and subscores, age, body mass index, disease duration, prior medication use (including corticosteroids), vital signs at baseline, and laboratory variables at baseline or week 8.

Prediction analyses were performed using the Mayo score [16]. Analyses performed using the empirical model included only the PMS, and analyses performed using the hypothesis-based model included both the PMS and total Mayo score, as defined in Table S1.

Logistic Regression

Logistic regression models were performed for each treatment group either with no selection method or using forward selection. Starting from an intercept-only model, one significant variable was added until no significant variables met the significance level to enter the model. An area under the receiver operating characteristic (AUROC) value of greater than 0.9 was considered predictive of response [17, 18].

Data were analyzed using three approaches. The first approach included the whole dataset of patients in the logistic regression model; the

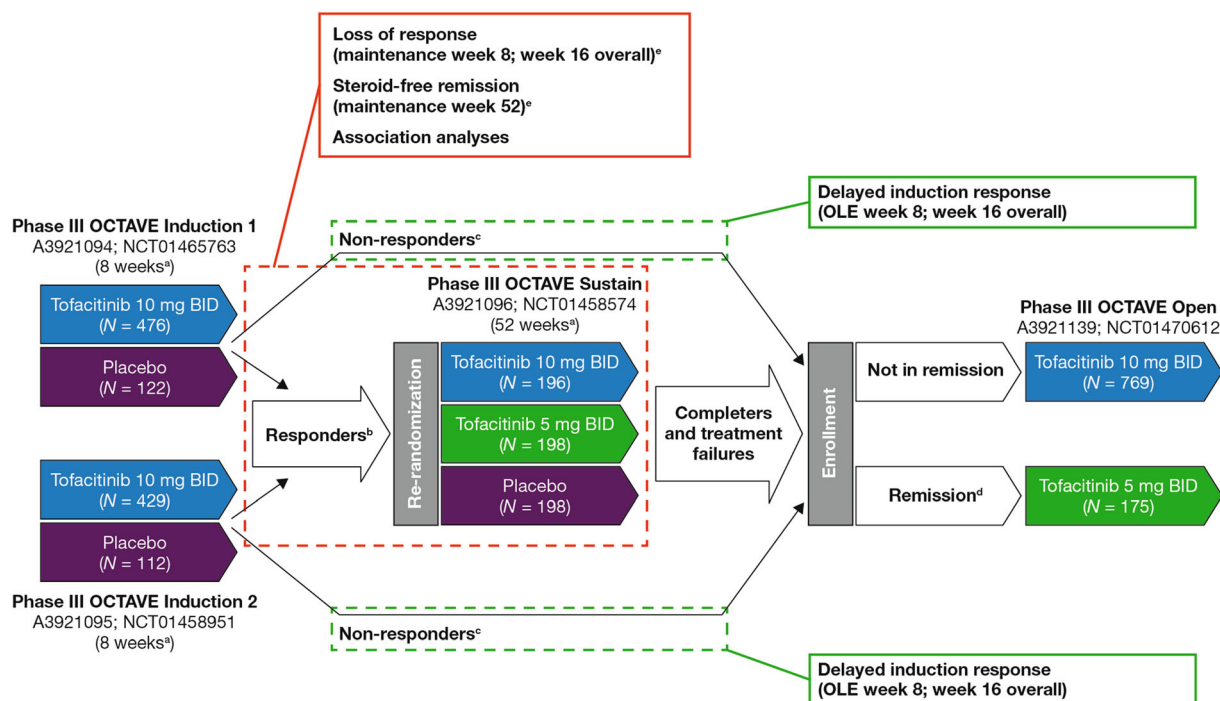


Fig. 1 Overview of the tofacitinib OCTAVE UC clinical program, and the outcomes of interest. ^aFinal complete efficacy assessment at week 8/52. Treatment continued up to week 9/53. ^bClinical response in OCTAVE Induction 1 and 2 was defined as a decrease from induction study baseline total Mayo score of ≥ 3 points and $\geq 30\%$, plus a decrease in rectal bleeding subscore of ≥ 1 point or an absolute rectal bleeding subscore of 0 or 1. ^cNon-responders from OCTAVE Induction 1 and 2 received tofacitinib 10 mg BID in OCTAVE Open. ^dRemission was defined as a total Mayo score of ≤ 2 with no

individual subscore > 1 and a rectal bleeding subscore of 0. ^eOnly patients receiving tofacitinib 5 or 10 mg BID were included in the analysis. *BID* twice daily, *OLE* open-label, long-term extension, *UC* ulcerative colitis. Figure has been adapted from Sandborn et al. [11] and Winthrop et al. [15]. These are both open access articles under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations were made

second approach used a 70:30 training–testing split for the identification of significant explanatory variables; and the third approach used *k*-fold cross-validation. For *k*-fold cross-validation, the whole dataset was randomly split into *k* different partitions (*k* = 5 or 10), and one partition at a time was iteratively considered. The remaining *k* – 1 partitions were used to predict outcomes for the patients in the omitted partition. Five repetitions of the whole cross-validation approach were performed for each *k*. Model performance was evaluated through the average AUROC value over the repetitions.

Least Absolute Shrinkage and Selection Operator Logistic Regression

LASSO logistic regression analyses, used to shrink or regularize coefficients to avoid overfitting [19, 20], were carried out in a similar way to the logistic regression modeling described above using the model with the smallest Akaike's information criterion instead of forward selection. Inclusion of the whole dataset of patients did not change the accuracy of the predictive response. Consequently, the LASSO logistic regression analyses reported here only utilized the 70:30 training–testing and *k*-fold cross-validation approaches.

All modeling analyses were performed using SAS 9.4 software.

