Effects of vedolizumab on health-related quality of life in patients with ulcerative colitis: results from the randomised GEMINI 1 trial

B. G. Feagan*, H. Patel[†], J.-F. Colombel[‡], D. T. Rubin[§], A. James**, R. Mody[††] & K. Lasch[‡‡]

*Robarts Clinical Trials, Robarts Research Institute, University of Western Ontario, London, ON, Canada.
[†]Immensity Consulting, Inc., Chicago, IL, USA.
[‡]Icahn School of Medicine at Mount Sinai, New York, NY, USA.
[§]University of Chicago Medicine Inflammatory Bowel Disease Center, Chicago, IL, USA.
**Takeda Development Centre Europe Ltd., London, UK.
[††]Takeda Development Center Americas, Inc., Deerfield, IL, USA.
[‡‡]Takeda Pharmaceuticals USA, Inc., Deerfield, IL, USA.

Correspondence to:
Dr B. G. Feagan, Robarts Clinical Trials Inc., 100 Dundas St, Suite 200, London, ON, Canada N6A 5B6.
E-mail: brian.feagan@robartsinc.com

SUMMARY

Background
Health-related quality of life (HRQL) is often diminished in patients with ulcerative colitis.

Aim
To evaluate the effects of vedolizumab on HRQL in patients with ulcerative colitis.

Methods
Using maintenance phase data from the GEMINI 1 study, an analysis of covariance model was used to calculate mean differences between the vedolizumab and placebo groups in changes from baseline to week 52 for 3 HRQL instruments: The Inflammatory Bowel Disease Questionnaire (IBDQ), 36-Item Short Form Health Survey (SF-36), and EQ-5D. Proportions of patients meeting minimal clinically important difference (MCID) thresholds for changes on these instruments were compared between treatment groups for the overall population and for clinically important subgroups. Concordance between clinical remission and remission defined using IBDQ scores was examined.

Results
Compared with placebo-treated patients, vedolizumab-treated patients had greater improvements (152–201%) in IBDQ, EQ-5D visual analogue scale (VAS), and EQ-5D utility scores. Greater proportions (6.9–19.9%) of vedolizumab-treated patients than placebo-treated patients met MCID thresholds for all the instruments. Vedolizumab-treated patients with lower baseline disease activity and those without prior tumour necrosis factor (TNF) antagonist failure had greater HRQL improvements. Among 127 patients with clinical remission based on complete Mayo Clinic scores, >80% also had IBDQ remission; >70% of the 150 patients with IBDQ remission demonstrated clinical remission.

Conclusions
Vedolizumab therapy was associated with significant improvements in HRQL measures compared with placebo. Benefits were greater in patients with lower disease activity and no prior TNF antagonist failure.

INTRODUCTION

Ulcerative colitis is a chronic inflammatory disease characterised by episodes of active disease interspersed with periods of remission.[1, 2] Its symptoms, which include rectal bleeding, increased stool frequency, and abdominal cramps, adversely affect health-related quality of life (HRQL).[1, 2] Results from a large European survey indicated that approximately three-quarters of patients with ulcerative colitis reported that symptoms interfere with their ability to enjoy leisure activities; almost 70% stated that symptoms negatively affect work performance.[3] Although disease activity is strongly inversely correlated with HRQL as assessed by both generic and disease-specific instruments,[1, 2, 4–6] it does not fully account for HRQL status; some patients report poor HRQL even during periods of low disease activity.[2, 7, 8] Numerous studies have suggested that perceived stress level, anxiety or depression, female sex, coexisting fatigue, and number of relapses are independent determinants of low HRQL in patients with inflammatory bowel disease[9–16]; however, this area remains poorly understood.

Vedolizumab (ENTYVIO; Takeda Pharmaceuticals America, Inc.; Deerfield, IL, USA) is a gut-selective anti-$\alpha_4\beta_7$ integrin monoclonal antibody that is approved in multiple jurisdictions for treatment of adults with moderately to severely active ulcerative colitis. The efficacy and safety of vedolizumab therapy for ulcerative colitis were established in GEMINI 1, a phase 3, randomised, double-blind, placebo-controlled, 52-week study with induction and maintenance phases.[17] Although multiple disease-specific and generic HRQL instruments were used in GEMINI 1, the effects of vedolizumab therapy on HRQL have not been examined thoroughly. Specifically, the magnitude of HRQL changes for individual instrument domains and data from clinically important subgroups have not been reported.

Our objectives were to use maintenance phase data from GEMINI 1 to evaluate (i) changes from baseline in HRQL by treatment group using disease-specific and generic instruments; (ii) effects of vedolizumab overall and in clinically important subgroups based on disease severity at baseline and prior tumour necrosis factor (TNF) antagonist failure status; (iii) proportions of patients by treatment group with clinically meaningful improvements from baseline in measures of HRQL; and (iv) the degree of concordance between remission as defined by Mayo Clinic scores and remission defined by Inflammatory Bowel Disease Questionnaire (IBDQ)

scores and examine potential explanations for discordance between these metrics.

MATERIALS AND METHODS

Study design

The methods of GEMINI 1 have been described previously (ClinicalTrials.gov number NCT00783718).[17] The study protocol, all applicable amendments, and informed consent documentation were reviewed and approved by the institutional review board(s) or independent ethics committee(s) at each participating investigational centre. Figure S1 (published online) provides an overview of the study design. Two patient cohorts were screened and enrolled. During the induction phase (weeks 0–6), patients in cohort 1 were randomly assigned to receive double-blind treatment with vedolizumab 300 mg or placebo at weeks 0 and 2. The second cohort was enrolled to fulfil sample size requirements for the maintenance phase; these patients received open-label vedolizumab 300 mg at weeks 0 and 2 during the induction phase. A maintenance phase (weeks 6–52) followed, wherein patients from cohorts 1 and 2 with a clinical response to vedolizumab (defined as a reduction in complete Mayo Clinic score of ≥3 points and a decrease of ≥30% from the baseline score with a decrease of ≥1 point on the rectal bleeding subscale or an absolute rectal bleeding score of 0 or 1) at week 6 of the induction phase were randomly assigned (1:1:1) to receive placebo, vedolizumab 300 mg every 4 weeks, or vedolizumab 300 mg every 8 weeks. These patients comprised the maintenance intention-to-treat population. Patients who did not respond to vedolizumab during the induction phase continued to receive vedolizumab 300 mg every 4 weeks during the maintenance phase, and patients who received placebo during the induction phase continued to receive placebo every 4 weeks during the maintenance phase.

