Full, Partial, and Modified Permutations of the Mayo Score: Characterizing Clinical and Patient-Reported Outcomes in Ulcerative Colitis Patients

April N. Naegeli, DrPH, MPH,* Theresa Hunter, PhD,* Yan Dong, PhD,* Ben Hoskin, BSc,[†] Chloe Middleton-Dalby, MRes,[†] James Hetherington, BSc,[†] Diana Stefani-Hunyady, MD, PhD,* and James B. Canavan, MD, PhD*

Background: Understanding ulcerative colitis disease activity assessed via the full, modified, or partial Mayo Score may help clinicians apply results from clinical trials to practice and facilitate interpretation of recent and older studies.

Methods: Mayo Score variables were assessed in a cross-sectional study of 2608 ulcerative colitis patients.

Results: Permutations of Mayo Scores were highly correlated, and models predicting the omitted variable from each permutation demonstrated significant agreement between predicted and observed values.

Conclusions: Partial/modified Mayo Scores may be used to predict endoscopic and Physician's Global Assessment scores, and serve as proxies for the full Mayo Score in clinical practice/trials.

Lay Summary

Assessing disease activity in patients with ulcerative colitis can be challenging. This study shows that disease activity can sometimes be evaluated without endoscopy or Physician's Global Assessment, and helps physicians interpret results of such evaluation.

Key Words: ulcerative colitis, patient-reported outcome measures, Mayo Disease Activity Score, partial Mayo Score, real world

INTRODUCTION

Ulcerative colitis (UC) is a complex, multifactorial disease characterized by inflammation of the colon and a chronic, intermittent relapsing–remitting disease course.[1–3] Symptoms include bloody diarrhea, abdominal pain, bowel urgency, and tenesmus.[2,3] The worldwide incidence and prevalence of UC are increasing, as is UC-related healthcare expenditure.[1,4] As such, the development of safe and effective treatments for UC is a priority. Accurate assessment of disease activity is important for evaluating the efficacy of investigational products in clinical trials and for assessing patients in clinical practice.

The Mayo Score was developed as a composite disease activity index for use in clinical trials.[5] The original description of the Mayo Score included an assessment of 2 patient-reported outcomes [PROs; stool frequency (SF) and rectal bleeding (RB)], the endoscopic appearance of the mucosa (endoscopic score, ES), and a Physician's Global Assessment (PGA), each of which were scored on a scale from 0 to 3, giving a maximum total score of 12.[5]

The Food and Drugs Administration (FDA) currently accepts the modified Mayo Score (mMayo; including SF, RB, and ES, but not PGA) for the assessment of disease activity

in pivotal UC clinical trials. This is subject to 2 modifications from the original description of the Mayo Score: the definition of an ES of 1 no longer includes mucosal friability and the PGA is no longer accepted as part of that composite index.[6,7] The mMayo Score gives a maximum total score of 9. However, the relationship between the mMayo Score and the full Mayo Score is not well defined. This is needed to facilitate side-by-side interpretation of older UC clinical trials that use full Mayo Score-based endpoints and more recent clinical trials that use mMayo Score-based endpoints.

In clinical practice, endoscopy is not performed at every visit. The partial Mayo Score (pMayo; including SF, RB, and PGA, but not ES) was initially developed to address this limitation and has subsequently been used for endpoint assessment in real-world observational cohort studies in UC.[8–10] Lewis et al showed that the pMayo Score correlated well with the full Mayo Score.[8] However, the relationship between the pMayo Score and the mMayo Score is poorly defined. An understanding of this is required to aid the interpretation and application of mMayo Score-based clinical trial results to patients in clinical practice, whose disease activity may be assessed using the Mayo Score or pMayo Score.

We used data from the Adelphi UC Disease Specific Programme (DSP), a large real-world database, to evaluate the relationship between individual components of the Mayo Score, and the relationship between disease activity scored by the full 12-point Mayo Score, the mMayo Score, and the pMayo Score, in patients with active UC. This study also explored if the score for the "missing variable" in the mMayo Score (ie, PGA) or pMayo Score (ie, ES) could be predicted from the 9-point score for those instruments. Finally, the study explored the association between changes in Mayo Scores and changes in selected clinical and health-related quality of life (HRQoL) outcomes using real-world data as available in the DSP, and attempted to model a minimal clinically important difference in Mayo Score change as determined by the EuroQol 5-dimension 5-level questionnaire (EQ-5D) and Short Inflammatory Bowel Disease Questionnaire (SIBDQ) to provide context to the Mayo Scores.

MATERIALS AND METHODS

Data Source

The Adelphi UC DSP is a noninterventional, cross-sectional, real-world survey of patients with UC and their treating gastroenterologists. The DSP includes patient- and physician-reported data, such as demographics, concomitant medication, UC symptoms and disease activity, and HRQoL. The DSP methodology has been described previously.[11]

Study Population

The Adelphi UC DSP from September 2017 to January 2018 included 100 gastroenterologists from the United States and 280 gastroenterologists from 5 European countries (France, Germany, Italy, Spain, and the United Kingdom; EU5). Treating gastroenterologists had at least 4 years of experience and were actively involved in the management of UC patients (seeing at least 7 UC patients per month).