Association Analyses

Association analyses were carried out to establish differences in on-response/discontinuation patterns between treatment groups. Bivariate analyses were performed to compare discontinuation as a result of insufficient response in the tofacitinib 5 and 10 mg BID groups using chi-square and Fisher's exact tests for categorical variables and the Kruskal–Wallis test for continuous variables. A *k*-nearest neighbor (kNN) analysis was performed to evaluate whether patients who received tofacitinib 5 mg BID and had loss of response in OCTAVE Sustain were more closely associated with patients who received tofacitinib 10 mg BID and had either loss of response or sustained response. Visualizations of individual response and remission patterns were performed using patient-level data with the aim of identifying the treatment patterns over time for each patient group. Three parameters were assessed: differences in the discontinuation and remission patterns among the treatment groups; 2- and 3-point PMS responder status between study baseline and week 52 of OCTAVE Sustain; and change in total Mayo score over time between study baseline and week 52 of OCTAVE Sustain.

In the analyses of discontinuation patterns, to determine *p* values, logistic regression was used to test if the average probability of discontinuation differed between treatment groups. Dummy variables were created to represent whether a patient had ceased participating in the study at any point. The dummy discontinuation variables were used as the dependent variables, and the treatment groups (tofacitinib 5 and 10 mg BID, and placebo) were used as the independent variable. In the analyses of change in total Mayo score over time, the proportion of total Mayo score comprising each subscore was calculated for all patients. Dummy variables were created to represent whether a patient had either an increase or decrease in total Mayo score over time. Regression models were run using the subscores as the dependent

variables, and the dummy variables as the independent variables. This was repeated for each treatment group. The statistical analyses were performed using R 4.1 software.

RESULTS

Loss of Responder Status at Week 8 of OCTAVE Sustain

Empirical logistic regression analyses that included variables from baseline resulted in AUROC values ranging from 0.50 to 0.79 to predict loss of response using all three modeling approaches. The equivalent values from the analyses that included variables from baseline to week 8 ranged from 0.61 to 0.86 (Table 1). For the models that could, to a certain extent, discriminate between the groups of interest that were assessed (i.e., tofacitinib 10 or 5 mg BID; AUROC > 0.6), the relative importance of the variables in the individual models is shown in Table S5 in the supplementary material.

The AUROC values were close to 0.9 (range 0.81–0.86) for the analysis that included the whole set of patients and used variables from baseline to week 8 (Table 1). Results of the hypothesis-based logistic regression analyses of loss of responder status were generally similar to the empirical analysis results (Table 2). The relative importance of each variable was not provided for the poorly performing hypothesis-based model because none of the variables selected by experts could discriminate between continued response and loss of response. In addition, results were generally similar with tofacitinib 5 and 10 mg BID (Tables 1 and 2).

Steroid-Free Remission at Week 52 of OCTAVE Sustain

Empirical logistic regression analyses that included variables that were selected from baseline only resulted in AUROC values that ranged from 0.65 to 0.66 for predicting steroid-free remission (based on PMS) when the treatments were combined and using the training–testing split and fivefold cross-validation

Table 1 Summary of AUROC values from logistic regression analyses of loss of responder status at week 8 of OCTAVE Sustain

	Whole set of patients		Testing set of training–testing split		Fivefold cross-validation	
	AUROC value (baseline only)	AUROC value (baseline to week 8)	AUROC value (baseline only)	AUROC value (baseline to week 8)	AUROC value (baseline only)	AUROC value (baseline to week 8)
2-point PMS loss of responder status						
Tofacitinib 10 mg BID	0.72	0.83	0.56	0.70	0.50	0.63
Tofacitinib 5 mg BID	0.73	0.84	0.50	0.73	0.56	0.65
3-point PMS loss of responder status						
Tofacitinib 10 mg BID	0.79	0.86	0.63	0.77	0.54	0.61
Tofacitinib 5 mg BID	0.77	0.81	0.44	0.68	0.49	0.62

LASSO logistic regression analyses did not improve the predictive power of the model (Table S6)

AUROC area under the receiver operating characteristic, *BID* twice daily, *LASSO* Least Absolute Shrinkage and Selection Operator, *PMS* partial Mayo score

models (Table 3). The equivalent values from the analyses that included variables from baseline to week 8 ranged from 0.73 to 0.79 when the treatments were combined (Table 3). Applying variable selection methods to improve the efficiency of the models did not increase the accuracy of the models. Results of the hypothesis-based logistic regression analyses were generally similar to the empirical analysis results (Tables 3 and 4; Table S8 in the supplementary material). For the models that could, to a certain extent, discriminate between the groups of interest that were assessed (i.e., tofacitinib 10 or 5 mg BID or placebo; AUROC > 0.6), the relative importance of the variables in the individual models is shown in Tables S9 and S10 in the supplementary material. In addition, results were generally similar across the two treatment groups (Tables 3 and 4; Table S8). Applying a tenfold cross-validation approach, and the addition of LASSO logistic regression analyses, did not improve the

predictive power of the model (Tables S11–14 in the supplementary material).

Delayed Response at Week 8 of OCTAVE Open

In total, 295 tofacitinib-treated patients who were non-responders in OCTAVE Induction 1 and 2 enrolled in OCTAVE Open as delayed responders. These patients represent only a portion of the patients who entered OCTAVE Open. As per Fig. 1, patients who had completed or had treatment failure during OCTAVE Sustain could also enter OCTAVE Open. Therefore, in total, 944 patients entered OCTAVE Open and received either tofacitinib 5 or 10 mg BID per protocol.