Health-related quality of life assessments

Health-related quality of life was assessed by the IBDQ, 36-Item Short Form Health Survey (SF-36), and EQ-5D, which were administered at screening (baseline) and before dosing at weeks 6, 30, and 52 and at the early termination visit if applicable. The IBDQ is a 32-item, disease-specific instrument that has been validated for use in ulcerative colitis[18] and comprises 4 subscales (bowel symptoms, systemic symptoms, social function, and emotional function).[19] Total scores range from 32 to 224 points, with higher scores indicating better HRQL. A

total IBDQ score ≥170 points is considered to constitute clinical remission in HRQL terms.[2] The SF-36[20] is a generic HRQL measure comprising 36 items that are divided into two summary components (mental component summary and physical component summary) and 8 subscales (physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, and mental health). The EQ-5D is a generic HRQL measure that evaluates five dimensions of health status (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) rated on a 3-point scale (1, *no problems*; 2, *moderate problems*; 3, *extreme problems*).[21–23] The EQ-5D includes a visual analogue scale on which health status is rated from 0 (*worst imaginable*) to 100 (*best imaginable*). EQ-5D utility scores were derived on the basis of UK population-based preference weights and a previously described scoring algorithm.[24, 25]

Patient subgroups

Health-related quality of life for each treatment group was assessed for the maintenance intention-to-treat population overall and for patient subgroups defined by baseline Mayo Clinic score (<9 or ≥9 points) and by prior TNF antagonist failure status.

Statistical analyses

Descriptive statistics were used to summarise patient demographics, baseline characteristics, and concomitant therapy use. All statistical analyses described herein were performed *post hoc*. For scores on the IBDQ (total), IBDQ subscales, SF-36 mental component summary and SF-36 physical component summary, SF-36 subscales, EQ-5D visual analogue scale, and EQ-5D utility index, adjusted mean changes from baseline to week 52 were calculated for each treatment group using an analysis of covariance model. The mean, standard error, and 95% confidence interval (CI) for differences in mean changes from baseline between each vedolizumab group and the placebo group were also calculated. Missing data were imputed using the last observation carried forward approach. The same methodology was used to evaluate changes from baseline for each HRQL instrument score among patient subgroups defined by baseline Mayo Clinic score and by prior TNF antagonist failure status.

The SF-36 mental component summary and physical component summary scores were normalised to the general US population, with a mean score of 50 points and a standard deviation (s.d.) of 10 points.[20] For each treatment group, means and associated 95% CIs for SF-36 mental component summary score, physical component summary score, and each subscale score were calculated for baseline and week 52. Mean scores at week 52 were compared with US population norms. The 95% CI of the US population norm was calculated as $50 \pm 1.96 \times [10/\sqrt{n}$ (*n* for the treatment group)]; if the 95% CI of the mean of the SF-36 parameter for the treatment group did not overlap with the 95% CI for the US population norm, the treatment mean was considered significantly different from the US norm.

For each HRQL instrument, the *minimal clinically important difference threshold* was defined on the basis of published literature as follows: a change of ≥16 points for IBDQ[26]; a change of ≥5 points for both the SF-36 mental component summary and SF-36 physical component summary[27]; a change of ≥10 points for the EQ-5D visual analogue scale[27]; and a change of 0.5 times the population s.d. at baseline for the EQ-5D utility index.[28] The number and proportion of patients in each treatment group who met the minimal clinically important difference threshold at week 52 for each HRQL instrument were calculated, as were the number and proportion of patients in each treatment group with a week-52 IBDQ total score of ≥170 points. The last observation carried forward approach was used for imputation of missing data before determining whether minimal clinically important difference thresholds were met.

Concordance between clinical and IBDQ-defined remission

Concordance between remission as defined by clinical measures and remission as defined by IBDQ scores was calculated. Two definitions of *clinical remission* were used: *complete Mayo Clinic score clinical remission* was defined as a complete Mayo Clinic score ≤2 points and no individual subscore >1 point, and *partial Mayo Clinic score clinical remission* was defined as a partial Mayo Clinic score ≤2 points and no individual subscale score >1 point.[29] *IBDQ remission* was defined by an IBDQ total score ≥170 points. Missing data were considered as non-remitters in the analyses. Two-by-two tables displaying concordance and discordance between these measures were generated, along with Kappa statistics to show the strength of agreement. The Kappa statistic can be interpreted as follows: less than chance agreement, <0; slight agreement, 0.01–0.20; fair agreement, 0.21–0.40; moderate agreement, 0.41–0.60; substantial agreement, 0.61–0.80; almost perfect agreement, 0.81–0.99.[30, 31] Differences between the concordant and discordant groups in demographic and other baseline characteristics were examined using descriptive statistics.

RESULTS

Study patients

In the GEMINI 1 study, 373 patients were randomly assigned to receive maintenance treatment with placebo ($n = 126$) or 300 mg of vedolizumab every 8 weeks ($n = 122$) or every 4 weeks ($n = 125$).[17] No clinically important differences between groups in demographics, baseline patient characteristics, or concomitant therapies were noted (Table 1). For all treatment groups, mean SF-36 mental component summary scores at baseline were >10 points below the average for the general US population (i.e. 50 points),[20] whereas mean baseline physical component summary scores were slightly higher. As required by the protocol, all patients had experienced a clinical response over the course of 6 weeks, having received two 300-mg infusions of vedolizumab during the induction phase.