Patients were at least 18 years of age with a diagnosis of UC, as assessed by the treating gastroenterologist. There was no requirement to perform any diagnostic test or assessments prior to inclusion. Patients had to have one of the following criteria: currently or previously received a corticosteroid, immunomodulator, or a biologic for UC; or classified at any point in their disease history as having moderate/severe disease; or received a Mayo Score of ≥4 in previous assessments. For the purpose of this study, biologics were defined as adalimumab, infliximab, certolizumab-pegol, golimumab, vedolizumab, natalizumab, ustekinumab, and biosimilars of these originator biologics.

Treating gastroenterologists completed questionnaires for 2608 patients, of which 1061 patients completed at least 1 HRQoL instrument.

Physician- and Patient-Reported Data

The Mayo Score comprises 4 categories (RB, SF, ES, and PGA), each rated with a score from 0 to 3 with a total score ranging from 0 to 12, and higher scores indicating greater disease severity.[5] Each of these parameters was derived from physician-reported information. The full Mayo Score was calculated by adding the score for each of these 4 parameters. The mMayo was calculated by adding the scores for RB, SF, and ES. The pMayo was calculated by adding the scores for RB, SF, and PGA.

The EQ-5D (a change of ∼0.07–0.11 can be considered clinically meaningful),[12,13] EuroQol Visual Analogue Scale (VAS), Work Productivity and Activity Impairment Questionnaire (WPAI),[14] and the SIBDQ (a change of 9 can be considered clinically meaningful)[15] were completed by patients. These instruments are described further in Supplementary Information.

Data Analysis and Statistical Methods

Descriptive statistics were used in this study. Continuous variables are presented as mean (SD) or median (minimum and maximum, or interquartile range). Categorical variables are presented as sample size, number of missing responses, and the number and percentage in each category. Confidence intervals or P values are provided as appropriate. Where statistical tests were performed, P values <0.05 were considered statistically significant.

The real-world nature of the DSP means that some variables contain missing data. Missing data were not imputed. The base sample size (n) for each variable was reported to enable assessment or calculation of the number of missing patients.

Linear regression (continuous outcomes) or ordered logistic regression (ordinal outcomes) was used. For count data (such as number of flares in the past year), negative binomial regression was used as appropriate. A number of covariates were assessed with various regression models, including age, time since diagnosis, disease severity, body mass index, Charlson comorbidity index, current treatment with a biologic and physician satisfaction, and were used to examine how changes in each Mayo Score individually affected clinical and PROs. To provide context to the Mayo Scores, and to estimate what may be considered as meaningful to a patient, the change in Mayo Score required to reach the published minimally important difference in EQ-5D and SIBDQ was calculated from this regression analysis.

Correlations between the Mayo permutations and subscores were explored using Spearman rho. Ordered logistic regression using a proportional odds model was used to predict the omitted item from the pMayo and mMayo, and weighted kappa analysis was used to measure agreement between predicted and observed values. The probability of a certain Mayo ES given a certain pMayo Score was predicted using the Stata post estimate command margins for each level of pMayo.

Ethical Considerations

Adelphi UC DSP questionnaires used in this study were approved by the Freiburger Ethik-Kommission International institutional review board (IRB) (Freiburg, Germany; IRB reference 017/1679) for EU5 participants and the Western IRB (Olympia, WA, USA; IRB reference 09-25-2017) for US participants. Consent was obtained from all patients whose physician-reported outcome or PRO data were used in this analysis.

RESULTS

Patient Demographic and Clinical Characteristics

Demographics and clinical data were available for 2608 patients (Table 1). In total, 1180 patients (45.3%) were female. Mean (SD) age was 40.8 (14.9) years, with a median (interquartile range) time from diagnosis of 1.9 years (0.5, 5.2). Treating gastroenterologists reported that over half of patients (54.0%) had mild disease, 44.2% of patients had moderate disease, and 3.7% of patients had severe disease. In comparison, at diagnosis, 10.5% of patients were classified as mild, 70.2% of patients as moderate, and 19.3% of patients as severe. Most patients were assessed to have a stable clinical course (58.2%), while 32.2% of patients were improving and 9.6% of patients deteriorating. In total, 61.0% of patients had experienced a flare since diagnosis

TABLE 1. Patient Demographics, Clinical Characteristics, and Current Medication Use

