Modified Mayo score *versus* Mayo score for evaluation of treatment efficacy in patients with ulcerative colitis: data from the tofacitinib OCTAVE program

William J. Sandborn, Bruce E. Sands, Séverine Vermeire, Yvette Leung, Xiang Guo, Irene Modesto, Chinyu Su, Wenjin Wang and Julian Panés

Abstract

Objectives: The subjectivity of the Physician Global Assessment (PGA) is a limitation of the Mayo score in assessing severity of ulcerative colitis (UC). We compared treatment efficacy using endpoint definitions based on modified Mayo (mMayo) score, *versus* those based on Mayo score, using data from the tofacitinib OCTAVE program.

Design: This post hoc analysis included data from two 8-week induction studies (OCTAVE Induction 1 and 2) and a 52-week maintenance study (OCTAVE Sustain).

Methods: Remission and clinical response [with nonresponder imputation (NRI)] were assessed using mMayo (without PGA) and Mayo scores, and further stratified by prior tumor necrosis factor inhibitor (TNFi) failure status.

Results: At week 8 of OCTAVE Induction 1 and 2, remission rates with placebo and tofacitinib 10 mg twice daily (BID), respectively, were 7.7% and 24.8% (mMayo) and 6.0% and 17.6% (Mayo). At week 52 of OCTAVE Sustain, remission rates with placebo, tofacitinib 5 and 10 mg BID, respectively, were 12.1%, 35.9%, and 42.1% (mMayo) and 11.1%, 34.3%, and 40.6% (Mayo). A statistically significant ($p < 0.05$) treatment effect of tofacitinib *versus* placebo was observed for remission and clinical response at all time points, regardless of scoring definition or prior TNFi failure status.

Conclusions: A significant effect of tofacitinib *versus* placebo was demonstrated across efficacy endpoints using mMayo score, consistent with previously reported data using Mayo score. Treatment effect sizes were generally similar regardless of scoring definition. This observation may help contextualize tofacitinib therapy outcomes with those of new UC therapies and support the use of Mayo score-based endpoints in UC clinical trials.

Trail registration: ClinicalTrials.gov identifiers: NCT01465763; NCT01458951; NCT01458574.

Keywords: Mayo score, tofacitinib, ulcerative colitis

Correspondence to:
William J. Sandborn
Division of Gastroenterology, University of California San Diego, 9500 Gilman Drive, MC 0956, La Jolla, 92093, CA, USA
wsandborn@health.ucsd. edu

Bruce E. Sands
Dr. Henry D. Janowitz Division of Gastroenterology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Séverine Vermeire
Department of Gastroenterology and Hepatology, University Hospitals Leuven, Leuven, Belgium

Yvette Leung
Department of Medicine, University of British Columbia, Vancouver, BC, Canada

Xiang Guo
Chinyu Su
Wenjin Wang
Pfizer Inc, Collegeville, PA, USA

Irene Modesto
Pfizer Inc, New York, NY, USA

Julian Panés
Inflammatory Bowel Disease Unit, Hospital Clínic de Barcelona, IDIBAPS, CIBERehd, Barcelona, Spain

Introduction

Clinical trials of ulcerative colitis (UC) therapies have typically used endpoint definitions based on the conventional Mayo score, comprising four subscores: stool frequency, rectal bleeding, endoscopic findings, and Physician Global Assessment (PGA). The PGA is an arbitrarily designed multicomponent evaluation of disease activity using the physician's assessment and is based on disease activity and symptom severity.[1] As a subjective measure, the PGA subscore may be a limitation of the Mayo score to assess UC disease severity. Guidance from the U.S. Food and Drug Administration (FDA), published in 2016,

recommends the use of a modified Mayo (mMayo) score in which the PGA subscore is omitted as an endpoint measure for UC clinical trials.[2] As a result, new therapies in development for the treatment of UC now include the mMayo score when assessing efficacy in clinical trials; therefore, evaluating efficacy using the mMayo score in already approved drugs, even in a post hoc fashion, will allow physicians to better contextualize all available data.

Tofacitinib is an oral small molecule Janus kinase inhibitor for the treatment of UC. The efficacy and safety of tofacitinib have been evaluated in patients with moderately to severely active UC in three phase III studies [two identical 8-week induction studies (OCTAVE Induction 1 and 2) and a 52-week maintenance study (OCTAVE Sustain)][3] and an open-label, long-term extension study (OCTAVE Open).[4] The start date for these studies was 2012; endpoint definitions were therefore based on the conventional Mayo score. As it is important to contextualize these data with those from future trials of UC therapies, we carried out a post hoc analysis of the phase III data to compare mMayo score with Mayo score for the evaluation of treatment efficacy using data from the tofacitinib OCTAVE program.

Materials and methods

Patients and study design
The full study design details of OCTAVE Induction 1 and 2 (ClinicalTrials.gov; NCT01465763 NCT01465763 and NCT01458951) and OCTAVE Sustain (ClinicalTrials.gov; NCT01458574) have been reported previously.[3]

In OCTAVE Induction 1 and 2, patients with moderately to severely active UC were randomized to receive placebo, or tofacitinib 10 or 15 mg twice daily (BID). The tofacitinib 15 mg BID dose in OCTAVE Induction 1 and 2 was subsequently discontinued following a protocol amendment.

Patients who completed OCTAVE Induction 1 and 2 and achieved clinical response (defined as a decrease from induction study baseline total Mayo score of ≥3 points and ≥30%, plus a decrease in rectal bleeding subscore of ≥1 point or an absolute rectal bleeding subscore of 0 or 1) at week 8 were eligible to enter OCTAVE Sustain,

in which patients were re-randomized to receive placebo, or tofacitinib 5 or 10 mg BID.

Oral corticosteroids (prednisone-equivalent up to 25 mg/day) were permitted during OCTAVE Induction 1 and 2, provided that the dose remained stable for at least 2 weeks prior to baseline and throughout the induction study period. Corticosteroid tapering was mandatory during OCTAVE Sustain. Concomitant immunosuppressant or tumor necrosis factor inhibitor (TNFi) therapy was prohibited throughout the induction and maintenance studies.

Efficacy assessments
Efficacy endpoints were derived from mMayo and Mayo scores, based on the centrally read endoscopic subscore, where a single read was performed by one reader who was blinded to the treatment assignment, study, visit and patient's clinical status.[5] mMayo score used in this post hoc analysis comprises three subscores: endoscopic findings, stool frequency, and rectal bleeding. The Mayo score—the scoring definition that was used in the prospective OCTAVE studies and primary efficacy analyses[3]—comprises four subscores: endoscopic findings, stool frequency, rectal bleeding, and PGA. Each component is scored from 0 to 3; therefore, the range of the mMayo score is 0–9, while the range of the Mayo score is 0–12.