Health-related quality of life

By week 52, mean differences from the placebo group in changes from baseline scores on the IBDQ (total and all subscales) and the EQ-5D visual analogue scale were statistically significant and generally of similar magnitude for both vedolizumab groups (Table 2). For EQ-5D utility scores, a significant difference from the placebo group was seen only for the vedolizumab every 4 weeks group (Table 2).

Among patients with lower baseline disease activity (Mayo Clinic score <9 points), both vedolizumab groups had significantly greater improvements from baseline to week 52 (relative to the placebo group) on all HRQL measures, with the exception of EQ-5D utility score (Table 3). In patients with higher baseline disease activity (Mayo Clinic score ≥9), the vedolizumab every 8 weeks group, but not the vedolizumab every 4 weeks group, had significantly greater improvements from baseline than the placebo group for all measures of HRQL (Table 3). Among patients without prior TNF antagonist failure, improvements from baseline to week 52 on all HRQL measures were significantly greater in both vedolizumab groups than in the placebo group (Table 4).

The mean SF-36 mental component summary and physical component summary scores for both vedolizumab groups that had responded to induction therapy were similar to US population norms at week 52 (Figure 1). In contrast, mean scores on both domains were significantly lower than US population norms for patients who had received induction therapy and had subsequently been assigned to receive placebo. General health was the only SF-36 subscale on which the mean scores of either vedolizumab group were significantly lower than US population norms; in contrast, mean scores for the placebo group were significantly lower than US population norms for all subscales except physical functioning (Figure S2, published online). Overall, compared with placebo-treated patients, patients treated with vedolizumab had 152–201% greater improvements in IBDQ, EQ-5D visual analogue scale, and EQ-5D utility scores.

Compared with the placebo group, both vedolizumab groups had significantly greater proportions of patients with an IBDQ total score ≥170 points at week 52 (Figure 2). The vedolizumab groups had significantly greater proportions of patients who met the minimal clinically important difference thresholds for IBDQ total score (every 8 weeks and every 4 weeks), SF-36 physical component summary score (every 8 weeks only), and EQ-5D visual analogue scale score (every 8 weeks and every 4 weeks) at week 52 (Figure 2) than the placebo group did. Overall, the percentage of patients who met minimal clinically important difference thresholds for the IBDQ, SF-36 physical component summary scale, SF-36 mental component summary scale, EQ-5D visual analogue scale, and EQ-5D utility score was 6.9–19.9% greater with vedolizumab than with placebo.

Concordance between clinical and IBDQ-defined remission

More than 80% of patients with clinical remission at week 52 also had IBDQ remission (108/127 based on complete Mayo Clinic score, 129/159 based on partial Mayo Clinic score). Among the 150 patients with IBDQ remission at week 52, more than 70% had concomitant clinical remission ($n = 108$ based on complete Mayo Clinic score, $n = 129$ based on partial Mayo Clinic score).

Approximately 84% of patients had concordance between remission defined by both IBDQ and clinical measures (Table 5). Concordance with IBDQ remission did not differ substantially (<3%) from clinical remission as defined by either complete or partial Mayo Clinic scores. Kappa statistics showed substantial agreement between IBDQ remission and clinical remission [0.65 (95% CI, 0.57–0.73) based on complete Mayo Clinic score; 0.72 (95% CI, 0.65–0.79) based on partial Mayo Clinic score]. Although most demographic and baseline characteristics were similar among the concordant and discordant groups, concordance appeared to be more common in males, patients with a disease duration <3 years, and patients with baseline Mayo Clinic scores of 9–12 points (data not shown).

Table 1 | Demographics, baseline characteristics, and prior and concomitant therapies

Characteristic	Placebo (n = 126)*	Vedolizumab every 8 weeks (n = 122)†	Vedolizumab every 4 weeks (n = 125)‡
Demographics			
Age, years, mean (s.d.)§	40.3 (14)	41.0 (13)	38.6 (14)
Male sex, No. (%)§	69 (55)	70 (57)	68 (54)
Current smoker, No. (%)§	8 (6)	7 (6)	8 (6)
Geographic region, No. (%)			
North America	36 (29)	49 (40)	37 (30)
Western/Northern Europe	20 (16)	23 (19)	25 (20)
Central Europe	26 (21)	20 (16)	25 (20)
Eastern Europe	10 (8)	12 (10)	11 (9)
Asia/Australia/Africa	34 (27)	18 (15)	27 (22)
Baseline medication use			
Corticosteroid use, No. (%)	72 (57)	70 (57)	73 (58)
Immunomodulator use, No. (%)	51 (40)	43 (35)	45 (36)
Disease characteristics			
Duration of disease, years, mean (s.d.)§	7.8 (7)	6.2 (5)	7.6 (7)
Mayo Clinic score, mean (s.d.)§	8.4 (1.8)	8.4 (1.8)	8.3 (1.7)
Baseline Mayo Clinic score ≥9, No. (%)	57 (45)	55 (45)	52 (42)
Mayo Clinic endoscopic subscale score, No. (%)			
0: Normal or inactive disease	0	0	0
1: Mild disease	0	0	1 (<1)¶
2: Moderate disease	59 (47)	60 (49)	61 (49)
3: Severe disease	67 (53)	62 (51)	63 (50)
Prior TNF antagonist failure, No. (%)§	38 (30)	43 (35)	40 (32)
HRQL measures			
IBDQ score, mean (s.d.)§			
Baseline	122 (34)	125 (34)	124 (34)
Week 6**	171 (33)	171 (35)	169 (33)
SF-36 physical component summary score, mean (s.d.)			
Baseline	39.7 (8.4)	40.0 (8.5)	41.1 (7.7)
Week 6**	46.7 (7.8)	47.0 (8.6)	46.8 (7.8)
SF-36 mental component summary score, mean (s.d.)			
Baseline	38.5 (12.0)	39.2 (11.6)	37.9 (11.2)
Week 6**	45.8 (11.6)	46.0 (12.5)	45.3 (10.5)
EQ-5D visual analogue scale score, mean (s.d.)			
Baseline	54.6 (20.2)	56.6 (20.9)	53.6 (20.3)
Week 6**	70.5 (18.4)	71.6 (19.9)	70.7 (17.2)
EQ-5D utility score, mean (s.d.)			
Baseline	0.677 (0.213)	0.673 (0.235)	0.674 (0.227)
Week 6**	0.798 (0.236)	0.799 (0.214)	0.786 (0.200)

HRQL, health-related quality of life; IBDQ, Inflammatory Bowel Disease Questionnaire; s.d., standard deviation; SF-36, 36-Item Short Form Health Survey; TNF, tumour necrosis factor.