Parameter (N = 2608 Unless Otherwise Indicated)	
Age, years	
Mean (SD)	40.75 (14.86)
Female, n (%)	1180 (45.3)
BMI, kg/m^2, n	2607
Mean (SD)	24.35 (3.80)
Physician severity: current, n (%)	
Mild	1407 (54.0)
Moderate	1105 (42.4)
Severe	96 (3.7)
Physician severity: at diagnosis, n (%)	
Mild	275 (10.5)
Moderate	1831 (70.2)
Severe	502 (19.3)
Time since diagnosis, years, n	2003
Median (IQR)	1.9 (0.5, 5.2)
Patient in remission, n (%)	
Not in remission	1172 (44.9)
In remission	1436 (55.1)
Current disease progression, n (%)	
Improving	839 (32.2)
Stable	1519 (58.2)
Deteriorating	250 (9.6)
Patient flaring, n	2356
Never flared, n (%)	920 (39.1)
Flared, n (%)	1436 (61.0)
Current flaring, n	2278
None, n (%)	2010 (88.2)
Currently flaring, n (%)	268 (11.8)
No. flares in the past year, n	2278
0, n (%)	1375 (60.4)
1, n (%)	593 (26.0)
2, n (%)	235 (10.3)
3, n (%)	63 (2.8)
4+, n (%)	12 (0.5)
Current medication use, n	2604
5-ASA, n (%)	1567 (60.2)
Biologic/biosimilar, n (%)	959 (36.8)
Immunomodulator, n (%)	724 (27.8)
Prednisolone, n (%)	364 (14.0)
Budesonide, n (%)	78 (3.0)
Other corticosteroids, n (%)	33 (1.3)
Other, n (%)	32 (1.2)

Biologic included the following medications: adalimumab, infliximab, certolizumab-pegol, golimumab, vedolizumab, natalizumab, ustekinumab, or biosimilars of these originator biologics.

5-ASA, 5-aminosalicylic acid; BMI, body mass index; IQR, interquartile range.

as determined by the treating gastroenterologist and 11.8% of patients were experiencing a flare at the time of data collection. The most frequently used UC medications included 5-aminosalycilates (5-ASA, 60.2%), biologics or biosimilars (36.8%), immunomodulators (27.8%), and corticosteroids (18.3%). Current symptoms and comorbidities are shown in Supplementary Tables S1 and S2, respectively.

Distribution of Mayo Score and Its Variants in UC Population

The mean (SD) values for the full Mayo, pMayo, and mMayo Scores were 4.07 (2.99), 2.92 (2.20), and 2.83 (2.24), respectively. The distribution of full Mayo, pMayo, and mMayo Scores are shown in Figure 1.

The full Mayo and pMayo Scores, and the full Mayo and mMayo Scores, were highly correlated (Spearman rho = 0.986, $P < 0.001$ for each comparison; Fig. 2), as were the individual parameters of ES and PGA (rho = 0.8837), ES and RB (rho = 0.6752), and ES and SF (rho = 0.6678, $P < 0.001$ for each comparison; Fig. 3). The correlation between SF and RB was moderate (rho = 0.5928, $P < 0.001$; Fig. 3).