Proportions of patients achieving remission and clinical response were assessed in OCTAVE Induction 1 and 2 (at week 8) and OCTAVE Sustain [at week 24, week 52, and at both time points (sustained)], using both mMayo and Mayo scoring definitions. Remission based on the mMayo score (modified remission) was defined as an endoscopic subscore of ≤1, a stool frequency subscore of ≤1, and a rectal bleeding subscore of 0. Remission based on the Mayo score was defined as a total Mayo score (including PGA subscore) of ≤2 with no individual subscore >1, and a rectal bleeding subscore of 0. Clinical response based on the mMayo score (modified clinical response) was defined as a decrease from induction study baseline mMayo score of ≥2 points and ≥30%, plus a decrease in rectal bleeding subscore of ≥1 point or an absolute rectal bleeding subscore of 0 or 1. Clinical response based on the Mayo score was defined as a decrease from induction study baseline total Mayo score of

≥3 points and ≥30%, plus a decrease in rectal bleeding subscore of ≥1 point or an absolute rectal bleeding subscore of 0 or 1.

In addition to remission and clinical response, sustained steroid-free remission was evaluated among patients who were in remission at baseline of OCTAVE Sustain using both mMayo and Mayo scoring definitions. Sustained steroid-free remission was defined as being in remission at both weeks 24 and 52 of OCTAVE Sustain and steroid-free (not requiring corticosteroid treatment for at least 4 weeks prior to study visit).

The proportions of patients who achieved remission, clinical response, or sustained steroid-free remission, by either mMayo or Mayo score definitions, were further stratified by prior TNFi failure status (yes or no).

While recent trials have adopted the regulatory guidance on omitting PGA in outcome measures, there is no consistent definition of remission using the mMayo scoring system.[2,6–8] Therefore, we sought to evaluate outcomes using three alternative definitions of modifed remission, to fully contextualize tofacitinib efficacy data with other trials of UC therapies. In alternative definition A, remission was defined as an endoscopic subscore of ≤1, a stool frequency subscore of ≤1 and no greater than at induction study baseline, and a rectal bleeding subscore of 0.[9] In alternative definition B, remission was defined as an endoscopic subscore of ≤1, a stool frequency subscore of ≤1 and a ≥1-point decrease from induction study baseline, and a rectal bleeding subscore of 0.[2] In alternative definition C, remission was defined as an endoscopic subscore of ≤1, a stool frequency subscore of 0, and a rectal bleeding subscore of 0.[2]

Statistical analysis

Efficacy endpoints are reported for the full analysis set, including all randomized patients assigned to placebo or tofacitinib 10 mg BID in OCTAVE Induction 1 and 2, and all randomized patients assigned to placebo, or tofacitinib 5 or 10 mg BID, in OCTAVE Sustain. Nonresponder imputation (NRI) was applied, meaning that patients with missing values were treated as nonresponders. Treatment effects [difference from placebo, 95% confidence intervals (CIs)] using mMayo and Mayo scores were evaluated, and treatment differences from placebo were compared using a Cochran-Mantel-Haenszel chi-squared test, stratified by study, prior TNFi exposure, corticosteroid use at baseline, and geographical region in OCTAVE Induction 1 and 2, and stratified by baseline remission status and induction study treatment in OCTAVE Sustain.

Ethical considerations

All studies were registered with ClinicalTrials.gov (NCT01465763; NCT01458951; NCT01458574) and were conducted in compliance with the Declaration of Helsinki and the International Conference on Harmonization Good Clinical Practice Guidelines, and were approved by the Institutional Review Boards and/or Independent Ethics Committees at each investigational center participating in the studies or at a central Institutional Review Board. All patients provided written informed consent.

Results

Patients

This post hoc analysis included 1139 patients from OCTAVE Induction 1 and 2 (placebo, $N = 234$; tofacitinib 10 mg BID, $N = 905$) and 593 patients with clinical response following OCTAVE Induction 1 and 2 who were enrolled into OCTAVE Sustain (placebo, $N = 198$; tofacitinib 5 mg BID, $N = 198$; tofacitinib 10 mg BID, $N = 197$).

Baseline demographics and clinical characteristics among patients in OCTAVE Induction 1 and 2, and OCTAVE Sustain, are presented in Table 1, and were generally similar between treatment groups.

Remission and clinical response in OCTAVE Induction 1 and 2

At week 8 of OCTAVE Induction 1 and 2, the observed effect of treatment with tofacitinib 10 mg BID *versus* placebo was statistically significant for remission and clinical response, using both the mMayo and Mayo scores (Figure 1). However, a higher proportion of patients achieved remission using the mMayo score compared with the Mayo score; rates of remission among patients who received placebo and tofacitinib 10 mg BID, respectively, were 7.7% and 24.8% using the mMayo score, and 6.0% and 17.6% using the

Table 1. Baseline demographics and clinical characteristics among patients in OCTAVE Induction 1 and 2, and OCTAVE Sustain.

	OCTAVE Induction 1 and 2		OCTAVE Sustain		
	Placebo *N* = 234	Tofacitinib 10 mg BID *N* = 905	Placebo *N* = 198	Tofacitinib 5 mg BID *N* = 198	Tofacitinib 10 mg BID *N* = 197
Male[a], *n* (%)	132 (56.4)	536 (59.2)	116 (58.6)	103 (52.0)	110 (55.8)
Mean age[a], years (SD)	41.1 (14.4)	41.2 (13.8)	43.4 (14.0)	41.9 (13.7)	42.9 (14.4)
Time since diagnosis[b] (years), mean (SD)	8.1 (7.0)	8.1 (7.0)	8.8 (7.5)	8.3 (7.2)	8.6 (7.0)
Extent of disease[c], *n* (%)					
Proctosigmoiditis	35 (15.0)	132 (14.6)	21 (10.6)	28 (14.3)	33 (16.8)
Left-sided colitis	76 (32.6)	307 (34.0)	68 (34.3)	66 (33.7)	60 (30.6)
Extensive colitis or pancolitis	122 (52.4)	463 (51.3)	108 (54.5)	102 (52.0)	103 (52.6)
Oral corticosteroid use at baseline, *n* (%)	113 (48.3)	412 (45.5)	105 (53.0)	101 (51.0)	92 (46.7)
Prior TNFi failure,[a] *n* (%)	124 (53.0)	465 (51.4)	89 (44.9)	83 (41.9)	93 (47.2)
Prior immunosuppressant failure,[a] *n* (%)	158 (67.5)	661 (73.0)	129 (65.2)	143 (72.2)	141 (71.6)
Baseline total Mayo score,[b] mean (SD)	9.0 (1.5)	9.0 (1.4)	3.3 (1.8)	3.3 (1.8)	3.4 (1.8)
Baseline mMayo score,[b] mean (SD)	6.7 (1.2)[d]	6.7 (1.2)[e]	2.5 (1.4)	2.5 (1.4)	2.6 (1.4)

[a]Indicated data are at baseline of OCTAVE Induction 1 and 2.
[b]For OCTAVE Sustain, indicated data are at baseline of OCTAVE Sustain.
[c]One patient with proctitis was enrolled into OCTAVE Induction 2 (received tofacitinib 10 mg BID) as a protocol deviation and progressed through to OCTAVE Sustain.
[d]*N* = 233.
[e]*N* = 903.
BID, twice daily; mMayo, modified Mayo (Mayo score excluding PGA); *N*, number of patients in the subgroup; *n*, number of patients in each category; PGA, Physician Global Assessment; SD, standard deviation; TNFi, tumor necrosis factor inhibitor.