Patients who received placebo during the randomised maintenance phase had received induction doses of vedolizumab at weeks 0 and 2.

* n = 125 for the baseline EQ-5D visual analogue scale score.

† n = 121 for the baseline IBDQ, SF-36 physical component summary, SF-36 mental component summary, EQ-5D visual analogue scale, and EQ-5D utility scores.

‡ n = 124 for baseline IBDQ and n = 123 for the baseline SF-36 physical component summary, SF-36 mental component summary, EQ-5D visual analogue scale, and EQ-5D utility scores.

§ From Feagan *et al.*[17]; also refer to Feagan *et al.* for disease localisation, faecal calprotectin level, and steroid dose.

¶ This patient from induction cohort 2 had a Mayo Clinic endoscopic subscale score of 1 point and a Mayo Clinic score of 8 points at baseline; therefore, the patient violated the inclusion criteria but was included in the maintenance intention-to-treat population.

** Week 6 of the study was the first week of the maintenance phase.

Table 2 | Differences vs. placebo for changes from baseline in disease-specific and generic measures of HRQL at week 52

Instrument	Placebo (n = 126)	Vedolizumab every 8 weeks (n = 122)	Vedolizumab every 4 weeks (n = 125)
IBDQ	n = 126	n = 121	n = 124
Total			
Baseline score, mean (S.E.)	122.2 (3.0)	124.5 (3.1)	123.7 (3.1)
Adjusted* change from baseline, mean (S.E.)	27.3 (3.3)	48.4 (3.4)	49.0 (3.3)
Difference from placebo, mean (95% CI)	NA	21.1 (**11.8–30.4**)	21.6 (**12.4–30.9**)
Bowel systems			
Baseline score, mean (S.E.)	37.3 (0.9)	38.3 (0.9)	38.7 (1.0)
Adjusted* change from baseline, mean (S.E.)	8.5 (1.1)	17.1 (1.1)	16.3 (1.1)
Difference from placebo, mean (95% CI)	NA	8.6 (**5.5–11.8**)	7.8 (**4.7–10.9**)
Emotional function			
Baseline score, mean (S.E.)	47.3 (1.4)	47.5 (1.3)	47.0 (1.2)
Adjusted* change from baseline, mean (S.E.)	9.1 (1.2)	16.2 (1.3)	17.0 (1.3)
Difference from placebo, mean (95% CI)	NA	7.1 (**3.6–10.6**)	7.9 (**4.5–11.4**)
Social function			
Baseline score, mean (S.E.)	19.9 (0.7)	20.7 (0.7)	20.1 (0.7)
Adjusted* change from baseline, mean (S.E.)	5.2 (0.6)	7.9 (0.7)	8.6 (0.6)
Difference from placebo, mean (95% CI)	NA	2.6 (**0.9–4.4**)	3.4 (**1.6–5.2**)
Systemic function			
Baseline score, mean (S.E.)	17.7 (0.5)	18.0 (0.6)	17.9 (0.6)
Adjusted* change from baseline, mean (S.E.)	4.4 (0.5)	7.3 (0.6)	7.1 (0.5)
Difference from placebo, mean (95% CI)	NA	2.9 (**1.4–4.4**)	2.7 (**1.2–4.2**)
EQ-5D			
Visual analogue scale	n = 125	n = 121	n = 123
Baseline score, mean (S.E.)	54.6 (1.8)	56.6 (1.9)	53.6 (1.8)
Adjusted* change from baseline, mean (S.E.)	9.7 (1.7)	19.0 (1.7)	19.4 (1.7)
Difference from placebo, mean (95% CI)	NA	9.3 (**4.6–14.0**)	9.7 (**5.0–14.4**)
Utility	n = 126	n = 121	n = 123
Baseline score, mean (S.E.)	0.677 (0.019)	0.673 (0.021)	0.674 (0.020)
Adjusted* change from baseline, mean (S.E.)	0.083 (0.019)	0.131 (0.019)	0.141 (0.019)
Difference from placebo, mean (95% CI)	NA	0.049 (−0.003–0.101)	0.058 (**0.006–0.110**)

CI, confidence interval; HRQL, health-related quality of life; IBDQ, Inflammatory Bowel Disease Questionnaire; NA, not applicable; S.E., standard error.

Missing data were imputed using the last observation carried forward approach. Bolded 95% CIs indicate significant differences from the placebo group.

* Adjusted mean changes from baseline to week 52 were calculated for each treatment group using an analysis of covariance model.

DISCUSSION

Health-related quality of life evaluations are now routinely accepted as outcome measures in clinical trials of new medical therapies for ulcerative colitis. Accordingly, the GEMINI 1 trial used several generic and disease-specific HRQL instruments at multiple time points throughout the study. At week 6 (i.e. the first week of the maintenance phase) of GEMINI 1, all participants in the maintenance intention-to-treat population had demonstrated a clinical response to vedolizumab induction therapy (Table 1). Following maintenance randomisation, clinically meaningful and statistically significant improvements in both vedolizumab groups compared with placebo were observed for most HRQL measures. For example, following induction, the mean week-6 IBDQ scores for the vedolizumab every 8 weeks and every 4 weeks groups and the placebo group were 171, 169, and 171 points, respectively. At week 52, the corresponding values were 172, 173, and 150 points. Given that a total IBDQ score ≥170 denotes remission, these data indicate that most patients treated with vedolizumab over the longer term can be expected to have improved HRQL. Similarly, although participants' week 6 SF-36 mental component summary scores were less