Ordered logistic regressions with a partial proportional odds model were used to predict ES from the pMayo, and PGA from the mMayo, respectively (pMayo model pseudo $R^2 = 0.4497$, $P < 0.001$; mMayo model pseudo $R^2 = 0.4622$, $P < 0.001$). Using the estimated margins from the model predicting ES from pMayo (Fig. 4), this model predicted that a pMayo Score of 0–1 was likely to be associated with an ES of 0, a pMayo Score of 2–3 with an ES of 1, a pMayo Score of 4–7 with an ES of 2, and a pMayo Score of 8–9 with an ES of 3 (Fig. 4A). Weighted kappa analysis showed substantial agreement between the values predicted by the models and those reported by the physicians. Agreement between model-predicted and physician-reported endoscopic results from the pMayo was
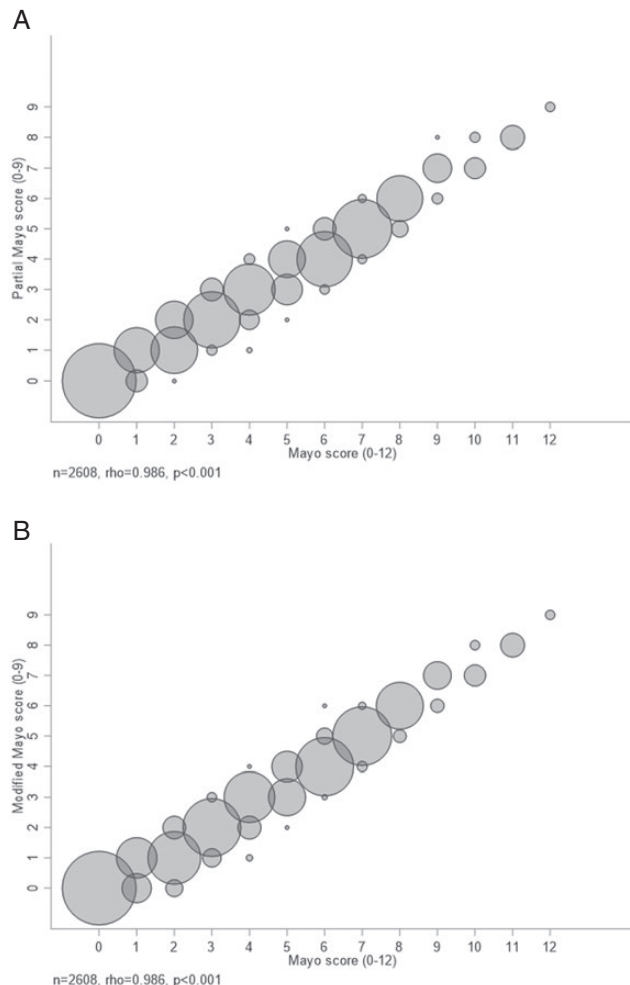
FIGURE 2. A comparison between the full Mayo and pMayo Score (A), and between the full Mayo and mMayo Score (B). The size of each circle is proportionate to the number of patients with the value at the center of each circle.
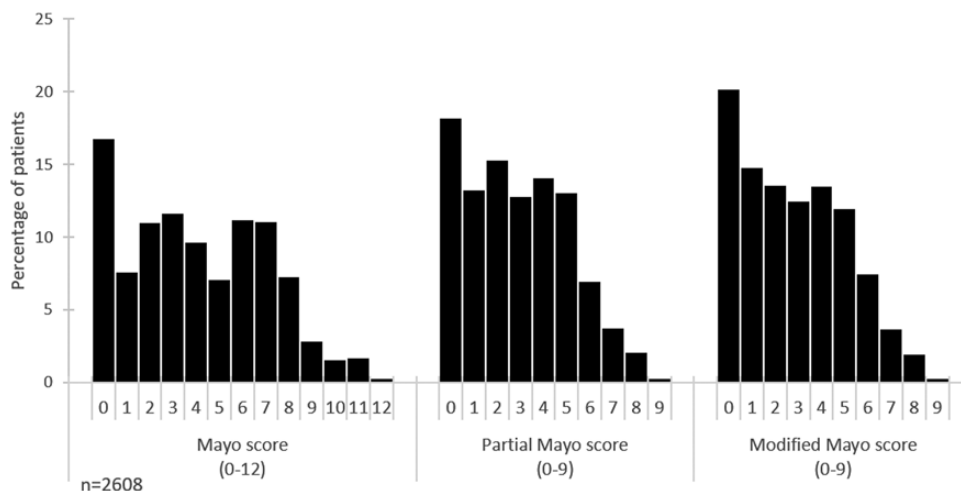
FIGURE 1. Percentage of patients with each score for Mayo, pMayo, and mMayo disease activity indices.
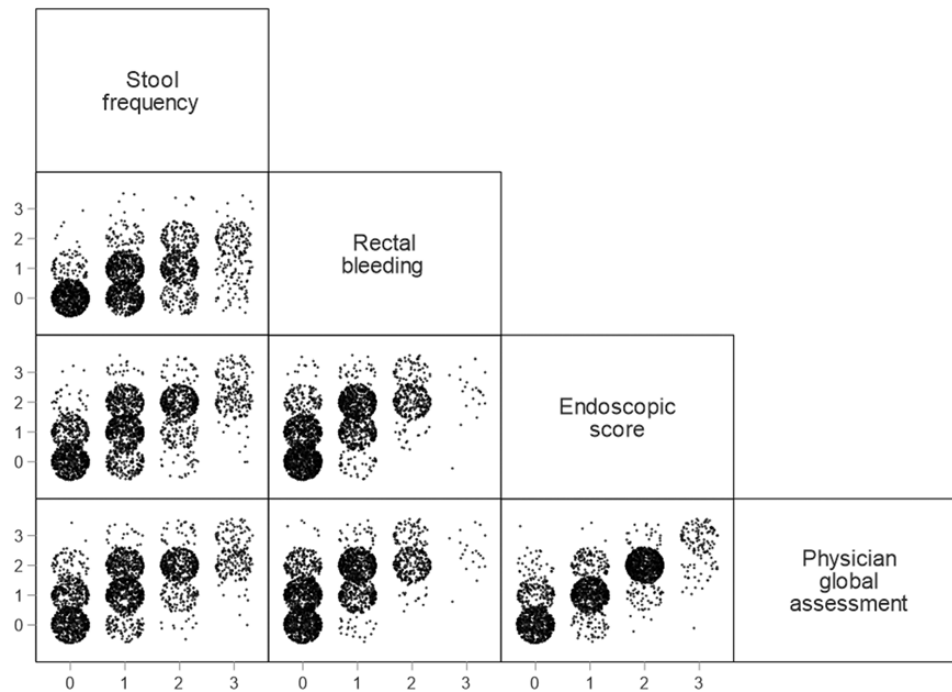
FIGURE 3. Scatter plots showing associations between different components of the full Mayo Score.

substantial, with 71.9% of results in agreement even when un-weighted. Values of kappa were 0.6994, *P* < 0.001.

Similarly, predicting PGA from mMayo showed that pre-dicted mMayo Scores of 0 were associated with a PGA of 0, mMayo Scores of 1–3 with a PGA of 1, mMayo Scores of 4–7 with a PGA of 2, and mMayo Scores of 8–9 with a PGA of 3 (Fig. 4B). Again, substantial agreement was observed between predicted and reported PGA results from the mMayo with 75.7% of results in agreement (kappa = 0.7160, *P* < 0.001). The substantial level of agreement between predicted and observed values of ES and PGA indicated that the omitted item from the mMayo and pMayo permutations could be predicted from the included items.

