IMPACT OF ADALIMUMAB ON DISEASE BURDEN IN MODERATE-TO-SEVERE ULCERATIVE COLITIS PATIENTS: THE ONE-YEAR, REAL-WORLD UCANADA STUDY

Supplementary material

Excerpt from the Statistical Analysis Plan:

"**6.5.5 Analysis for secondary objective:** ***To determine the correlation between effectiveness (clinical response and remission) rates and PRO measures***

The association between effectiveness (clinical response and remission; both binary variables) and each of the PRO measures was assessed through modelling approaches for repeated measures including observations from all follow-up visits. To ensure an unbiased coefficient of response/remission status variable, the models were adjusted for potential confounders and the unbalanced characteristics between responders and non-responders (or patients achieving remission and patients not achieving remission).

The association between effectiveness (clinical response and remission; both binary variables) and each of the PRO measures except for VOLP outcomes was assessed using a mixed model for repeated measures including observations from all follow up visits. Regression analysis for cross-sectional data was applied to VOLP outcomes and effectiveness at week 52. Model building followed the methods described next in principle.

All models with repeated measures included a random intercept with the effectiveness variable (fixed, forced-in), visit (fixed, forced-in), baseline value of the PRO measure (fixed, forced-in) and other covariates, which were selected based on the model selection mechanism described below. Cross-sectional regression models included an intercept with the effectiveness variable (forced-in), baseline value of the PRO measure (fixed) and other covariates, which were selected based on the same proposed method described below.

Least squares means, *p*-value and 2-sided 95% confidence interval of the difference between the two groups defined by the clinical effectiveness were determined. Significance

tests were based on least-squares means using a two-sided $\alpha = .05$.

For each model appropriate distribution was specified depending on the distribution of the PRO measures. For example, logistic regression was used for binary outcomes. Models assuming normal distribution were used for continuous outcomes, which did not severely deviate from normal distribution. Since paid or unpaid productivity loss in days/hours are highly skewed count data with excessive zeros, Poisson, Negative Biominal, zero-inflated Poisson or zero- inflated Negative Binomial models were examined and compared. Likelihood Ratio test was used to select between Poisson and Negative Binomial or between zero-inflated Poisson and zero-inflated Negative Binomial models, while Vuong's test was used to select between Poisson and zero-inflated Poisson or between Negative Binomial and zero-inflated Negative Binomial models. These model selections were based on only the forced-in independent variables. For the total costs of lost productivity, two-part model (logistic regression for whether cost > 0 and generalized linear regression model with gamma distribution for cost > 0) was performed.

To facilitate model building, we divided all baseline and patient characteristic variables into five groups: Group 1: social-demographics (age, gender, and ethnicity, etc.); Group 2: medical history (comorbidities) and medication use (e.g., prescriptions of systemic corticosteroids and steroid use, historical use of AZA/6-MP/MTX including since UC diagnosis and in the prior 6 months and concomitant use of AZA/6-MP/MTX with adalimumab at start of study or addition during study period, etc.); Group 3: baseline clinical outcomes (Montreal classification of extent of ulcerative colitis [E1, E2, E3] , disease duration, family history of UC, healthcare use [endoscopy, UC-related emergency room visit and hospitalization, patient/physician global assessment of disease activity, etc.]; Group 4: baseline quality of life (QoL) measures (EQ-5D); Group 5: job/workplace characteristics (working status, job type, work habit, etc.). Group 5 was only relevant for the productivity related outcomes.

Baseline variables were first selected within each group of the independent variables (group variable selection) The selection criteria for group variable selections included entry

criterion p value ≤0.2 and the Bayesian Information Criterion (BIC). The smaller the BIC was, the better the model fitted. The final model selection was then be constructed among the variables selected in each group in the first step along with the variables being forced in. The selection criteria for the final model selections were p value ≤0.1 and the BIC. In case of a zero-inflated model or two-part model, the same covariates was included in both parts of the models, i.e., if one covariate met the selection criteria and was selected into one part of the model, it was included in both parts. Using the variable selection method, we avoided over-adjustment and address the issue that the variables within the same group are usually highly correlated (multicollinearity issue).