Table 3 | Differences from placebo for changes in measures of HRQL at week 52 by baseline disease activity

Instrument Difference from placebo in adjusted* change from baseline score	Baseline Mayo Clinic score <9†		Baseline Mayo Clinic score ≥9‡	
	Vedolizumab every 8 weeks (n = 67)	Vedolizumab every 4 weeks (n = 73)	Vedolizumab every 8 weeks (n = 55)	Vedolizumab every 4 weeks (n = 52)
IBDQ total	n = 66	n = 73	n = 55	n = 51
Mean (95% CI)	20.8 (**8.4–33.1**)	27.6 (**15.6–39.6**)	21.7 (**7.6–35.9**)	13.8 (−0.7–28.2)
SF-36				
Physical component summary	n = 66	n = 73	n = 55	n = 50
Mean (95% CI)	2.6 (**0.2–5.0**)	3.8 (**1.5–6.2**)	4.1 (**1.3–6.9**)	1.4 (−1.5–4.3)
Mental component summary	n = 66	n = 73	n = 55	n = 50
Mean (95% CI)	5.1 (**1.6–8.6**)	6.1 (**2.7–9.5**)	4.3 (**0.8–7.8**)	2.8 (−0.9–6.4)
EQ-5D				
Visual analogue scale	n = 66	n = 73	n = 55	n = 50
Mean (95% CI)	8.2 (**1.7–14.6**)	11.5 (**5.3–17.8**)	10.5 (**3.4–17.5**)	7.2 (0–14.5)
Utility index	n = 66	n = 73	n = 55	n = 50
Mean (95% CI)	0.007 (−0.052–0.066)	0.028 (−0.029–0.086)	0.098 (**0.006–0.190**)	0.092 (−0.001–0.186)

CI, confidence interval; HRQL, health-related quality of life; IBDQ, Inflammatory Bowel Disease Questionnaire; SF-36, 36-Item Short Form Health Survey.

Missing data were imputed using the last observation carried forward approach. Bolded 95% CIs indicate significant differences from the placebo group.

* Adjusted mean changes from baseline to week 52 were calculated for each treatment group using an analysis of covariance model.

† n = 69 for the placebo group except for the EQ-5D visual analogue scale score (n = 68).

‡ n = 57 for the placebo group.

than US population normative values, patients who continued vedolizumab treatment achieved, on average, normal scores at the end of 52 weeks of therapy, whereas those who received placebo showed minimal improvement.

These treatment-related benefits parallel those observed in clinical outcomes such as remission, response, and mucosal improvement. At one level, this observation is not surprising because clinical disease activity is a major determinant of HRQL in patients with ulcerative colitis[1, 2, 4–6, 32, 33]; however, ulcerative colitis is a complex disease that may adversely affect patients' social life, employment status, and psychological functioning,[34–36] and these domains may not improve after a relatively short course of therapy during a lifetime disease experience. More than 80% concordance was demonstrated between patients in clinical remission and those in IBDQ-defined remission, and Kappa statistics supported substantial agreement between the two (Table 5). Both discordant patient groups are interesting. Persistence of impaired HRQL in the presence of clinical remission may relate to other covariates such as coexisting depression or socioeconomic status. Unfortunately,

this study had limited capacity to assess whether these factors were responsible for discordance. A comparison of demographic and baseline characteristics collected for the trial suggested that concordance was more common among males, patients with shorter disease duration, and those with higher baseline Mayo Clinic scores. These *post hoc* analyses were limited to descriptive statistics but could be useful in hypothesis generation for further research.

Patients who fail to meet the criteria for clinical remission while in IBDQ-defined remission are also of interest. A substantial proportion of this subset of discordance (54.8%) was due to failure to achieve a Mayo Clinic endoscopic score of 0 or 1. This observation further emphasises the need for objective assessment of inflammatory disease activity in guiding therapy for ulcerative colitis given the evolving belief that endoscopically defined treatment targets are preferable to those based on patient-reported outcomes[35]; however, 35.7% of patients who met endoscopic remission criteria failed to meet patient-reported outcome criteria, including bleeding (4.8%) and stool frequency (4.8%) criteria. Persistence of increased stool frequency in the presence of

Table 4 | Differences from placebo for changes in measures of HRQL at week 52 by prior TNF antagonist failure status

Instrument Difference from placebo in adjusted* change from baseline score	Prior TNF antagonist failure†		No prior TNF antagonist failure‡	
	Vedolizumab every 8 weeks (n = 43)	Vedolizumab every 4 weeks (n = 40)	Vedolizumab every 8 weeks (n = 79)	Vedolizumab every 4 weeks (n = 85)
IBDQ total	n = 43	n = 40	n = 78	n = 84
Mean (95% CI)	14.1 (−2.5–30.7)	13.4 (−3.4–30.2)	25.9 (**14.6–37.3**)	25.8 (**14.7–36.9**)
SF-36				
Physical component summary	n = 43	n = 40	n = 78	n = 83
Mean (95% CI)	2.2 (−1.0–5.4)	1.1 (−2.1–4.4)	3.9 (**1.7–6.2**)	3.6 (**1.4–5.8**)
Mental component summary	n = 43	n = 40	n = 78	n = 83
Mean (95% CI)	3.3 (−1.2–7.8)	2.2 (−2.3–6.7)	6.0 (**2.9–9.0**)	6.2 (**3.2–9.1**)
EQ-5D				
Visual analogue scale	n = 43	n = 40	n = 78	n = 83
Mean (95% CI)	6.8 (−1.8–15.5)	6.9 (−2.0–15.7)	10.6 (**4.9–16.3**)	11.1 (**5.5–16.7**)
Utility index	n = 43	n = 40	n = 78	n = 83
Mean (95% CI)	0.033 (−0.073–0.139)	0.046 (−0.061–0.153)	0.062 (**0.003–0.120**)	0.066 (**0.008–0.123**)

CI, confidence interval; HRQL, health-related quality of life; IBDQ, Inflammatory Bowel Disease Questionnaire; SF-36, 36-Item Short Form Health Survey; TNF, tumour necrosis factor.