In the empirical logistic regression analyses of delayed response among patients who received tofacitinib 10 mg BID for an additional 8 weeks following induction therapy using all three modeling approaches, AUROC values, based on the PMS and individual partial Mayo

Table 2 Summary of AUROC values from the hypothesis-based logistic regression analyses of loss of responder status at week 8 of OCTAVE Sustain

	Whole set of patients				Testing set of training–testing split				Fivefold cross-validation			
	AUROC value (baseline only)		AUROC value (baseline to week 8)		AUROC value (baseline only)		AUROC value (baseline to week 8)		AUROC value (baseline only)		AUROC value (baseline to week 8)	
	PMS	Total Mayo score	PMS	Total Mayo score	PMS	Total Mayo score	PMS	Total Mayo score	PMS	Total Mayo score	PMS	Total Mayo score
2-point PMS loss of responder status												
Tofacitinib 10 mg BID	0.50	0.50	0.63	0.61	Global null	Global null	0.48	0.48	0.46	0.46	0.59	0.59
Tofacitinib 5 mg BID	0.60	0.60	0.70	0.70	0.62	0.62	0.61	0.61	0.53	0.53	0.53	0.53
3-point PMS loss of responder status												
Tofacitinib 10 mg BID	0.50	0.50	0.64	0.65	Global null	Global null	0.52	0.52	0.40	0.40	0.51	0.50
Tofacitinib 5 mg BID	0.58	0.58	0.66	0.65	Global null	Global null	0.57	0.57	0.49	0.49	0.44	0.45

“Global null” is defined as meaning that at least one of the statistical tests on the global null hypothesis did not reject the null hypothesis that all regression coefficients were equal to 0; consequently, the results obtained were not statistically significant (i.e., $p > 0.05$), and LASSO logistic regression analyses did not improve the predictive power of the model (Table S7 in the supplementary material)

AUROC area under the receiver operating characteristic, BID twice daily, LASSO Least Absolute Shrinkage and Selection Operator, PMS partial Mayo score

subscores, were less than 0.80 in the models that included variables from baseline only (PMS, range 0.52–0.63; individual subscores, range 0.52–0.66) or variables from baseline to week 8 (PMS, range 0.55–0.76; individual subscores, range 0.58–0.75; Table 5).

The results of the hypothesis-based logistic regression analyses were generally similar to those of the empirical analysis (Table 6). For the models that could, to a certain extent, discriminate between the groups of interest that were assessed (i.e., tofacitinib 10 mg BID; AUROC > 0.6), the relative importance of the variables in the individual models is shown in Table S15 in the supplementary material. In addition, the findings from the analyses based on PMS and total Mayo score were generally similar

(Table 6). Applying a tenfold cross-validation approach did not improve the predictive power of the model (Tables S16 and S17 in the supplementary material).

Association Analyses in OCTAVE Sustain

In the bivariate analysis, the following categorical variables were identified as significant: sex ($p = 0.005$), presence of extraintestinal manifestations at induction study baseline ($p = 0.027$), immunosuppressant use prior to induction ($p = 0.003$), and a history of musculoskeletal and connective tissue disorders ($p = 0.037$). In the analysis of continuous variables, only uric acid at week 8 of the induction

Table 3 Summary of AUROC values from logistic regression analyses predicting steroid-free remission^a at week 52 of OCTAVE Sustain without applying variable selection

	Testing set of training–testing split		Fivefold cross-validation	
	AUROC value (baseline only)	AUROC value (baseline to week 8)	AUROC value (baseline only)	AUROC value (baseline to week 8)
Tofacitinib 10 mg BID	0.70	0.75	0.56	0.67
Tofacitinib 5 mg BID	0.69	0.71	0.46	0.65
Placebo	0.55	0.68	0.48	0.53
Treatments combined	0.66	0.79	0.65	0.73

AUROC area under the receiver operating characteristic, *BID* twice daily, *PMS* partial Mayo score

^aSteroid-free remission was defined as a PMS of 0–1 in the absence of corticosteroids

Table 4 Summary of AUROC values from hypothesis-based logistic regression analyses predicting steroid-free remission^a at week 52 of OCTAVE Sustain applying variable selection

	Testing set of training–testing split				Fivefold cross-validation			
	AUROC value (baseline only)		AUROC value (baseline to week 8)		AUROC value (baseline only)		AUROC value (baseline to week 8)	
	PMS	Total Mayo score	PMS	Total Mayo score	PMS	Total Mayo score	PMS	Total Mayo score
Tofacitinib 10 mg BID	0.50	0.50	0.77	0.78	0.58	0.58	0.72	0.71
Tofacitinib 5 mg BID	0.50	0.50	0.73	0.69	0.45	0.45	0.66	0.62
Placebo	0.50	0.50	0.66	0.66	0.42	0.42	0.61	0.64
Treatments combined	0.67	0.67	0.78	0.79	0.64	0.66	0.74	0.73

AUROC area under the receiver operating characteristic, *BID* twice daily, *PMS* partial Mayo score

^aSteroid-free remission was defined as a PMS of 0–1 in the absence of corticosteroids

studies ($p = 0.042$) was identified as significant. Consecutive multivariable analyses were performed; however, no significant variables were identified (data not shown).

Using the significant variables extracted from the bivariate analysis, the kNN analysis

determined that patients with loss of response in the tofacitinib 5 mg BID group were more closely associated with patients who received tofacitinib 10 mg BID and had sustained response, compared with patients who had loss

Table 5 Summary of AUROC values from logistic regression analyses predicting delayed response at week 8 of OCTAVE Open

	Whole set of patients		Testing set of training–testing split		Fivefold cross-validation	
	AUROC value (baseline only)	AUROC value (baseline to week 8)	AUROC value (baseline only)	AUROC value (baseline to week 8)	AUROC value (baseline only)	AUROC value (baseline to week 8)
PMS						
Measured values	0.63	0.76	0.52	0.55	0.52	0.61
Differences	–	0.71	–	0.57	–	0.63
Individual partial Mayo subscores						
Measured values	0.66	0.75	0.54	0.58	0.52	0.63
Differences	–	0.71	–	0.49	–	0.58

AUROC area under the receiver operating characteristic, *PMS* partial Mayo score

Table 6 Summary of AUROC values from hypothesis-based logistic regression analyses predicting delayed response at week 8 of OCTAVE Open and applying variable selection

	Whole set of patients		Testing set of training–testing set		Fivefold cross-validation	
	AUROC value (baseline only)	AUROC value (baseline to week 8)	AUROC value (baseline only)	AUROC value (baseline to week 8)	AUROC value (baseline only)	AUROC value (baseline to week 8)
PMS and subscores ^a	0.57	0.69	0.42	0.59	0.47	0.63
Total Mayo score and subscores ^b	0.56	0.69	0.42	0.59	0.46	0.65

AUROC area under the receiver operating characteristic, *PMS* partial Mayo score

^aStool frequency, rectal bleeding, and Physician Global Assessment

^bStool frequency, rectal bleeding, endoscopic appearance, and Physician Global Assessment

of response in the tofacitinib 10 mg BID group (Table 7).