Mayo score. There was a numerically greater difference from placebo with mMayo score (17.1%; 95% CI, 12.6–21.5) *versus* Mayo score (11.6%; 95% CI, 7.7–15.5) [Figure 1(a)]. In contrast, a similar proportion of patients achieved clinical response with both scoring definitions [Figure 1(b)].

The same trends were noted among patients with and without prior TNFi failure, with numerically more patients in the placebo and tofacitinib groups achieving remission, but not clinical response, using the mMayo score compared with the Mayo score (Figure 1). Treatment effect (difference *versus* placebo) was similar between those with and without prior TNFi failure, regardless of scoring definition.

Proportions of patients in remission at week 8 of OCTAVE Induction 1 and 2, assessed using alternative definitions of modified remission using the mMayo score, are shown in Supplemental Figure 1. Treatment effect sizes were generally similar across the definitions used for modified remission (including alternative definitions A and B), and were similar to results using remission
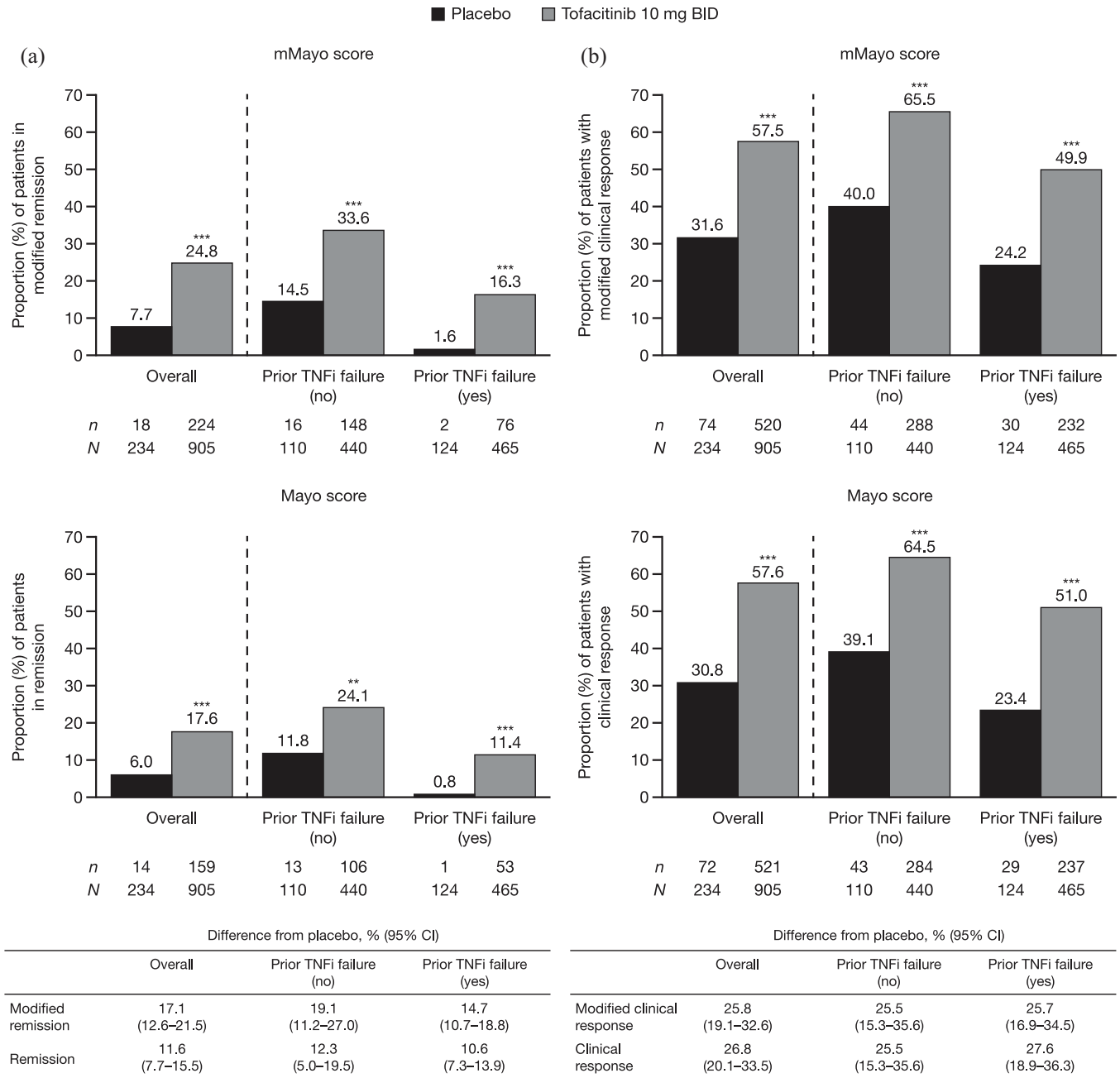
■ Placebo ▨ Tofacitinib 10 mg BID

Figure 1. Proportions of patients (a) in remission and (b) with clinical response, defined by mMayo score and Mayo score at week 8 of OCTAVE Induction 1 and 2, overall and stratified by prior TNFi failure status (FAS, NRI).

Remission was defined as a total Mayo score (including PGA subscore) of ≤2 with no individual subscore >1, and a rectal bleeding subscore of 0. Modified remission was defined as an endoscopic subscore of ≤1, a stool frequency subscore of ≤1, and a rectal bleeding subscore of 0. Clinical response was defined as a decrease from induction study baseline total Mayo score (including PGA subscore) of ≥3 points and ≥30%, plus a decrease in rectal bleeding subscore of ≥1 point or an absolute rectal bleeding subscore of 0 or 1. Modified clinical response was defined as a decrease from induction study baseline mMayo score (excluding PGA subscore) of ≥2 points and ≥30%, plus a decrease in rectal bleeding subscore of ≥1 point or an absolute rectal bleeding subscore of 0 or 1. Endoscopic subscore was based on central read. *p* values *versus* placebo were based on the Cochran-Mantel-Haenszel chi-squared test, stratified by study, prior treatment with TNFi, corticosteroid use at baseline, and geographical region. ****p<0.01 *versus* placebo. ***p<0.001 *versus* placebo.**