Missing data were imputed using the last observation carried forward approach. Bolded 95% CIs indicate significant differences from the placebo group.

* Adjusted mean changes from baseline to week 52 were calculated for each treatment group using an analysis of covariance model.

† n = 38 for the placebo group except for the EQ-5D visual analogue scale score (n = 37).

‡ n = 88 for the placebo group.
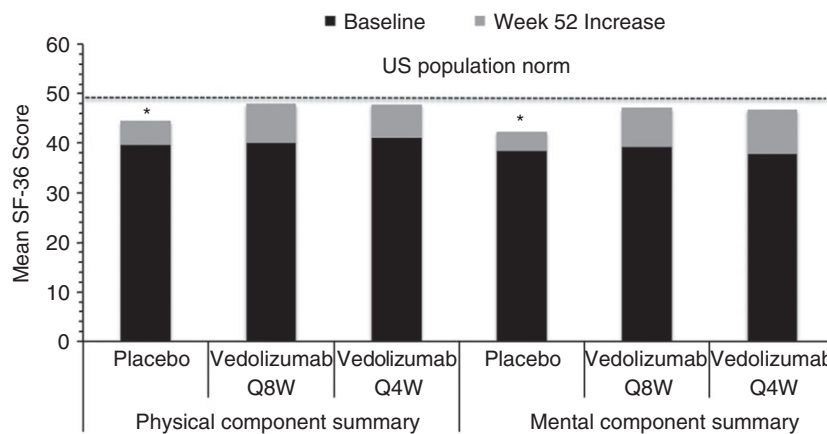
Figure 1 | Mean scores at baseline (black portion of the bars) and increases in mean scores at week 52 (grey portion of the bars) for the SF-36 physical component summary and SF-36 mental component summary. Missing data were imputed using the last observation carried forward approach. *Indicates week 52 value that is considered significantly different from the US population norm (i.e. the 95% CI of the mean of the SF-36 parameter for the treatment group does not overlap with the 95% CI of the US population norm). For all parameters, at baseline: placebo n = 126, vedolizumab every 8 weeks n = 121, and vedolizumab every 4 weeks n = 123; at week 52: placebo n = 126, vedolizumab every 8 weeks n = 122, and vedolizumab every 4 weeks n = 125. Patients who received placebo during the randomised maintenance phase had received induction doses of vedolizumab at weeks 0 and 2. CI, confidence interval; Q4W, every 4 weeks; Q8W, every 8 weeks; SF-36, 36-Item Short-Form Health Survey.
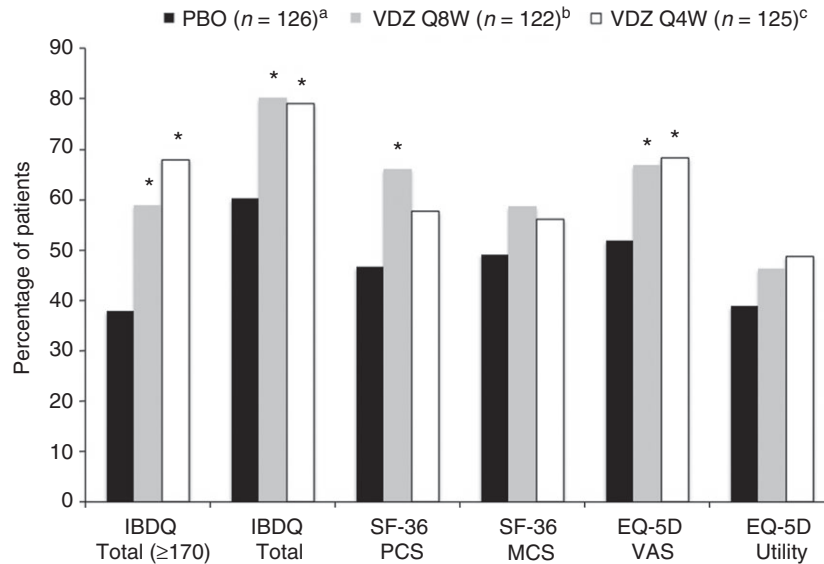
Figure 2 | Proportion of patients with IBDQ total score ≥170 points or with clinically meaningful improvements (i.e. minimal clinically important difference thresholds met) in HRQL measurements at week 52. Patients who received placebo during the randomised maintenance phase had received induction doses of vedolizumab at weeks 0 and 2. Missing data were imputed using the last observation carried forward approach. *Indicates that a value is considered significantly different from that of placebo (i.e. 95% CIs for the mean of the difference from placebo do not contain 0). [a]n = 125 for EQ-5D VAS score. [b]n = 121 for IBDQ Total minimal clinically important difference, SF-36 mental component summary, SF-36 physical component summary, EQ-5D VAS, and EQ-5D utility scores. [c]n = 124 for IBDQ Total minimal clinically important difference; n = 123 for SF-36 mental component summary, SF-36 physical component summary, EQ-5D VAS, and EQ-5D utility scores. CI, confidence interval; HRQL, health-related quality of life; IBDQ, Inflammatory Bowel Disease Questionnaire; MCS, mental component summary; PCS, physical component summary; Q4W, every 4 weeks; Q8W, every 8 weeks; SF-36, 36-Item Short Form Health Survey; VAS, visual analogue scale.

mucosal improvement is a recognised phenomenon that may be due to residual histological inflammation or submucosal fibrosis that results in impaired colonic motility.[37] Further research into possible mechanisms underlying persistently increased stool frequency should be conducted in this population. The persistence of bleeding in patients with endoscopically defined remission is poorly understood and may be due to haemorrhoids or persistent disease beyond the range of the sigmoid colon that was examined.