Note that if the study were stopped early, a simplified model following the same method described in principle was to be used due to smaller sample size."

Supplementary Table 1 Sensitivity analysis on the proportion of patients with improvement in Patient Health Questionnaire – 9 Items total score at week 52/final visit

Analysis population	Imputation method	Total	*N*	Proportion (%)	95% CI
Intent-to-treat	NRI	94	40	42.6%	(32.6%, 52.6%)
	LOCF	94	59	62.8%	(53.0%, 72.5%)
Per protocol	NRI	48	29	60.4%	(46.6%, 74.3%)
	LOCF	48	32	66.7%	(53.3%, 80.0%)

CI: Confidence interval, NRI: Non-responder imputation; LOCF: Last observation carried forward; PHQ-9: Patient Health Questionnaire – 9 Items.

Supplementary Table 2 Patient Health Questionnaire – 9 Items: Proportion of response by item by visit

PHQ-9 item	Response	Baseline *n* (%)	Week 8 *n* (%)	Week 52 *n* (%)
1. Little interest or pleasure in doing things	Not at all	33 (35.5%)	38 (46.3%)	33 (50.8%)
	Several days	31 (33.3%)	29 (35.4%)	23 (35.4%)
	More than half the days	12 (12.9%)	9 (11.0%)	5 (7.7%)
	Nearly every day	17 (18.3%)	6 (7.3%	4 (6.2%)
2. Feeling down, depressed, or hopeless	Not at all	37 (39.4%)	44 (53.7%)	29 (45.3%)
	Several days	37 (39.4%)	24 (29.3%)	22 (34.4%)
	More than half the days	13 (13.8%)	7 (8.5%)	8 (12.5%)
	Nearly every day	7 (7.4%)	7 (8.5%)	5 (7.8%)
3. Trouble falling or staying asleep, or sleeping too much	Not at all	22 (23.4%)	21 (25.9%)	17 (26.2%)
	Several days	26 (27.7%)	34 (42.0%)	29 (44.6%)
	More than half the days	17 (18.1%)	13 (16.0%)	7 (10.8%)
	Nearly every day	29 (30.9%)	13 (16.0%)	12 (18.5%)
4. Feeling tired or having little energy	Not at all	14 (14.9%)	9 (11.1%)	15 (23.1%)
	Several days	30 (31.9%)	41 (50.6%)	25 (38.5%)
	More than half the days	15 (16.0%)	11 (13.6%)	14 (21.5%)
	Nearly every day	35 (37.2%)	20 (24.7%)	11 (16.9%)

5. Poor appetite or overeating	Not at all	27 (29.0%)	34 (42.0%)	34 (52.3%)
	Several days	26 (28.0%)	30 (37.0%)	18 (27.7%)
	More than half the days	14 (15.1%)	11 (13.6%)	8 (12.3%)
	Nearly every day	26 (28.0%)	6 (7.4%)	5 (7.7%)
6. Feeling bad about yourself – or that you are a failure or have let yourself or your family down	Not at all	53 (57.0%)	53 (64.6%)	36 (55.4%)
	Several days	19 (20.4%)	18 (22.0%)	20 (30.8%)
	More than half the days	11 (11.8%)	7 (8.5%)	6 (9.2%)
	Nearly every day	10 (10.8%)	4 (4.9%)	3 (4.6%)
7. Trouble concentrating on things, such as reading the newspaper or watching television	Not at all	51 (54.8%)	41 (50.0%)	33 (50.8%)
	Several days	21 (22.6%)	32 (39.0%)	23 (35.4%)
	More than half the days	15 (16.1%)	3 (3.7%)	8 (12.3%)
	Nearly every day	6 (6.5%)	6 (7.3%)	1 (1.5%)
8. Moving or speaking so slowly that other people could have noticed?[1]	Not at all	68 (73.1%)	59 (72.0%)	50 (76.9%)
	Several days	17 (18.3%)	17 (20.7%)	10 (15.4%)
	More than half the days	4 (4.3%)	4 (4.9%)	4 (6.2%)
	Nearly every day	4 (4.3%)	2 (2.4%)	1 (1.5%)
9. Thoughts that you would be better off dead or of hurting yourself in some way	Not at all	84 (90.3%)	76 (92.7%)	59 (90.8%)
	Several days	6 (6.5%)	5 (6.1%)	3 (4.6%)
	More than half the days	2 (2.2%)	0 (0.0%)	3 (4.6%)
		1 (1.1%)	1 (1.2%)	0 (0.0%)

	Nearly every day

[1]Or the opposite–being so fidgety or restless that you have been moving around a lot more than usual. % were calculated based on the number of patients with non-missing assessments.