BID, twice daily; CI, confidence interval; FAS, full analysis set; mMayo, modified Mayo; *N*, number of patients in the analysis set; *n*, number of patients achieving endpoint; NRI, nonresponder imputation; PGA, Physician Global Assessment; TNFi, tumor necrosis factor inhibitor.
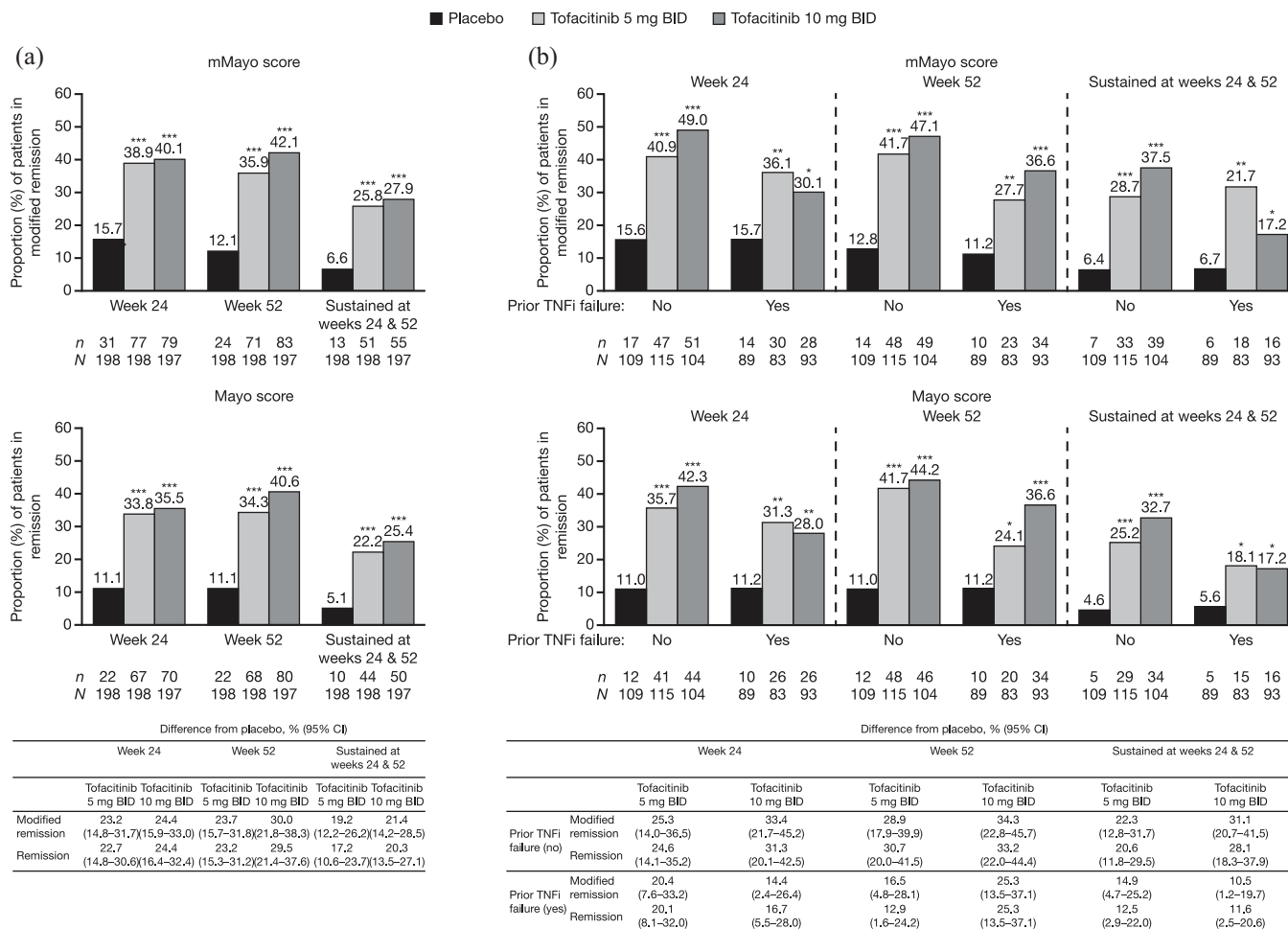
Figure 2. Proportions of patients in remission, defined by mMayo score and Mayo score at week 24, week 52, and at both time points (sustained) in OCTAVE Sustain, (a) overall and (b) stratified by prior TNFi failure status (FAS, NRI).
Remission was defined as a total Mayo score (including PGA subscore) of ≤2 with no individual subscore >1, and a rectal bleeding subscore of 0. Modified remission was defined as an endoscopic subscore of ≤1, a stool frequency subscore of ≤1, and a rectal bleeding subscore of 0. Endoscopic subscore was based on central read. *p* values *versus* placebo were based on the Cochran-Mantel-Haenszel chi-squared test, stratified by induction study treatment and baseline remission status.
*$p < 0.05$ *versus* placebo. **$p < 0.01$ *versus* placebo. ***$p < 0.001$ *versus* placebo.
BID, twice daily; CI, confidence interval; FAS, full analysis set; mMayo, modified Mayo; *N*, number of patients in the analysis set; *n*, number of patients achieving endpoint; NRI, nonresponder imputation; PGA, Physician Global Assessment; TNFi, tumor necrosis factor inhibitor.

based on Mayo score. Treatment effect sizes using alternative definition C were smaller than for other definitions.

Remission, sustained remission, and clinical response in OCTAVE Sustain
At week 24 and/or week 52 of OCTAVE Sustain, there was a significant treatment effect for tofacitinib 5 or 10 mg BID *versus* placebo for remission, sustained remission, and clinical response using both the mMayo and Mayo scores [Figures 2(a) and 3(a)]. Using the mMayo score, rates of