Results from the subgroup analyses demonstrated that vedolizumab-treated patients with lower disease activity at baseline and those with no prior TNF antagonist failure generally had greater improvements (vs. placebo-treated patients) in HRQL. The latter is unsurprising when one considers that patients with prior TNF antagonist failure typically demonstrate an overall poorer response to treatment than do patients with no prior TNF antagonist failure,[17, 38, 39] possibly because of prevalent non-inflammatory conditions such as stricture and malabsorption, which result from persistent

inflammation over a long duration of disease. Currently, little else is known about whether ulcerative colitis therapy differentially affects HRQL among various patient subgroups, and more research is needed. In this regard, we could not confirm results from previous studies that identified female sex as a potential covariate.[13, 14] Also, we did not collect data to evaluate potential effects of coexistent depression and fatigue or the number of previous relapses.

Our study had some limitations. First, SF-36 data from GEMINI 1, which was an international study, were compared with norms from the US population. Population-based norms are useful for assessing whether scores differ from expected values in a reference population, and a similar approach has been employed in previous studies.[33, 40, 41] Determining whether SF-36 scores for patients from each geographic region differ from population norms in that region would require additional analysis and comparison against normative data for each region of interest. Second, in the current analyses, a change of ≥10 points was used as the minimal clinically

Table 5 | Concordance between clinical remission and IBDQ remission at week 52 (intention-to-treat population, *n* = 373)

	IBDQ remission*	
	Yes	No
(A) Clinical remission (complete Mayo Clinic score)†		
Yes	108 (29.0%)	19 (5.1%)
No	42 (11.3%)	204 (54.7%)
Kappa statistic (95% CI)‡	0.65 (0.57–0.73)	
(B) Clinical remission (partial Mayo Clinic score)§		
Yes	129 (34.6%)	30 (8.0%)
No	21 (5.6%)	193 (51.7%)
Kappa statistic (95% CI)‡	0.72 (0.65–0.79)	

CI, confidence interval; IBDQ, Inflammatory Bowel Disease Questionnaire.

Data are shown as No. (%) of patients.

* *IBDQ remission* was defined by an IBDQ total score ≥170 points.

† *Clinical remission* was defined by a complete Mayo Clinic score ≤2 points and no individual subscore >1 point (A).

‡ Kappa statistics: <0, less than chance agreement; 0.01–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; 0.81–0.99, almost perfect agreement.

§ *Clinical remission* was defined by a partial Mayo Clinic score ≤2 points and no individual subscore >1 point (B).

important difference threshold for the EQ-5D visual analogue scale. Although Coteur et al.[27] estimated that the EQ-5D visual analogue scale minimal clinically important difference ranges between 4.2 and 14.8, a specific value has not been firmly established. Finally, despite the relatively large size of the trial, limited statistical power was available to identify factors predictive of HRQL improvements following vedolizumab therapy in *post hoc* analyses. In future research, data from larger samples will be needed to evaluate potential predictors.

CONCLUSION

These analyses indicate that patients with a clinical response to vedolizumab induction therapy have improved HRQL with continued treatment. Further research is needed to identify determinants of HRQL in patients with clinically quiescent disease.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. GEMINI 1 study design. VDZ doses were 300 mg. [a]Induction doses were administered at weeks 0

and 2. [b]Patients who received PBO during the randomised maintenance phase had received induction doses of VDZ at weeks 0 and 2. PBO, placebo; VDZ, vedolizumab.

Figure S2. Mean scores at baseline (black portion of the bars) and increases in mean scores at week 52 (grey portion of the bars) for SF-36 subscales. Missing data were imputed using the last observation carried forward approach. * Indicates week 52 value that is considered significantly different from the US population norm (ie, the 95% CI of the mean of the SF-36 parameter for the treatment group does not overlap with the 95% CI of the US population norm). For all parameters, at baseline: placebo n=126, vedolizumab every 8 weeks n=121, and vedolizumab every 4 weeks n=123; at week 52: placebo n=126, vedolizumab every 8 weeks n=122, and vedolizumab every 4 weeks n=125. Patients who received placebo during the randomised maintenance phase had received induction doses of vedolizumab at weeks 0 and 2. CI, confidence interval; Q4W, every 4 weeks; Q8W, every 8 weeks; SF-36, 36-Item Short Form Health Survey.

AUTHORSHIP

ACKNOWLEDGMENTS

REFERENCES

1. 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.

2. Irvine EJ. Quality of life of patients with ulcerative colitis: past, present, and future. *Inflamm Bowel Dis* 2008; **14**: 554–65.

3. Ghosh S, Mitchell R. Impact of inflammatory bowel disease on quality of life: results of the European Federation of Crohn's and Ulcerative Colitis Associations (EFCCA) patient survey. *J Crohns Colitis* 2007; **1**: 10–20.

4. Kalafateli M, Triantos C, Theocharis G, et al. Health-related quality of life in patients with inflammatory bowel disease: a single-center experience. *Ann Gastroenterol* 2013; **26**: 243–8.

5. Kuriyama M, Kato J, Kuwaki K, et al. Clinical factors that impair health-related quality of life in ulcerative colitis patients vary with the disease duration. *Eur J Gastroenterol Hepatol* 2008; **20**: 634–41.

6. Lix LM, Graff LA, Walker JR, et al. Longitudinal study of quality of life and psychological functioning for active, fluctuating, and inactive disease patterns in inflammatory bowel disease. *Inflamm Bowel Dis* 2008; **14**: 1575–84.

7. Casellas F, Barreiro de Acosta M, Iglesias M, et al. Mucosal healing restores normal health and quality of life in patients with inflammatory bowel disease. *Eur J Gastroenterol Hepatol* 2012; **24**: 762–9.

8. Rochelle TL, Fidler H. The importance of illness perceptions, quality of life and psychological status in patients with ulcerative colitis and Crohn's disease. *J Health Psychol* 2013; **18**: 972–83.

9. Tabibian A, Tabibian JH, Beckman LJ, Raffals LL, Papadakis KA, Kane SV. Predictors of health-related quality of life and adherence in Crohn's disease and ulcerative colitis: implications for clinical management. *Dig Dis Sci* 2015; **60**: 1366–74.