Disease activity graphs were generated to evaluate the patterns of response for individual patients in each treatment group. Visualization of data from OCTAVE Sustain demonstrated

that the patterns of remission, lack of remission, and discontinuation (based on OCTAVE Open study criteria) in patients receiving tofacitinib 5 vs 10 mg BID were generally similar over time ($p > 0.05$; Fig. 2). Moreover, patients with prior tumor necrosis factor inhibitor (TNFi) exposure

Table 7 Summary of results from k -nearest neighbor analysis evaluating loss of response in patients who received tofacitinib 5 mg BID in OCTAVE Sustain, using patients who received tofacitinib 10 mg BID as neighbors

	Number of neighbors (k)							
	$k = 1$	$k = 3$	$k = 5$	$k = 7$	$k = 9$	$k = 11$	$k = 13$	$k = 15$
2-point PMS loss of responder status								
Loss of response, n (%)	21 (30.9)	21 (30.9)	42 (61.8)	30 (44.1)	24 (35.3)	18 (26.5)	28 (41.2)	23 (33.8)
Response, n (%)	47 (69.1)	47 (69.1)	26 (38.2)	38 (55.9)	44 (64.7)	50 (73.5)	40 (58.8)	45 (66.2)
3-point PMS loss of responder status								
Loss of response, n (%)	22 (32.8)	21 (31.3)	41 (61.2)	31 (46.3)	26 (38.8)	22 (32.8)	26 (38.8)	24 (35.8)
Response, n (%)	45 (67.2)	46 (68.7)	26 (38.8)	36 (53.7)	41 (61.2)	45 (67.2)	41 (61.2)	43 (64.2)

BID twice daily, n number of patients who received tofacitinib 5 mg BID and had loss of response who were more closely associated with patients who received tofacitinib 10 mg BID and had either loss of response or sustained response, *PMS* partial Mayo score

demonstrated numerically higher rates of lack of remission and discontinuation compared with patients who were TNFi-naïve; however, these differences were not statistically significant. Patients treated with placebo demonstrated a significantly different pattern of disease activity compared with patients who were treated with tofacitinib 5 or 10 mg BID ($p < 0.05$).

Visualization patterns of response, lack of response, and discontinuation were generally similar in analyses that defined response as changes in PMS of 2 points vs 3 points (Figs. S1 and S2, respectively, in the supplementary material).

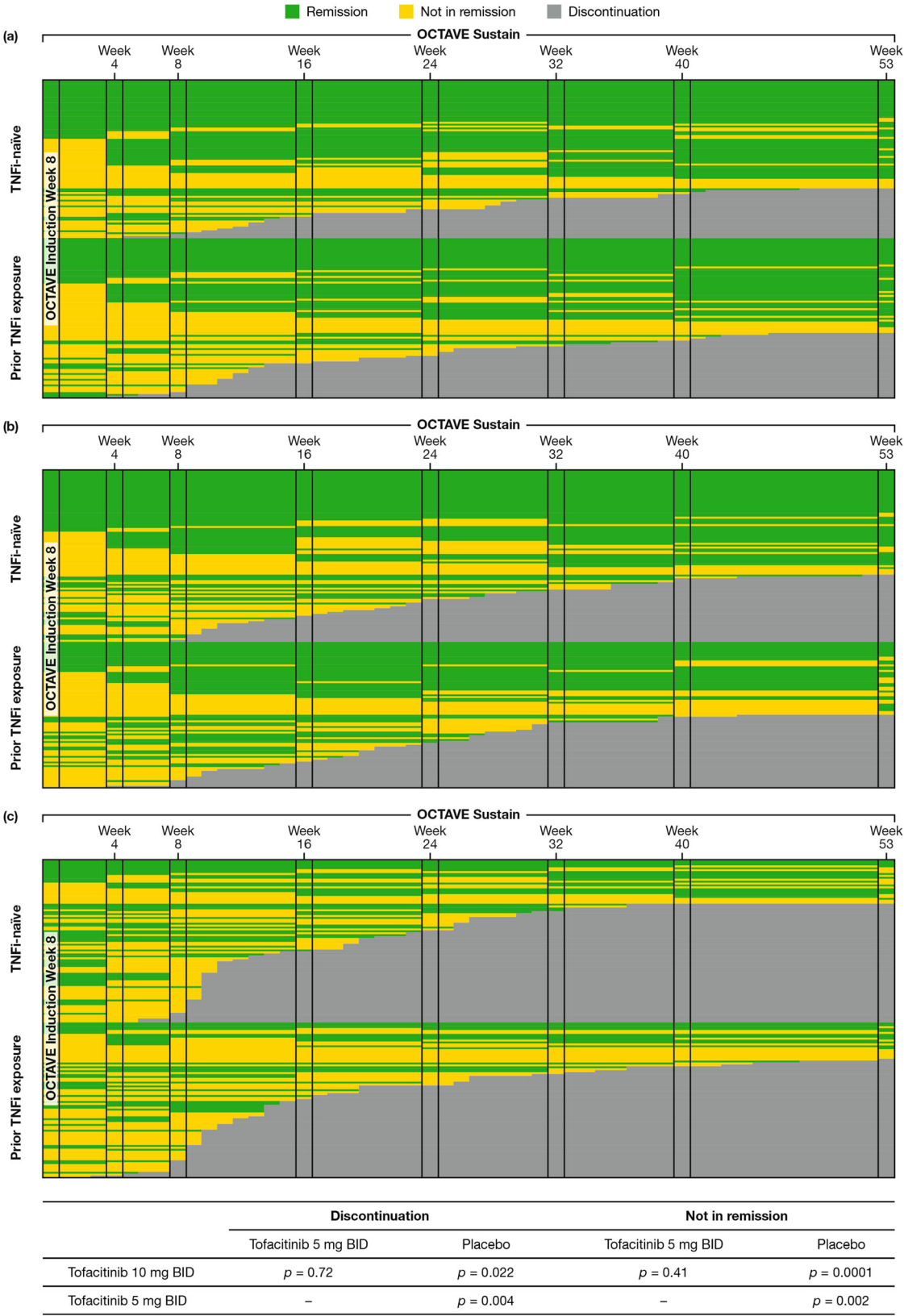
Interestingly, in the analysis of the contribution of individual subscores to the change in total Mayo score over time, the rectal bleeding subscore was the predominant subscore in patients with disease worsening (indicated by an increase in total Mayo score) between baseline and week 52 of OCTAVE Sustain (contribution of the rectal bleeding subscore was 24.7% [$p < 0.001$], 19.6% [$p < 0.001$], and 25.2% [$p < 0.001$] higher for patients with disease worsening compared with patients with disease improvement in those receiving tofacitinib 5 mg BID, tofacitinib 10 mg BID, and placebo, respectively; Fig. 3). The difference in the proportion of the total Mayo score made up by the rectal bleeding subscore was not statistically