Association Between Mayo Score and Clinical Outcomes

The association between increases in the full Mayo, pMayo, and mMayo Scores, and changes in physician-reported outcomes and PROs (Table 2) are shown in Table 3. Point-increases in each of the permutations of the Mayo Score were associated with increases in the risk of adverse physician-reported outcomes including a current flare, the number of flares in past year, deterioration in clinical status, and patient-reported overall work productivity impairment. Point-increases in each of the permutations of the Mayo Score were also associated with decreases in the EQ-5D index and VAS, and in the SIBDQ, where lower scores in each of

these indices represent a lower HRQoL. A 1-point increase in pMayo or mMayo was associated with a 0.02-unit decrease in EQ-5D, suggesting that a change of pMayo or mMayo Score of ≥4 may be associated with a clinically meaningful reduc-tion in HRQoL, as measured by EQ-5D. Similarly, a 1-point increase in pMayo or mMayo was associated with a 2.18 and 2.73 point decrease in SIBDQ, respectively, suggesting that a change of pMayo or mMayo Score of ≥4 may be associated with a clinically meaningful reduction in HRQoL, as meas-ured by SIBDQ.

DISCUSSION

Accurate assessment of disease activity is important for evaluating the efficacy of investigational products in UC clinical trials and assessing UC patients in clinical practice. Understanding the relationship between disease activity scores employed in clinical trials and clinical practice is needed to in-terpret clinical trial results and apply them to patient care.

The Mayo Score was developed as an instrument to measure UC disease activity in clinical trials.[8] The mMayo Score was developed from the full Mayo Score by making 2 changes from the original description. Firstly, mucosal fria-bility was removed from the definition of ES = 1, due to the difficulties in distinguishing between mild friability (ES = 1) and friability (ES = 2), and because the presence of friability was not consistent with the definition of clinical remission that includes a Mayo ES of 0–1.[9,10] Secondly, PGA was removed from the mMayo Score, as the concept that the PGA purported
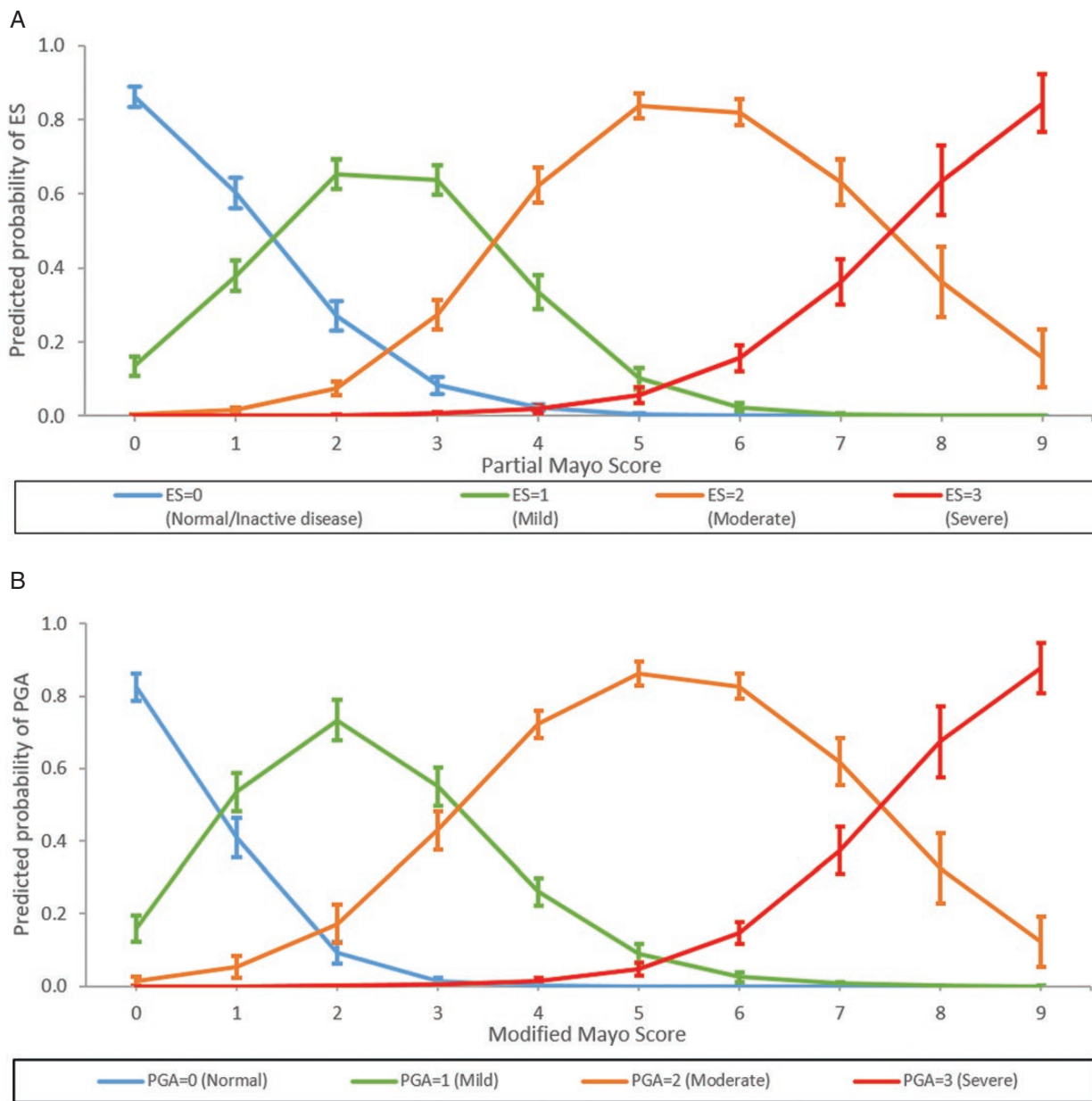
A