10. van der Have M, van der Aalst KS, Kaptein AA, et al. Determinants of health-related quality of life in Crohn's disease: a systematic review and meta-analysis. *J Crohns Colitis* 2014; **8**: 93–106.

11. Iglesias-Rey M, Barreiro-de Acosta M, Caamaño-Isorna F, et al. Psychological factors are associated with changes in the health-related quality of life in inflammatory bowel disease. *Inflamm Bowel Dis* 2014; **20**: 92–102.

12. Kim ES, Cho KB, Park KS, et al. Predictive factors of impaired quality of life in Korean patients with inactive inflammatory bowel disease: association with functional gastrointestinal disorders and mood disorders. *J Clin Gastroenterol* 2013; **47**: e38–44.

13. Moradkhani A, Beckman LJ, Tabibian JH. Health-related quality of life in inflammatory bowel disease: psychosocial, clinical, socioeconomic, and demographic predictors. *J Crohns Colitis* 2013; **7**: 467–73.

14. Magalhães J, Castro FD, Carvalho PB, Moreira MJ, Cotter J. Quality of life in patients with inflammatory bowel disease: importance of clinical, demographic and psychosocial factors. *Arq Gastroenterol* 2014; **51**: 192–7.

15. Zhang CK, Hewett J, Hemming J, et al. The influence of depression on quality of life in patients with inflammatory bowel disease. *Inflamm Bowel Dis* 2013; **19**: 1732–9.

16. Cohen BL, Zoëga H, Shah SA, et al. Fatigue is highly associated with poor health-related quality of life, disability and depression in newly-diagnosed patients with inflammatory bowel disease, independent of disease activity. *Aliment Pharmacol Ther* 2014; **39**: 811–22.

17. Feagan BG, Rutgeerts P, Sands BE, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. *N Engl J Med* 2013; **369**: 699–710.

18. Han SW, McColl E, Steen N, Barton JR, Welfare MR. The inflammatory bowel disease questionnaire: a valid and reliable measure in ulcerative colitis patients in the North East of England. *Scand J Gastroenterol* 1998; **33**: 961–6.

19. Guyatt G, Mitchell A, Irvine EJ, et al. A new measure of health status for clinical trials in inflammatory bowel disease. *Gastroenterology* 1989; **96**: 804–10.

20. Ware JE Jr. SF-36 health survey update. Available at: http://www.SF-36.org/tools/sf36.shtml (accessed 21 October 2014).

21. König H-H, Ulshöfer A, Gregor M, et al. Validation of the EuroQol questionnaire in patients with inflammatory bowel disease. *Eur J Gastroenterol Hepatol* 2002; **14**: 1205–15.

22. Brooks R. EuroQol: the current state of play. *Health Policy* 1996; **37**: 53–72.

23. EuroQol Group. EuroQol–a new facility for the measurement of health-related quality of life. *Health Policy* 1990; **16**: 199–208.

24. Agency for Healthcare Research and Quality. Calculating the U.S. population-based EQ-5D index score. Available at: http://www.ahrq.gov/professionals/clinicians-providers/resources/rice/EQ5Dscore.html-weights. Updated August 2005 (accessed 5 January 2015).

25. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care* 2005; **43**: 203–20.

26. Irvine EJ. Development and subsequent refinement of the inflammatory bowel disease questionnaire: a quality-of-life instrument for adult patients with inflammatory bowel disease. *J Pediatr Gastroenterol Nutr* 1999; **28**: S23–7.

27. Coteur G, Feagan B, Keininger DL, Kosinski M. Evaluation of the meaningfulness of health-related quality of life improvements as assessed by the SF-36 and the EQ-5D VAS in patients with active Crohn's disease. *Aliment Pharmacol Ther* 2009; **29**: 1032–41.

28. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003; **41**: 582–92.

29. Travis SP, Higgins PD, Orchard T, et al. Review article: defining remission in ulcerative colitis. *Aliment Pharmacol Ther* 2011; **34**: 113–24.

30. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960; **20**: 37–46.

31. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–74.

32. Hodgkins P, Yen L, Yarlas A, Karlstadt R, Solomon D, Kane S. Impact of MMX® mesalamine on improvement and maintenance of health-related quality of life in patients with ulcerative colitis. *Inflamm Bowel Dis* 2013; **19**: 386–96.

33. Feagan BG, Reinisch W, Rutgeerts P, et al. The effects of infliximab therapy on health-related quality of life in ulcerative colitis patients. *Am J Gastroenterol* 2007; **102**: 794–802.

34. Jedel S, Hood MM, Keshavarzian A. Getting personal: a review of sexual functioning, body image, and their impact on quality of life in patients with inflammatory bowel disease. *Inflamm Bowel Dis* 2015; **21**: 923–38.

35. Peyrin-Biroulet L, Sandborn W, Sands BE, et al. Selecting therapeutic targets in inflammatory bowel disease (STRIDE): determining therapeutic goals for treat-to-target. *Am J Gastroenterol* 2015; **110**: 1324–38.

36. Reiss M, Sandborn WJ. The role of psychosocial care in adapting to health care reform. *Clin Gastroenterol Hepatol* 2015; **13**: 2219–24.

37. 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**: 1082–92.

38. Sandborn WJ, Feagan BG, Rutgeerts P, et al. Vedolizumab as induction and maintenance therapy for Crohn's disease. *N Engl J Med* 2013; **369**: 711–21.

39. Feagan BG, Gasink C, Lang Y, et al. Ustekinumab induction therapy for Crohn's disease refractory to conventional tumor necrosis factor antagonist therapy. *N Engl J Med.* in press.

40. Feagan BG, Sandborn WJ, Hass S, Niecko T, White J. Health-related quality of life during natalizumab maintenance therapy for Crohn's disease. *Am J Gastroenterol* 2007; **102**: 2737–46.

41. Feagan BG, Coteur G, Tan S, Keininger DL, Schreiber S. Clinically meaningful improvement in health-related quality of life in a randomized controlled trial of certolizumab pegol maintenance therapy for Crohn's disease. *Am J Gastroenterol* 2009; **104**: 1976–83.