significant when the three treatment groups were compared.

By contrast, the endoscopic subscore was the predominant subscore in patients who had disease improvement, indicated by a decrease in the total Mayo score between baseline and week 52 of OCTAVE Sustain in all three treatment groups (contribution of the endoscopic subscore was 15.9% [$p = 0.021$], 14.5% [$p = 0.046$], and 12.4% [$p = 0.096$] higher for patients with disease improvement compared with patients with disease worsening in those receiving tofacitinib 5 mg BID, tofacitinib 10 mg BID, and placebo, respectively; Fig. 3). The difference in the proportion of the total Mayo score made up by the endoscopic subscore was not statistically significant when the three treatment groups were compared.

DISCUSSION

This post hoc analysis of data from patients with UC in OCTAVE Sustain and OCTAVE Open utilized statistical modeling to predict patient loss of responder status (based on change in PMS) at week 8 of OCTAVE Sustain, steroid-free remission (based on PMS) at week 52 of OCTAVE Sustain, and delayed response at week 8 of OCTAVE Open. In all analyses, AUROC values were below the threshold for accurate prediction of response (AUROC



◀ **Fig. 2** Visualization of remission rates over time stratified by TNFi exposure among patients in OCTAVE Sustain receiving **a** tofacitinib 10 mg BID, **b** tofacitinib 5 mg BID, or **c** placebo. Remission was defined as having a PMS ≤ 1 . BID twice daily, PMS partial Mayo score, TNFi tumor necrosis factor inhibitor

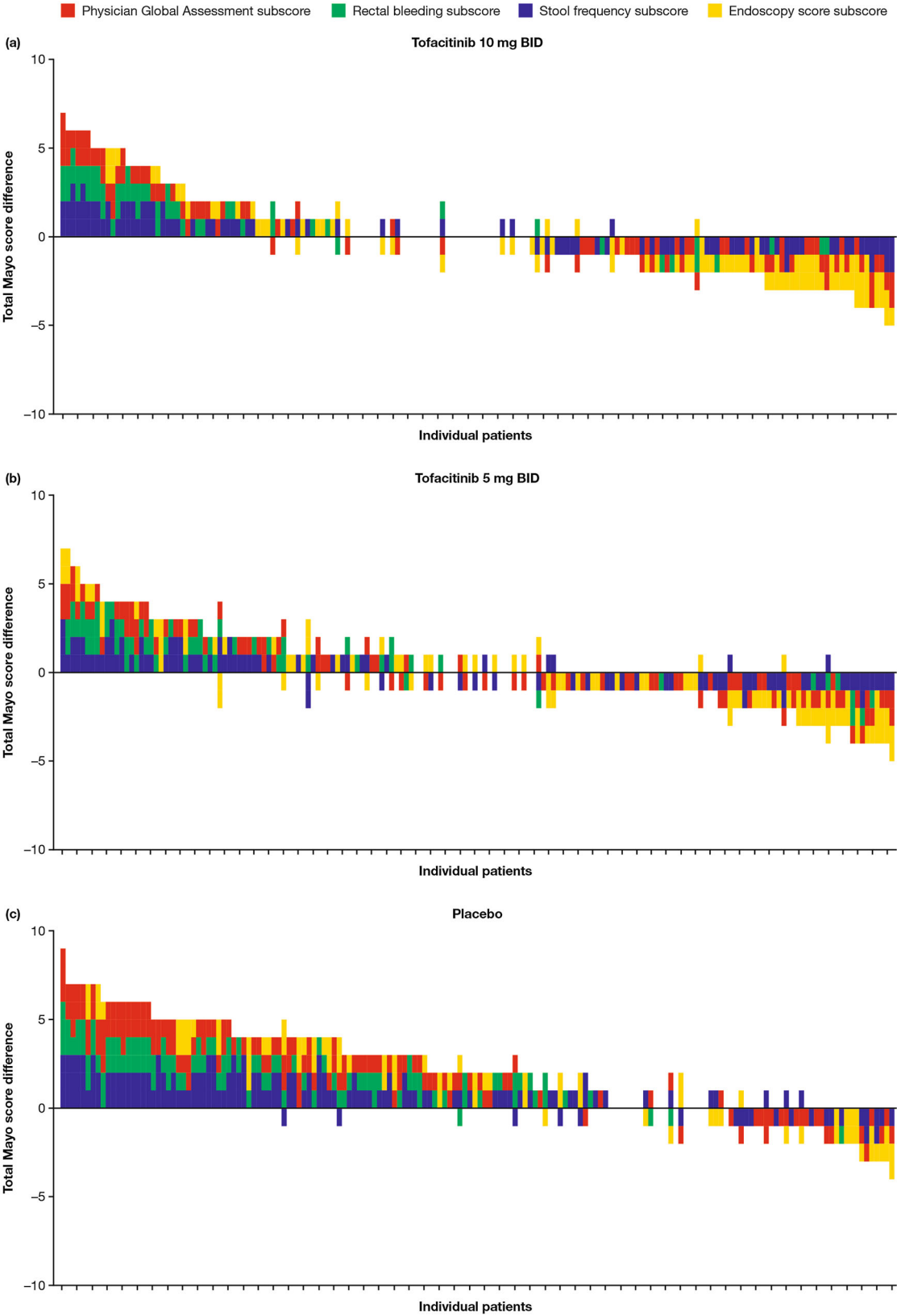
value < 0.9), and there were no notable differences in model outcomes when 2- and 3-point changes in PMS were used as definitions of response. The AUROC values were moderately acceptable at discriminating responders vs non-responders (i.e., AUROC > 0.8) in the prediction modeling analyses that used the whole dataset, including data from baseline to week 8 of OCTAVE Induction 1 and 2. However, models validated using the training–testing split or the cross-validation datasets generally had AUROC values below the threshold for moderate to accurate prediction (AUROC values < 0.8) [17, 18]. Comparisons of the two active treatment arms of OCTAVE Sustain (tofacitinib 5 and 10 mg BID) demonstrated that patient characteristics associated with responses to treatment were generally similar when evaluated using the variables that were available in this clinical trial. However, the variables that were identified as important within a particular model should be interpreted with caution. As the overall performance of the models was rather poor, the variables that were driving the outcomes could also not be considered specific enough to discriminate between the tofacitinib 5 and 10 mg BID groups. However, the variables listed such as age, PMS, and corticosteroid use have been previously described as factors associated with UC disease activity [14, 21, 22]. Despite the multiple efforts that were made to refine the research methodology by applying advanced prediction modeling methods and multi-omic approaches, which allows for more robust prediction by combining individual components and sharpening the approach to disease prediction, further refinement is suggested [23–25].