B

FIGURE 4. Prediction of ES from the pMayo Score (A) and prediction of PGA from the mMayo Score (B). The figure shows the model-predicted probability that a patient will be at each level of ES/PGA, for any given pMayo/mMayo Score and the predictions show mean probabilities 95% confidence intervals.

to measure were not distinct from the other components of the Mayo Score: PROs (SF and RB) and objective findings (ES).[6] The pMayo Score was developed as a composite index that did not require endoscopy, as colonoscopy or flexible sigmoidoscopy are performed less frequently than clinic visits, are expensive, relatively unpleasant for a patient and may be associated with complications. It has subsequently been used for endpoint assessment in real-world observational studies in UC.[8–10] Lewis et al showed that the pMayo Score correlated well with the full Mayo Score, suggesting that the pMayo Score has a

role in the noninvasive assessment of disease activity in clinical practice.[8] However, the relationship between the full Mayo and the mMayo Scores, and between the pMayo and the mMayo Scores is not well defined. This study used real-world evidence to confirm the strong correlation between the full and pMayo Scores and extended this finding by also showing strong correlations between the full Mayo, mMayo, and pMayo Score. Furthermore, this study identified values of the mMayo and pMayo Scores that correspond with values of the full Mayo Score, providing guidance for inter-interpretation between

TABLE 2. PROs

PRO	
EQ-5D index, n	1033
Mean (SD)	0.84 (0.15)
EQ-5D VAS, n	1047
Mean (SD)	74.92 (15.90)
WPAI overall work impairment, n	471
Mean (SD)	28.52 (28.38)
SIBDQ, n	1017
Mean (SD)	51.41 (11.34)

EQ-5D, EuroQol 5-dimension 5-level questionnaire, cross-walked to the EuroQol 5-dimension 3-level questionnaire.

TABLE 3. Association Between the Full Mayo, pMayo, and mMayo Score With Clinical or PROs

	Mayo	pMayo	mMayo
Physician-reported outcome			
Current flaring, n = 1821			
OR (SE)	1.41 (0.08)	1.58 (0.12)	1.52 (0.10)
No. flares in past year, n = 1821			
IRR (SE)	1.13 (0.02)	1.16 (0.03)	1.17 (0.03)
Disease Progression, n = 2000			
OR (SE)	1.41 (0.07)	1.56 (0.11)	1.48 (0.10)
PROs			
EQ-5D index, n = 860			
Score (SE)	−0.016 (0.003)	−0.021 (0.004)	−0.020 (0.003)
EQ-5D VAS, n = 870			
Score (SE)	−1.748 (0.283)	−2.363 (0.374)	−2.248 (0.358)
WPAI, n = 382			
Score (SE)	5.175 (0.745)	6.913 (1.006)	6.940 (0.888)
SIBDQ, n = 849			
Score (SE)	−2.105 (0.221)	−2.817 (0.299)	−2.728 (0.273)

Disease Progression: physician-perceived worsening of disease progression (order of responses: improving, stable, and deteriorating). The table shows the increase or decrease in the odds ratio, incidence rate ratio, or mean score with a change in Mayo Score of 1 unit.

EQ-5D, EuroQol 5-dimension 5-level questionnaire, cross-walked to the EuroQol 5-dimension 3-level questionnaire; IRR, incidence rate ratio; OR, odds ratio; SE, standard error.

these permutations of the Mayo Score, which may aid the application of clinical trial results to real-world clinical settings.

The Adelphi UC DSP is a large, real-world database that includes both patient- and physician-reported data. The mean age of patients, their range of symptoms, and the number and type of comorbidities were broadly comparable to those noted in other real-world populations, and were also similar to those seen in previous DSP populations.[16–18] The frequency of 5-ASA

use was similar to that seen in other studies.[16,18] The higher rate of biologic use, compared to immunomodulators, suggests appreciable use of biologic monotherapy. Approximately 96% of the study population had mild–moderate disease, which may limit the interpretability of these data for patients with severe UC. The change in the prevalence of severe UC over time (~19% at diagnosis to ~4% at data acquisition) suggests that severe UC is recognized and treated in clinical practice. In total, 42% of patients had moderate disease, 12% of patients were flaring at the time of survey completion and 40% of patients experienced a flare in the preceding year, suggesting that significant unmet need still exists in the treatment of patients in this real-world cohort.