remission among patients who received placebo, tofacitinib 5 mg BID, and tofacitinib 10 mg BID, respectively, were: 15.7%, 38.9%, and 40.1% of patients at week 24; 12.1%, 35.9%, and 42.1% of patients at week 52; and 6.6%, 25.8%, and 27.9% of patients at both weeks 24 and 52 (sustained remission) [Figure 2(a)]. Furthermore, the differences in the rates of remission from placebo with tofacitinib 5 or 10 mg BID, respectively, using the mMayo score were: 23.2% (95% CI, 14.8–31.7) and 24.4% (95% CI, 15.9–33.0) at week 24; 23.7% (95% CI, 15.7–31.8) and 30.0% (95% CI, 21.8–38.3) at week 52; and 19.2% (95%
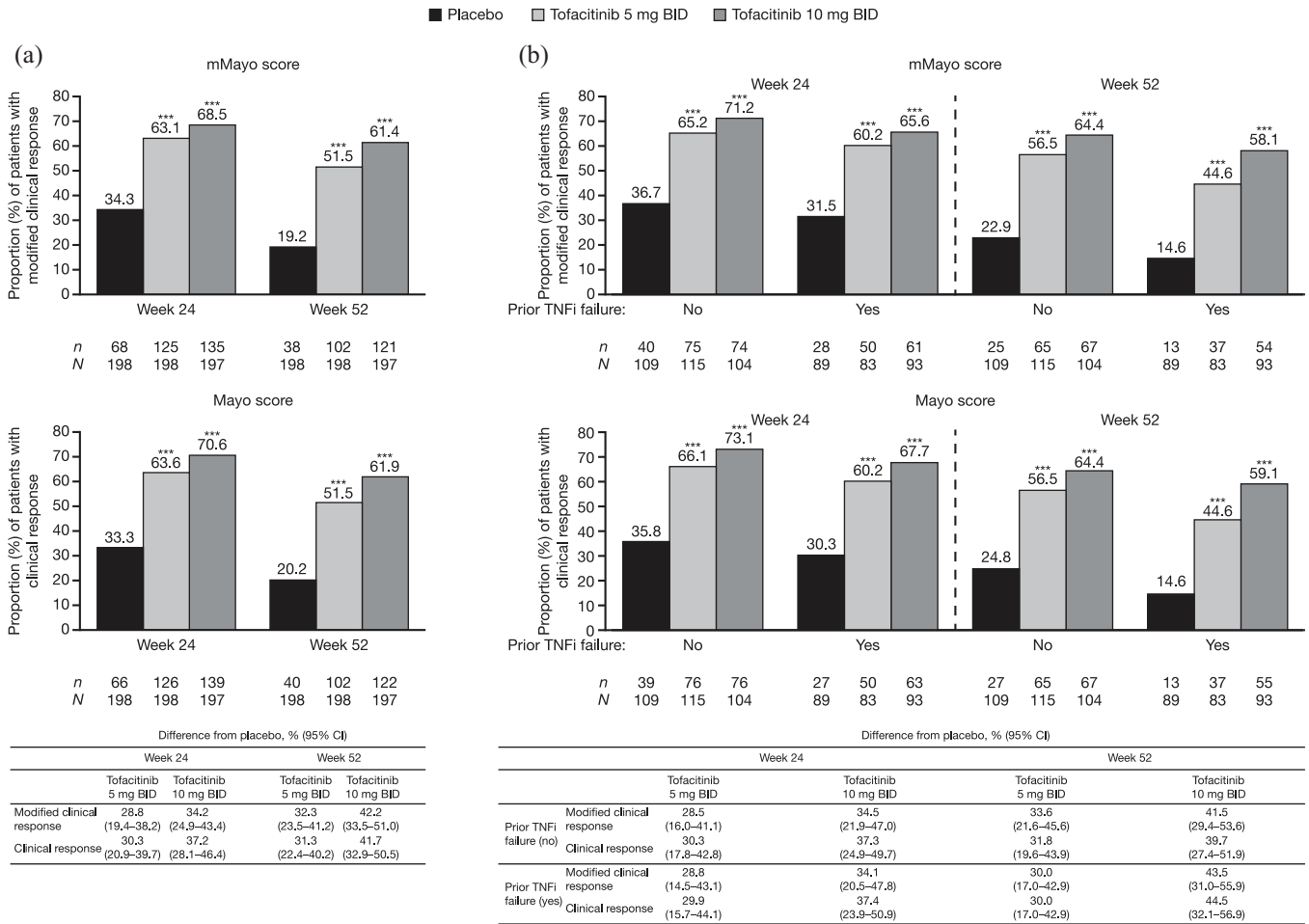
Figure 3. Proportions of patients with clinical response, defined by mMayo score and Mayo score at week 24 or week 52 in OCTAVE Sustain, (a) overall and (b) stratified by prior TNFi failure status (FAS, NRI).

Clinical response was defined as a decrease from induction study baseline total Mayo score (including PGA subscore) of ≥3 points and ≥30%, plus a decrease in rectal bleeding subscore of ≥1 point or an absolute rectal bleeding subscore of 0 or 1. Modified clinical response was defined as a decrease from induction study baseline mMayo score (excluding PGA subscore) of ≥2 points and ≥30%, plus a decrease in rectal bleeding subscore of ≥1 point or an absolute rectal bleeding subscore of 0 or 1. Endoscopic subscore was based on central read. *p* values *versus* placebo were based on the Cochran–Mantel–Haenszel chi-squared test, stratified by induction study treatment and baseline remission status.

***p < 0.001 *versus* placebo.

BID, twice daily; CI, confidence interval; FAS, full analysis set; mMayo, modified Mayo; *N*, number of patients in the analysis set; *n*, number of patients achieving endpoint; NRI, nonresponder imputation; PGA, Physician Global Assessment; TNFi, tumor necrosis factor inhibitor.

CI, 12.2–26.2) and 21.4% (95% CI, 14.2–28.5) at both weeks 24 and 52 (sustained remission). When using the Mayo score, remission at week 24 and/or week 52 of OCTAVE Sustain was achieved by slightly lower proportions of patients compared with using the mMayo score; however, the differences in proportions determined with each scoring definition were small and the treatment effect was similar.

Proportions of patients in remission at week 24 and/or week 52, assessed using alternative definitions of modified remission, are shown in

Supplemental Figure 2. Treatment effect sizes were generally similar across the alternative definitions used for modified remission (including alternative definitions A and B), and were similar to results using remission based on Mayo score. Treatment effect sizes using alternative definition C were smaller than for other definitions.

Using the mMayo score, rates of clinical response among patients who received placebo, tofacitinib 5 mg BID, and tofacitinib 10 mg BID, respectively, were: 34.3%, 63.1%, and 68.5% of patients at week 24; and 19.2%, 51.5%, and 61.4% of

patients at week 52 [Figure 3(a)]. Clinical response was achieved by a similar proportion of patients when using the Mayo score. Using the mMayo score, differences in the rates of clinical response from placebo with tofacitinib 5 or 10 mg BID, respectively, were: 28.8% (95% CI, 19.4–38.2) and 34.2% (95% CI, 24.9–43.4) at week 24; and 32.3% (95% CI, 23.5–41.2) and 42.2% (95% CI, 33.5–51.0) at week 52. Treatment effect was similar when using the Mayo score.

There was a significant treatment effect for tofacitinib 5 or 10 mg BID *versus* placebo for remission, sustained remission, and clinical response using both the mMayo and Mayo scores, regardless of prior TNFi failure status [Figures 2(b) and 3(b)]. However, among patients with prior TNFi failure, a lower proportion of tofacitinib-treated patients achieved these efficacy endpoints at week 24 and/or week 52, *versus* patients with no prior TNFi failure, regardless of scoring definition used.

Sustained steroid-free remission in OCTAVE Sustain

Among patients who were in remission at baseline of OCTAVE Sustain, a statistically significant treatment effect of tofacitinib 5 or 10 mg BID *versus* placebo was observed for sustained steroid-free remission at weeks 24 and 52 using both the mMayo and Mayo scores (Figure 4).

Overall, numerically similar proportions of patients achieved sustained steroid-free remission using both mMayo and Mayo scoring definitions. Rates of sustained steroid-free remission among patients who received placebo, tofacitinib 5 mg BID, and tofacitinib 10 mg BID, respectively, were 8.8%, 29.2%, and 46.2% using the mMayo score, and 5.1%, 35.4%, and 47.3% using the Mayo score. There was a numerically greater difference from placebo with tofacitinib 5 or 10 mg BID, respectively, using the Mayo score (30.3%; 95% CI, 17.4–43.2 and 42.2%; 95% CI, 27.9–56.5) *versus* the mMayo score (20.5%; 95% CI, 9.2–31.8 and 37.4%; 95% CI, 24.7–50.1).