Notably, among patients with an increase in total Mayo score from study baseline to week 52 of OCTAVE Sustain, the rectal bleeding subscore was identified as the primary determinant of

disease worsening. By contrast, among patients with a decrease in total Mayo score from study baseline to week 52 of OCTAVE Sustain, the endoscopic subscore was identified as the primary determinant of improvement.

A previous assessment of data from OCTAVE Induction 1 and 2 used statistical modeling techniques to identify predictors of early treatment response among patients receiving induction therapy with tofacitinib 10 mg BID [14]. This study reported that logistic regression and random forest models were able to predict week 4 and week 8 responder status using only four measures (PMS, partial Mayo subscores, total cholesterol, and C-reactive protein) at different timepoints. An early change in the PMS was identified as the most important predictor of responder status. Here, we explored the extent to which treatment outcomes after week 8 of OCTAVE Induction 1 and 2 could be predicted on the basis of a change in PMS, as well as steroid-free remission (based on PMS) or delayed response among patients with UC. Patients included in this analysis were receiving either tofacitinib 5 or 10 mg BID as maintenance therapy or placebo in OCTAVE Sustain, or were non-responders at the end of OCTAVE Induction 1 and 2 and received extended induction for an additional 8 weeks in OCTAVE Open. There is no unambiguous reason why the models in this analysis were unable to accurately predict treatment response; however, factors such as the inclusion/exclusion criteria for the patient population in the global OCTAVE clinical program, number of patients analyzed, study design (i.e., omission of variables that could have been more predictive), the subjective nature of patient-reported symptoms in the PMS, and flare severity may have influenced the models' predictive capabilities.

Visualization of remission and discontinuation patterns across treatment groups demonstrated that disease activity patterns were generally similar in patients receiving tofacitinib 5 or 10 mg BID among TNFi-naïve patients. These visualizations suggest that the differences in measurable disease activity characteristics between the tofacitinib treatment groups were too small to be accurately assessed, even with advanced data analysis techniques. The



◀ **Fig. 3** Visualization of change in total Mayo score from study baseline to week 52 of OCTAVE Sustain in patients receiving **a** tofacitinib 10 mg BID, **b** tofacitinib 5 mg BID, or **c** placebo. Individual patients are shown along the x-axis. *BID* twice daily

relatively low patient numbers included in this analysis may also have influenced the performance of the models, and this was partly confirmed by the small increase in power with the models that utilized the whole set of patients.

Previous studies have evaluated the role of biomarkers, such as fecal calprotectin, in predicting endoscopic response and possible relapse among patients with UC [26, 27]. A prospective, real-world study of patients with UC in Spain demonstrated that fecal calprotectin levels could act as a potential surrogate marker for endoscopic remission [27]. The absence of such biomarkers in the OCTAVE clinical program may have influenced the overall predictive value of the models, and future studies might consider including a different array of variables.

This analysis included visualizations of disease activity and discontinuation patterns that provided some insight into the effects of maintenance treatment over time. Notably, while patients who received tofacitinib 5 or 10 mg BID exhibited generally similar patterns of response, patients who received placebo demonstrated a distinct outcome pattern, compared with those who received tofacitinib, and this is consistent with previously reported results of induction study efficacy data [10].

Furthermore, the visualizations established that a change in rectal bleeding subscore was the primary determinant of disease worsening, as indicated by an increase in total Mayo score between baseline and week 52 of OCTAVE Sustain. This is consistent with a previous meta-analysis, which found that a normal rectal bleeding subscore correlated with the achievement of endoscopic remission among patients with UC [28]. The rectal bleeding subscore is a semiquantitative, subjective, patient-reported variable that has demonstrated fair correlation with endoscopic activity in patients with UC

[29]. However, similar to stool frequency, it is subject to substantial bias. Conversely, endoscopic activity (measured using the Mayo endoscopic subscore) is relatively objective and, in the global OCTAVE clinical program, was reviewed centrally. Here, the rectal bleeding subscore was not normally distributed. It is possible that the variance in the total Mayo score among patients with disease worsening was driven by a subgroup of outliers who were reporting higher rectal bleeding subscores compared with other variables, such as stool frequency, Physician Global Assessment, and endoscopic subscores, which either remained the same or improved. However, a change in the rectal bleeding subscore may be an early indicator of change in disease activity in patients with UC [30]. Furthermore, a strong correlation between clinical symptoms and endoscopic disease severity has been reported [31]. Therefore, it is possible that improvements in the endoscopic subscore may be indicative of an improvement in a patient's clinical symptoms.