We showed that a strong correlation exists between the pMayo, mMayo, and full Mayo Scores, and HRQoL indices, and extends the association established between PROs, disease activity, and HRQoL seen previously in recent studies.[19–21] Our study presents ranges of the full Mayo Score that corresponds to levels of the pMayo and mMayo Score, respectively, which may assist in the inter-interpretability of these scores and the application of clinical trial results to real-world clinical practice. This study also presents the probability of a Mayo ES given a pMayo Score, which may assist in the assessment of disease activity in clinical practice. Use of the PGA subscore is not recommended by the FDA as a component of an endpoint measure to support a marketing application, because the concept it purports to measure is not distinct from the other components of the Mayo Score.[6] However, we showed that the PGA correlated strongly with Mayo endoscopic findings, suggesting that a PGA by an experienced gastroenterologist continues to have utility in clinical practice.

Our results show that greater UC disease activity is associated with worse clinical outcomes, including flare, a reduction in HRQoL, as measured by EQ-5D and SIBDQ, and greater work impairment, as assessed by WPAI, which has been observed in other UC populations.[13,22–24] Previous studies have suggested that a reduction of ≥3 points on the Mayo and pMayo Score represent a clinically meaningful change.[8] A change from baseline of ~0.07–0.11 in EQ-5D has been estimated as clinically meaningful in other UC populations.[13] Similarly, a change from baseline of 9 in SIBDQ has been estimated to be clinically meaningful.[15] Extrapolating the associations between change in EQ-5D or SIBDQ and change in permutations of the Mayo Score, this study supports this finding, with changes of ≥4 units in Mayo permutations associated with clinically meaningful differences in EQ-5D and SIBDQ. Little differences in effect were observed between the pMayo and mMayo Scores. The unit changes in Mayo were associated with consistently smaller changes in PROs than for the pMayo and mMayo Scores, which can be explained by the Mayo Score having a greater range than the pMayo and mMayo, and therefore any given unit change in Mayo Score is a smaller step than in the other permutations, rather than a fundamental difference between scores.

Strengths of the current study include the use of a well-defined RWE dataset with prospective acquisition of clinical and HRQoL data from a large number of UC patients, using established and validated methodology. However, this study has limitations that may impact the observed results. Study patients were cared for by experienced, inflammatory bowel disease (IBD)-focused gastroenterologists who actively managed at least 7 UC patients per month and their clinical experience in IBD care may have impacted prescribing patterns.[25] Physician selection may have also been influenced by unmeasured factors, such as willingness to participate, capacity, practice setting, and geographical considerations. Taken together, these may be sources of selection bias and may limit the applicability of our findings to practices that see a low volume of IBD patients.

Patient participation in this survey may have been subject to motivational bias, and both patient and physician completion of study questionnaires may have been subject to recall bias. It is also possible that a small degree of random misclassification of diagnosis may have occurred in this dataset, as 5/2608 patients were reported to have received biologics that were not approved for the treatment of UC at the time of use, including ustekinumab (n = 2), certolizumab-pegol (n = 1), and natalizumab (n = 1).

Given these limitations, the predictive findings outlined above should be confirmed using data from an independent RWE UC patient cohort, or independently validated in a substudy that is nested within a prospective, adequate, and well-controlled clinical trial.

CONCLUSIONS

Our study demonstrates that the pMayo, mMayo, and full Mayo Scores are highly correlated, with increased pMayo Score associated with higher ES and increased mMayo Score associated with higher PGA in UC patients treated in real-world settings. Furthermore, a >3–4-unit change in Mayo Score was established as a clinically meaningful difference in relation to EQ-5D and SIBDQ. These findings support the use of the pMayo Score and the mMayo Score in clinical practice as proxies for the full Mayo Score and also help contextualize changes in Mayo Score in relation to meaningful impact on HRQoL. These findings should be confirmed in independent RWE datasets, or validated in prospective clinical trials.

SUPPLEMENTARY MATERIAL

Supplementary data are available at *Crohn's & Colitis 360* online.

DATA AVAILABILITY

All data supporting the survey are the intellectual property of "Adelphi Real World" and can be made available upon request to ben.hoskin@adelphigroup.com.