When stratified by prior TNFi failure status, there was a significant treatment effect for tofacitinib 10 mg BID *versus* placebo for remission, using both the mMayo and Mayo scores, regardless of prior TNFi failure status. Lower proportions of tofacitinib-treated patients with prior TNF

failure achieved sustained steroid-free remission *versus* those without prior TNFi failure, regardless of the scoring definition used (Figure 4).
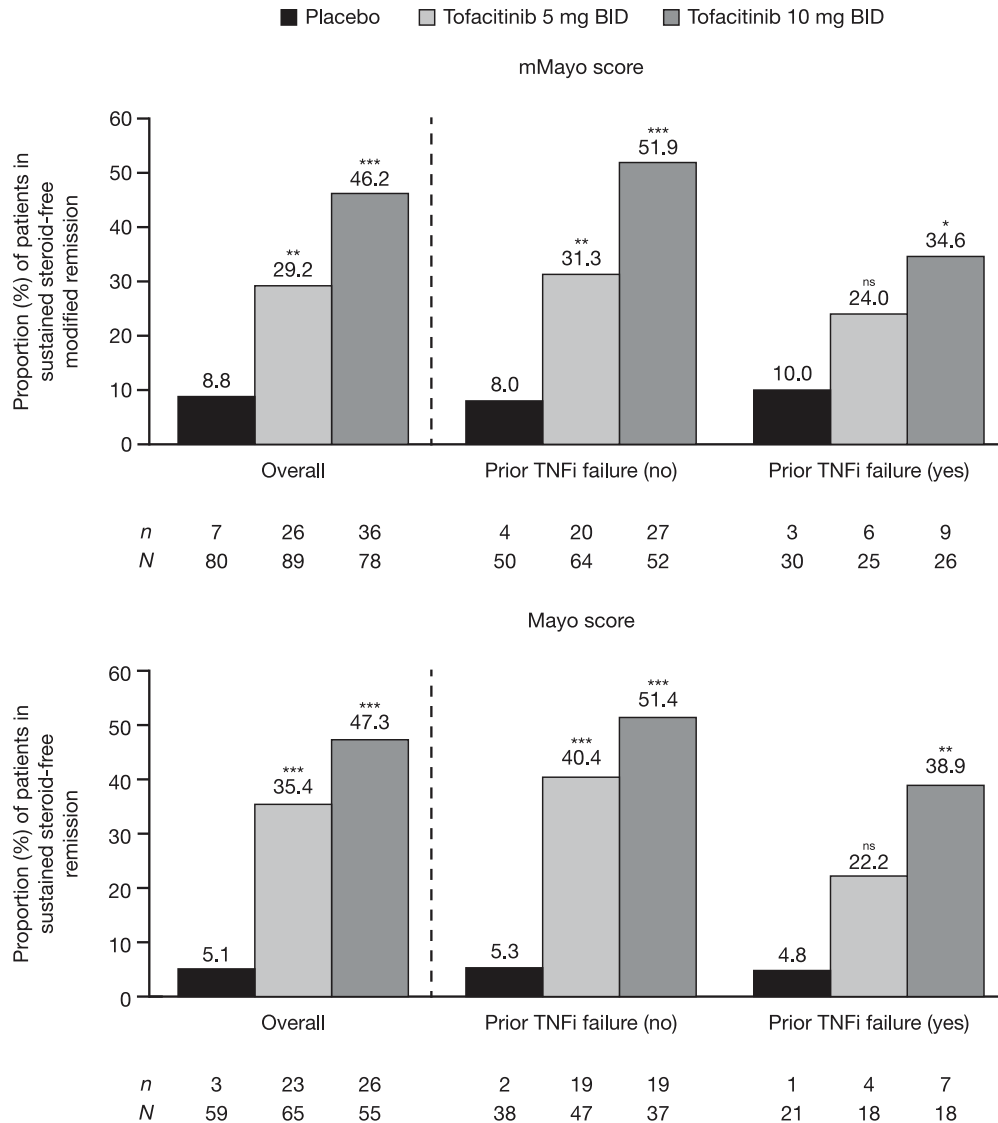
Proportions of patients in remission at baseline of OCTAVE Sustain who achieved sustained steroid-free remission at weeks 24 and 52 of OCTAVE Sustain, assessed using alternative definitions of modified remission, are shown in Supplemental Figure 3. Proportions of patients in remission were generally similar across the alternative definitions used for modified remission (including alternative definitions A and B), and were similar to results using remission based on Mayo score. Treatment effect size using alternative definition C was slightly greater than for other definitions.

Discussion

In this post hoc analysis, we compared efficacy endpoints, using scoring definitions based on the mMayo and Mayo scores; we used data from phase III tofacitinib studies (OCTAVE Induction 1 and 2, and OCTAVE Sustain) in patients with moderately to severely active UC, to enable contextualization with data from future trials of UC therapies.

A statistically significant effect of tofacitinib treatment *versus* placebo was demonstrated across efficacy endpoints based on the mMayo score, consistent with previously reported data from the OCTAVE studies based on the Mayo score.[3] In OCTAVE Induction 1 and 2, there was a greater treatment effect for remission, but not clinical response, based on the mMayo score compared with the Mayo score. In OCTAVE Sustain, treatment effect sizes for remission, sustained steroid-free remission, and clinical response, based on the mMayo score, were generally similar to those based on the Mayo score.

Patients with UC who have previously failed TNFi therapy are generally considered more refractory to treatment than those who have not;[10,11] therefore, it is of interest to evaluate the efficacy of tofacitinib treatment in this subgroup of patients. In this analysis, tofacitinib was efficacious regardless of prior TNFi failure status or scoring definition; a generally similar dose–response pattern was observed when using the mMayo or Mayo scores in OCTAVE Induction 1 and 2, and OCTAVE Sustain. These results expand on previous analyses from OCTAVE

■ Placebo ☐ Tofacitinib 5 mg BID ▨ Tofacitinib 10 mg BID

mMayo score

Proportion (%) of patients in sustained steroid-free modified remission

Overall: Placebo 8.8; Tofacitinib 5 mg BID ** 29.2; Tofacitinib 10 mg BID *** 46.2

Prior TNFi failure (no): Placebo 8.0; Tofacitinib 5 mg BID ** 31.3; Tofacitinib 10 mg BID *** 51.9

Prior TNFi failure (yes): Placebo 10.0; Tofacitinib 5 mg BID ns 24.0; Tofacitinib 10 mg BID * 34.6

	Overall			Prior TNFi failure (no)			Prior TNFi failure (yes)		
n	7	26	36	4	20	27	3	6	9
N	80	89	78	50	64	52	30	25	26

Mayo score

Proportion (%) of patients in sustained steroid-free remission

Overall: Placebo 5.1; Tofacitinib 5 mg BID *** 35.4; Tofacitinib 10 mg BID *** 47.3

Prior TNFi failure (no): Placebo 5.3; Tofacitinib 5 mg BID *** 40.4; Tofacitinib 10 mg BID *** 51.4

Prior TNFi failure (yes): Placebo 4.8; Tofacitinib 5 mg BID ns 22.2; Tofacitinib 10 mg BID ** 38.9

	Overall			Prior TNFi failure (no)			Prior TNFi failure (yes)		
n	3	23	26	2	19	19	1	4	7
N	59	65	55	38	47	37	21	18	18

Difference from placebo, % (95% CI)

	Overall		Prior TNFi failure (no)		Prior TNFi failure (yes)	
	Tofacitinib 5 mg BID	Tofacitinib 10 mg BID	Tofacitinib 5 mg BID	Tofacitinib 10 mg BID	Tofacitinib 5 mg BID	Tofacitinib 10 mg BID
Sustained steroid-free modified remission	20.5 (9.2–31.8)	37.4 (24.7–50.1)	23.3 (9.6–36.9)	43.9 (28.4–59.4)	14.0 (–5.9–33.9)	24.6 (3.4–45.8)
Sustained steroid-free remission	30.3 (17.4–43.2)	42.2 (27.9–56.5)	35.2 (19.4–50.9)	46.1 (28.5–63.7)	17.5 (–3.8–38.7)	34.1 (9.8–58.4)

Figure 4. Proportions of patients in remission[a] at baseline of OCTAVE Sustain who achieved sustained steroid-free remission at weeks 24 and 52 of OCTAVE Sustain, defined by mMayo score and Mayo score, overall and stratified by prior TNFi failure status (FAS, NRI).