This study had some limitations. Firstly, the number of patients included in this prediction analysis was relatively small, especially in some treatment subgroups, which limits the ability to interpret the data. Also, the prediction models have not been confirmed or validated by other data sources. External validation is key to determining the reproducibility and generalizability of the models to other patient populations. In addition, all patients who enrolled in OCTAVE Sustain had at least achieved a clinical response following OCTAVE Induction 1 and 2, and this may have contributed to the generally similar patterns of response with both tofacitinib doses in OCTAVE Sustain. Finally, this study only included data from patients who met the inclusion/exclusion criteria for the global OCTAVE clinical program. This may limit the generalizability of the results in the general population of patients with UC, and further studies that include prospective and real-world validation of the predictive models should be performed.

CONCLUSION

Using data from the global OCTAVE clinical program, we could not generate sufficiently accurate models that could predict loss of responder status with maintenance therapy and/or steroid-free remission. However, in-depth analyses of the individual disease score patterns indicated that there were particular differences in Mayo subscores when patients with disease worsening and disease improvement were compared; the rectal bleeding subscore was the primary determinant of disease worsening, while the endoscopic subscore was the primary determinant of disease improvement. The development of prediction models contributes to a growing body of research that is exploring the use of various statistical techniques to predict treatment outcomes with potential application in clinical practice. Notably, it is possible that there were missing data related to yet unknown key variables that were not collected during the clinical trials included in this analysis that could have significantly improved the accuracy of the models, and further studies with larger sample sizes are required.

ACKNOWLEDGMENTS

The authors would like to thank the patients, investigators, and study teams involved in the OCTAVE trials.

Funding. This study, the journal's Rapid Service and Open Access Fee was sponsored by Pfizer.

Medical Writing and Editorial Assistance. Medical writing support, under the direction of the authors, was provided by Chimwemwe Chibambo, MBChB, CMC Connect, a division of IPG Health Medical Communications, and was funded by Pfizer, New York, NY, USA, in accordance with Good Publication Practice (GPP 2022) guidelines (Ann Intern Med. 2022;175:1298–304). Gianluca Bonfanti has a contract with Health Services Consulting

Corporation, which was a paid contractor to Pfizer in connection with the formal analysis of this manuscript. Roger A. Edwards is a Principal in Health Services Consulting Corporation, which was a paid contractor to Pfizer in connection with the formal analysis of this manuscript.

Author Contributions. All authors have made a significant contribution to the research described in this manuscript. Michael Chiorean: conceptualization, writing – review & editing. Marco Daperno: conceptualization, writing – review & editing. Charlie W. Lees: conceptualization, writing – review & editing. Gianluca Bonfanti: formal analysis, writing – review & editing. Dimitrios Soudis: formal analysis, writing – review & editing. Irene Modesto: conceptualization, writing – review & editing. J. Jasper Deuring: conceptualization, methodology, writing – review & editing. Roger A. Edwards: conceptualization, methodology, writing – review & editing. All authors approved the final version of the article, including the authorship list.

Disclosures. Michael Chiorean has served as a speaker for AbbVie, Bristol Myers Squibb, Janssen, Medtronic, Pfizer Inc, and Takeda; and has served as a consultant for AbbVie, Arena, Bristol Myers Squibb, Eli Lilly, Janssen, Medtronic, Pfizer Inc, Prometheus, and Takeda. Marco Daperno has received personal fees from AbbVie, Chiesi, Ferring, Janssen, MSD, Pfizer Inc, and Quintiles; has received grants and personal fees from Takeda; has received non-financial support from SOFAR; and has been involved in clinical trials for Eli Lilly, Galapagos, Gilead, Janssen, MSD, Roche, and Takeda. Charlie W. Lees has served as a speaker for AbbVie, Dr Falk, Ferring, Hospira, Janssen, MSD, Pfizer Inc, Shire, Takeda, and Warner-Chilcott; has served as a consultant for AbbVie, Dr Falk, Gilead, GSK, Hospira, Iterative Scopes, Janssen, MSD, Oshi Health, Pfizer Inc, Pharmacosmos, Takeda, Topivert, Trellus Health, and Vifor Pharma; and has received research funding from AbbVie and Gilead. Gianluca Bonfanti has a contract with Health Services Consulting Corporation, which was a paid contractor to Pfizer in connection with the formal analysis of this manuscript. Dimitrios

Soudis, Irene Modesto, and J. Jasper Deuring are employees of, and own stocks and shares in, Pfizer Inc. Roger A. Edwards is a Principal in Health Services Consulting Corporation, which was a paid contractor to Pfizer in connection with the formal analysis of this manuscript.

Compliance with Ethics Guidelines. The study protocols were approved by the Institutional Review Board or Independent Ethics Committee for each participating center. Written informed consent was obtained from all patients, per the ethics committee-approved protocols. All studies were conducted in compliance with the ethical principles derived from the Declaration of Helsinki and in compliance with all International Council for Harmonisation Good Clinical Practice Guidelines.

Data Availability. The datasets generated and analyzed during the current study are not publicly available in order to ensure the protection of patient privacy in compliance with the EU General Data Protection Regulation (GDPR) but are available from the corresponding author on reasonable request.