REFERENCES

1. Ungaro R, Mehandru S, Allen PB, et al. Ulcerative colitis. *Lancet.* 2017;389:1756–1770.
2. Feuerstein JD, Cheifetz AS. Ulcerative colitis: epidemiology, diagnosis, and management. *Mayo Clin Proc.* 2014;89:1553–1563.
3. Danese S, Fiocchi C. Ulcerative colitis. *N Engl J Med.* 2011;365:1713–1725.
4. Ananthakrishnan AN, Kaplan GG, Ng SC. Changing global epidemiology of inflammatory bowel diseases: sustaining health care delivery into the 21st century. *Clin Gastroenterol Hepatol.* 2020;18:1252–1260. PMID:32007542. doi:10.1016/j.cgh.2020.01.028
5. Schroeder KW, Tremaine WJ, Ilstrup DM. Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. A randomized study. *N Engl J Med.* 1987;317:1625–1629.
6. Food and Drug Administration. Guidance for Industry. Ulcerative Colitis: Clinical Trial Endpoints. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM515143.pdf (1 November 2019 date last accessed).
7. Scherl EJ, Pruitt R, Gordon GL, et al. Safety and efficacy of a new 3.3 g b.i.d. tablet formulation in patients with mild-to-moderately-active ulcerative colitis: a multicenter, randomized, double-blind, placebo-controlled study. *Am J Gastroenterol.* 2009;104:1452–1459.
8. Lewis JD, Chuai S, Nessel L, et al. Use of the noninvasive components of the Mayo score to assess clinical response in ulcerative colitis. *Inflamm Bowel Dis.* 2008;14:1660–1666.
9. Amiot A, Filippi J, Abitbol V, et al. Effectiveness and safety of ustekinumab induction therapy for 103 patients with ulcerative colitis: a GETAID multicentre real-world cohort study. *Aliment Pharmacol Ther.* 2020;51:1039–1046.
10. Honap S, Chee D, Chapman TP, et al. Real-world effectiveness of tofacitinib for moderate to severe ulcerative colitis: a multicentre UK experience. *J Crohns Colitis.* 2020;14:1385–1393.
11. Anderson P, Benford M, Harris N, et al. Real-world physician and patient behaviour across countries: disease-specific programmes—a means to understand. *Curr Med Res Opin.* 2008;24:3063–3072.
12. Dolan P. Modeling valuations for EuroQol health states. *Med Care.* 1997;35:1095–1108.
13. Stark RG, Reitmeir P, Leidl R, et al. Validity, reliability, and responsiveness of the EQ-5D in inflammatory bowel disease in Germany. *Inflamm Bowel Dis.* 2010;16:42–51.
14. Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics.* 1993;4:353–365.
15. Irvine EJ, Zhou Q, Thompson AK. The Short Inflammatory Bowel Disease Questionnaire: a quality of life instrument for community physicians managing inflammatory bowel disease. CCRPT Investigators. Canadian Crohn's Relapse Prevention Trial. *Am J Gastroenterol.* 1996;91:1571–1578.
16. Brady JE, Stott-Miller M, Mu G, et al. Treatment patterns and sequencing in patients with inflammatory bowel disease. *Clin Ther.* 2018;40:1509–1521.e5.
17. Naegeli A, Zhang X, Hunter T, et al. Full, partial or modified: understanding relationships between permutations of the Mayo score in real-world ulcerative colitis patients. 2018; Program No. P1381. ACG 2018 Annual Scientific Meeting Abstracts. Philadelphia, PA: American College of Gastroenterology.
18. Rubin DT, Mody R, Davis KL, et al. Real-world assessment of therapy changes, suboptimal treatment and associated costs in patients with ulcerative colitis or Crohn's disease. *Aliment Pharmacol Ther.* 2014;39:1143–1155.
19. Aniwan S, Bruining DH, Park SH, et al. The combination of patient-reported clinical symptoms and an endoscopic score correlates well with health-related quality of life in patients with ulcerative colitis. *J Clin Med.* 2019;8:1171. PMID:31387259. PMCID:PMC6723355. doi:10.3390/jcm8081171
20. Taleban S, Stewart KO, Li DK, et al. Clinical activity and quality of life indices are valid across ulcerative colitis but not Crohn's disease phenotypes. *Dig Dis Sci.* 2016;61:2627–2635.
21. Theede K, Kiszka-Kanowitz M, Nordgaard-Lassen I, et al. The impact of endoscopic inflammation and mucosal healing on health-related quality of life in ulcerative colitis patients. *J Crohns Colitis.* 2015;9:625–632.
22. Gibson PR, Vaizey C, Black CM, et al. Relationship between disease severity and quality of life and assessment of health care utilization and cost for ulcerative colitis in Australia: a cross-sectional, observational study. *J Crohns Colitis.* 2014;8:598–606.
23. Vaizey CJ, Gibson PR, Black CM, et al. Disease status, patient quality of life and healthcare resource use for ulcerative colitis in the UK: an observational study. *Frontline Gastroenterol.* 2014;5:183–189.
24. Jowett SL, Seal CJ, Barton JR, et al. The short inflammatory bowel disease questionnaire is reliable and responsive to clinically important change in ulcerative colitis. *Am J Gastroenterol.* 2001;96:2921–2928.
25. Swaminath A, Lebwohl B, Capiak KM, et al. Practice patterns in the use of anti-tumor necrosis factor alpha agents in the management of Crohn's disease: a US national practice survey comparing experts and non-experts. *Dig Dis Sci.* 2011;56:1160–1164.