Remission was defined as a total Mayo score (including PGA subscore) of ≤2 with no individual subscore >1, and a rectal bleeding subscore of 0. Modified remission was defined as an endoscopic subscore of ≤1, a stool frequency subscore of ≤1, and a rectal bleeding subscore of 0. Sustained steroid-free remission was defined as being in remission and steroid-free (not requiring corticosteroid treatment for at least 4 weeks prior to study visit) at both weeks 24 and 52. Endoscopic subscore was based on central read. *p* values *versus* placebo were based on the Cochran-Mantel-Haenszel chi-squared test, stratified by induction study treatment.

[a]Sustained steroid-free modified remission at weeks 24 and 52 was evaluated among patients in modified remission at baseline; sustained steroid-free remission at weeks 24 and 52 was evaluated among patients in remission at baseline.

*$p < 0.05$ *versus* placebo. **$p < 0.01$ *versus* placebo. ***$p < 0.001$ *versus* placebo.

BID, twice daily; CI, confidence interval; FAS, full analysis set; mMayo, modified Mayo; *N*, number of patients in the analysis set; *n*, number of patients achieving endpoint; NRI, nonresponder imputation; ns, not significant; PGA, Physician Global Assessment; TNFi, tumor necrosis factor inhibitor.

Induction 1 and 2, and OCTAVE Sustain, which also showed that the treatment effect was similar regardless of prior TNFi failure, using endpoint definitions based on the Mayo score.[3,12]

While the Mayo score has been a commonly used tool for registration trials, including the OCTAVE studies, the FDA now recommends the use of the mMayo score[2] in order to address the subjectivity and arbitrary scoring associated with the PGA subscore.[13] By excluding this subscore, the mMayo score benefits from reduced subjectivity, with outcomes focused on patient-reported symptoms and objective endoscopic findings. Furthermore, this guidance recommends that trials use the primary endpoint of clinical remission, defined as a stool frequency subscore of 0, a rectal bleeding subscore of 0, and an endoscopic subscore of ≤1. A permitted alternative to a stool frequency subscore of 0 is a stool frequency subscore of ≤1, and a ≥1-point decrease from baseline.[2]

A recent phase IIb randomized controlled trial of the Janus kinase inhibitor upadacitinib for the treatment of moderately to severely active UC included clinical remission based on the mMayo score and the Mayo score as primary and secondary endpoints, respectively, with similar results observed between scoring methods.[7] The phase III clinical trial program went on to measure the primary endpoint of clinical remission as defined by the mMayo score.[14] Furthermore, a recent phase III randomized controlled trial, evaluating the interleukin-12/-23 antagonist ustekinumab for the treatment of moderately to severely active UC, utilized the mMayo score in an alternative definition for the primary endpoint of clinical remission, in addition to clinical remission based on the Mayo score.[6] The study reported that the proportion of patients with clinical remission following induction or maintenance therapy was similar regardless of inclusion of PGA subscore in the assessment, supporting the findings reported in this post hoc analysis of tofacitinib efficacy. This supports the validity of the mMayo score as an accurate measure of clinical disease activity. A previous report has further demonstrated a strong correlation between the total, modified and partial Mayo scores in measuring clinical efficacy, suggesting that modified or partial Mayo scores can be used as a proxy for the total Mayo score in clinical trials and clinical practice.[15] The modified partial Mayo score also facilitates more frequent assessment either virtually or remotely, as in-person visits are no longer required. Moreover, there was consensus among gastroenterologists, colorectal surgeons, methodologists, and clinical trialists to use the mMayo score, omitting PGA, as a primary outcome measure in UC clinical trials.[16]

Of note, while both the upadacitinib and ustekinumab studies used a rectal bleeding subscore of 0 and an endoscopic subscore of ≤1 in their definition of modified clinical remission, their stool frequency criteria differed, using a stool frequency subscore of ≤1 and an absolute stool number of ≤3, respectively.[6,7] The definition of modified remission in this post hoc analysis of data from the OCTAVE studies was the same as that used in the upadacitinib study.

Treatment effect sizes for tofacitinib in the OCTAVE studies were similar across alternative definitions of modified remission, including (in addition to an endoscopic subscore of ≤1 and a rectal bleeding subscore of 0) a stool frequency subscore of ≤1 and no greater than at induction study baseline, and the alternative FDA definition (a stool frequency subscore of ≤1 and a ≥1-point decrease from induction study baseline). When modified remission was defined as requiring a stool frequency subscore of 0, treatment effect sizes were generally smaller for remission at week 8, or week 24 and/or week 52. However, this was not true for sustained steroid-free remission among patients who were in modified remission at baseline of OCTAVE Sustain. Studies have shown that there is a discrepancy between endoscopic and histological remission and patient-reported outcomes, with observational and clinical study data of patients with UC showing that endoscopically inactive disease was not associated with complete normalization of stool frequency.[17,18]

Taken together, we demonstrate that the mMayo score is a robust endpoint assessment of clinical efficacy, supporting its use as an endpoint in UC clinical trials. One disadvantage of both the mMayo and total Mayo scores, however, is that they only focus on two patient-reported outcomes (rectal bleeding, stool frequency). Other patient-reported symptoms, such as stool urgency, and pain, have been perceived as most important by patients with UC[19,20] and therefore should also be considered in clinical trial data to fully capture disease activity pertinent to patients.

A limitation of this analysis is that NRI was applied, a conservative method that avoids bias by assuming that the patient is a nonresponder at the time of trial withdrawal, which may limit the differences between scoring definitions compared with use of observed data. NRI was used in this analysis to align with the primary efficacy analyses of OCTAVE Induction 1 and 2, and OCTAVE Sustain,[3] and for better contextualization with other studies.[6,7]

Conclusion

In conclusion, these results suggest that use of the mMayo scoring definition in OCTAVE Induction 1 and 2, and OCTAVE Sustain, generates efficacy findings that are consistent with the previously reported findings based on Mayo scores. This may help to better contextualize outcomes with tofacitinib therapy with newly developed therapies for treatment of UC and supports the use of mMayo score-based endpoints, instead of Mayo score-based endpoints, in UC trials.