Open Access. This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. Ungaro R, Mehandru S, Allen PB, Peyrin-Biroulet L, Colombel J-F. Ulcerative colitis. *Lancet*. 2017;389:1756–70.
2. Ordás I, Eckmann L, Talamini M, Baumgart DC, Sandborn WJ. Ulcerative colitis. *Lancet*. 2012;380:1606–19.
3. Weissshof R, El Jurdi K, Zmeter N, Rubin DT. Emerging therapies for inflammatory bowel disease. *Adv Ther*. 2018;35:1746–62.
4. Verstockt B, Ferrante M, Vermeire S, Van Assche G. New treatment options for inflammatory bowel diseases. *J Gastroenterol*. 2018;53:585–90.
5. Danese S, Fiorino G, Peyrin-Biroulet L. Positioning therapies in ulcerative colitis. *Clin Gastroenterol Hepatol*. 2020;18:1280–90.e1.
6. Kayal M, Shah S. Ulcerative colitis: current and emerging treatment strategies. *J Clin Med*. 2019;9:94.
7. Borg-Bartolo SP, Boyapati RK, Satsangi J, Kalla R. Precision medicine in inflammatory bowel disease: concept, progress and challenges. *F1000Res*. 2020;9:F1000.
8. Viennois E, Zhao Y, Merlin D. Biomarkers of inflammatory bowel disease: from classical laboratory tools to personalized medicine. *Inflamm Bowel Dis*. 2015;21:2467–74.
9. Sandborn WJ, Ghosh S, Panes J, et al. Tofacitinib, an oral Janus kinase inhibitor, in active ulcerative colitis. *N Engl J Med*. 2012;367:616–24.
10. Sandborn WJ, Su C, Sands BE, et al. Tofacitinib as induction and maintenance therapy for ulcerative colitis. *N Engl J Med*. 2017;376:1723–36.
11. Sandborn WJ, Lawendy N, Danese S, et al. Safety and efficacy of tofacitinib for treatment of ulcerative colitis: final analysis of OCTAVE Open, an open-label, long-term extension study with up to 7.0 years of treatment. *Aliment Pharmacol Ther*. 2022;55:464–78.
12. Vermeire S, Su C, Lawendy N, et al. Outcomes of tofacitinib dose reduction in patients with ulcerative colitis in stable remission from the randomised RIVETING trial. *J Crohns Colitis*. 2021;15:1130–41.
13. US Food and Drug Administration. XELJANZ® (tofacitinib): highlights of prescribing information. 2022. <http://labeling.pfizer.com/ShowLabeling.aspx?id=959>. Accessed 23 Aug 2022.

14. Lees CW, Deuring JJ, Chiorean M, et al. Prediction of early clinical response in patients receiving tofacitinib in the OCTAVE Induction 1 and 2 studies. *Therap Adv Gastroenterol*. 2021;14:17562848211054710.
15. Winthrop KL, Melmed GY, Vermeire S, et al. Herpes zoster infection in patients with ulcerative colitis receiving tofacitinib. *Inflamm Bowel Dis*. 2018;24:2258–65.
16. Lewis JD, Chuai S, Nessel L, Lichtenstein GR, Aberra FN, Ellenberg JH. Use of the noninvasive components of the Mayo score to assess clinical response in ulcerative colitis. *Inflamm Bowel Dis*. 2008;14:1660–6.
17. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol*. 2010;5:1315–6.
18. Hosmer DW, et al. Assessing the fit of the model. In: *Applied logistic regression*. Hoboken; 2000. p. 143–202. <https://doi.org/10.1002/0471722146.ch5>.
19. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B Methodol*. 1996;58:267–88.e3.
20. Göbl CS, Bozkurt L, Tura A, Pacini G, Kautzky-Willer A, Mittlböck M. Application of penalized regression techniques in modelling insulin sensitivity by correlated metabolic parameters. *PLoS ONE*. 2015;10: e0141524.
21. Sandborn WJ, Armuzzi A, Liguori G, et al. Predictors of sustained response with tofacitinib therapy in patients with ulcerative colitis. *Inflamm Bowel Dis*. 2022;28:1338–47.
22. Torres J, Caprioli F, Katsanos KH, et al. Predicting outcomes to optimize disease management in inflammatory bowel diseases. *J Crohns Colitis*. 2016;10:1385–94.
23. Lamb CA, Saifuddin A, Powell N, Rieder F. The future of precision medicine to predict outcomes and control tissue remodeling in inflammatory bowel disease. *Gastroenterology*. 2022;162:1525–42.
24. Buescher JM, Driggers EM. Integration of omics: more than the sum of its parts. *Cancer Metab*. 2016;4:4.
25. Fiocchi C, Iliopoulos D. IBD systems biology is here to stay. *Inflamm Bowel Dis*. 2021;27:760–70.
26. Theede K, Holck S, Ibsen P, Kallemose T, Nordgaard-Lassen I, Nielsen AM. Fecal calprotectin predicts relapse and histological mucosal healing in ulcerative colitis. *Inflamm Bowel Dis*. 2016;22:1042–8.
27. Lobatón T, Rodríguez-Moranta F, Lopez A, Sánchez E, Rodríguez-Alonso L, Guardiola J. A new rapid quantitative test for fecal calprotectin predicts endoscopic activity in ulcerative colitis. *Inflamm Bowel Dis*. 2013;19:1034–42.
28. Narula N, Alshahrani AA, Yuan Y, Reinisch W, Colombel JF. Patient-reported outcomes and endoscopic appearance of ulcerative colitis: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol*. 2019;17:411–8.e3.
29. Colombel J-F, Lakatos PL, Santana GO, et al. Diagnostic accuracy of patient-reported outcomes in predicting endoscopic subscore in patients with ulcerative colitis. *GastroHep*. 2021;3:161–8.
30. Hanauer S, Panaccione R, Danese S, et al. Tofacitinib induction therapy reduces symptoms within 3 days for patients with ulcerative colitis. *Clin Gastroenterol Hepatol*. 2019;17:139–47.
31. Restellini S, Chao CY, Martel M, et al. Clinical parameters correlate with endoscopic activity of ulcerative colitis: a systematic review. *Clin Gastroenterol Hepatol*. 2019;17:1265–75.e8.