Authors' note

Conference presentation: Some of the data in this manuscript were previously presented at the Crohn's and Colitis Congress, 23–25 January 2020.

Declarations

Ethics approval and consent to participate

All studies were registered with ClinicalTrials.gov (NCT01465763; NCT01458951; NCT01458574) and were conducted in compliance with the Declaration of Helsinki and the International Conference on Harmonization Good Clinical Practice Guidelines, and were approved by the Institutional Review Boards and/or Independent Ethics Committees at each investigational center participating in the studies or at a central Institutional Review Board. All patients provided written informed consent.

Consent for publication

Not applicable. This manuscript does not contain any data from an individual person.

Author contributions

William J. Sandborn: Conceptualization; Writing – original draft; Writing – review & editing.

Bruce E. Sands: Conceptualization; Writing – original draft; Writing – review & editing.

Séverine Vermeire: Conceptualization; Writing – original draft; Writing – review & editing.

Yvette Leung: Conceptualization; Writing – original draft; Writing – review & editing.

Xiang Guo: Conceptualization; Methodology; Writing – original draft; Writing – review & editing.

Irene Modesto: Conceptualization; Methodology; Writing – original draft; Writing – review & editing.

Chinyu Su: Conceptualization; Methodology; Writing – original draft; Writing – review & editing.

Wenjin Wang: Conceptualization; Formal analysis; Writing – original draft; Writing – review & editing.

Julian Panés: Conceptualization; Formal analysis; Writing – original draft; Writing – review & editing.

Competing interests

WJS has received grant support, personal fees, and non-financial support from Pfizer Inc during the conduct of the studies. Outside of the submitted work, WJS has received grant support from AbbVie, Abivax, Arena Pharmaceuticals, Boehringer Ingelheim, Celgene, Eli Lilly,

Availability of data and material

Upon request, and subject to review, Pfizer will provide the data that support the findings of this study. Subject to certain criteria, conditions, and exceptions, Pfizer may also provide access to the related individual de-identified participant data. See https://www.pfizer.com/science/clinical-trials/trial-data-and-results for more information.

Supplemental material

Supplemental material for this article is available online.

References

1. D'Haens G, Sandborn WJ, Feagan BG, *et al.* A review of activity indices and efficacy end points for clinical trials of medical therapy in adults with ulcerative colitis. *Gastroenterology* 2007; 132: 763–786.

2. U.S. Department of Health and Human Services, Food and Drug Administration and Center for Drug Evaluation and Research (CDER). Ulcerative colitis: clinical trial endpoints. Guidance for industry. http://www.fda.gov/downloads/Drugs/Guidances/UCM515143.pdf (2016, accessed 8 Jul 2021).

3. Sandborn WJ, Su C, Sands BE, *et al.* Tofacitinib as induction and maintenance therapy for ulcerative colitis. *N Engl J Med* 2017; 376: 1723–1736.

4. Lichtenstein GR, Loftus EV Jr, Wei SC, *et al.* Tofacitinib, an oral, small-molecule Janus kinase inhibitor, in the treatment of ulcerative colitis: analysis of an open-label, long-term extension study with up to 5.9 years of treatment [Abstract]. *J Crohns Colitis* 2020; 14: DOP61.

5. Feagan BG, Khanna R, Sandborn WJ, *et al.* Agreement between local and central reading of endoscopic disease activity in ulcerative colitis: results from the tofacitinib OCTAVE trials. *Aliment Pharmacol Ther* 2021; 54: 1442–1453.

6. Sands BE, Sandborn WJ, Panaccione R, *et al.* Ustekinumab as induction and maintenance therapy for ulcerative colitis. *N Engl J Med* 2019; 381: 1201–1214.

7. Sandborn WJ, Ghosh S, Panes J, *et al.* Efficacy of upadacitinib in a randomized trial of patients with active ulcerative colitis. *Gastroenterology* 2020; 158: 2139–2149.

8. European Medicines Agency. Guideline on the development of new medicinal products for the treatment of ulcerative colitis. https://www.ema.europa.eu/documents/scientific-guideline/guideline-development-new-medicinal-products-treatment-ulcerative-colitis-revision-1_en.pdf (2018, accessed July 8, 2021).

9. Danese S, Vermeire S, Zhou W, *et al.* OP24 Efficacy and safety of upadacitinib induction therapy in patients with moderately to severely active ulcerative colitis: results from the phase 3 U-ACHIEVE study [Abstract]. *J Crohns Colitis* 2021; 15: OP24.

10. 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.

11. Sands BE, Peyrin-Biroulet L, Marano C, *et al.* Efficacy in biologic-failure and nonbiologic-failure populations in a phase 3 study of ustekinumab in moderate-severe ulcerative colitis: UNIFI [Abstract]. *Gastroenterology* 2019; 156: 833a.

12. Sandborn WJ, Peyrin-Biroulet L, Sharara AI, *et al.* Efficacy and safety of tofacitinib in ulcerative colitis based on prior tumor necrosis factor inhibitor failure status. *Clin Gastroenterol Hepatol* 2022; 20: 591.e8–601.e8.

13. U.S. Department of Health and Human Services, Food and Drug Administration and (CDER). Ulcerative colitis: developing drugs for treatment. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/ulcerative-colitis-developing-drugs-treatment (2022, accessed 22 Aug 2022).

14. Danese S, Vermeire S, Zhou W, *et al.* Upadacitinib as induction and maintenance therapy for moderately to severely active ulcerative colitis: results from three phase 3, multicentre, double-blind, randomised trials. *Lancet* 2022; 399: 2113–2128.

15. Naegeli AN, Hunter T, Dong Y, *et al.* Full, partial, and modified permutations of the Mayo score: characterizing clinical and patient-reported outcomes in ulcerative colitis patients. *Crohns Colitis 360* 2021; 3: otab007.

16. Ma C, Hanzel J, Panaccione R, *et al.* CORE-IBD: a multidisciplinary international consensus initiative to develop a core outcome set for randomized controlled trials in inflammatory bowel disease. *Gastroenterology* 2022; 163: 950–964.

17. Colombel JF, Keir ME, Scherl A, *et al.* Discrepancies between patient-reported outcomes, and endoscopic and histological appearance in UC. *Gut* 2017; 66: 2063–2068.

18. Jharap B, Sandborn WJ, Reinisch W, *et al.* Randomised clinical study: discrepancies between patient-reported outcomes and endoscopic appearance in moderate to severe ulcerative colitis. *Aliment Pharmacol Ther* 2015; 42: 1082–1092.

19. Nag A and Romero B. Development and content validation of patient-reported outcomes tools for ulcerative colitis and Crohn's disease in adults with moderate-to-severe disease. *Health Qual Life Outcomes* 2022; 20: 75.

20. Teich N, Schulze H, Knop J, *et al.* Novel approaches identifying relevant patient-reported outcomes in patients with inflammatory bowel diseases—LISTEN 1. *Crohns Colitis 360* 2021; 3: otab050.

Visit SAGE journals online journals.sagepub.com/home/tag

SAGE journals