PHQ-9: Patient Health Questionnaire–9 Items.

Supplementary Table 3 Clinical endpoints at week 8 and week 52/final visit–completers population

Clinical outcome	Week 8 (*N* = 47)	Week 52/FV (*N* = 45)
Clinical response (SCCAI)	29 (65.9%)	35 (85.4%)
Clinical remission (SCCAI)	21 (47.7%)	30 (73.2%)
Endoscopic healing (Mayo Endoscopic Subscore)	5 (71.4%)	8 (80.0%)
Endoscopic healing (Fcal)		
Active disease	11 (84.6%)	10 (58.8%)
Clinical remission	1 (7.7%)	5 (29.4%)
Endoscopic healing	1 (7.7%)	2 (11.8%)
PGA		
Normal	16 (39.0%)	31 (70.5%)
Mild disease	12 (29.3%)	7 (15.9%)
Moderate disease	13 (31.7%)	6 (13.6%)
Severe disease		
PGA responder	30 (73.2%)	36 (83.7%)
Extracolonic feature: current		
Arthritis	3 (6.8%)	1 (2.4%)
None	41 (93.2%)	40 (97.6%)
Changes in the extracolonic feature from baseline		
Arthritis at both baseline and follow up	1 (2.3%)	0 (0%)
None at baseline and arthritis at follow-up	2 (4.5%)	1 (2.4%)
None at both baseline and follow up	35 (79.5%)	34 (82.9%)
None at follow-up and arthritis at baseline	5 (11.4%)	5 (12.2%)
None at follow up and pyoderma gangrenosum at baseline	1 (2.3%)	1 (2.4%)
With ≥ 1 hospitalization	0 (0%)	0 (0%)

With ≥ 1 important medical event	0 (0%)	0 (0%)
With ≥ 1 hospitalization or important medical event	0 (0%)	0 (0%)
Steroid use: since baseline	22 (48.9%)	23 (48.9%)
Steroid use: since last visit	22 (48.9%)	10 (21.3%)
Steroid use: current	10 (22.2%)	6 (12.8%)

Percentages were calculated based on the number of patients with non-missing values. Clinical response based on SCCAI: Decrease from baseline of ≥ 2. Clinical remission: SCCAI ≤ 2. Endoscopic healing (Mayo Endoscopic Subscore): subscore of 0 or 1. Fecal calprotectin: active disease: > 250 µg/g, clinical remission: 250 to 50 µg/g, endoscopic healing < 50 µg/g. PGA responder: Decrease from baseline of ≥ 1 point. PGA: Physician's global assessment; SCCAI: Simple Clinical Colitis Activity Index.

Supplementary Table 4 Patients with complications and steroid use – Intent-to-treat population

Clinical outcome	Week 8 (*N* = 94)	Week 52/FV (*N* = 73)
With ≥ 1 hospitalization	0 (0%)	3 (4.1%)
With ≥ 1 important medical event	0 (0%)	2 (2.7%)
With ≥ 1 hospitalization or important medical event	0 (0%)	4 (5.5%)
Steroid use: since baseline	43 (50.6%)	36 (49.3%)
Steroid use: since last visit	43 (50.6%)	21 (28.8%)
Steroid use: current	25 (29.4%)	14 (19.2%)

FV: Final visit.

Supplementary table 5 Inflammatory Bowel Disease Disability Index: Changes from baseline and proportion with improvement at week 8 and week 52/Final visit – Completers population

Visit	Change from baseline			With improvement	
	N	Mean (SD)	95% CI	*P*-Value	*N* (%)	95% CI
Week 8	43	-14.69 (17.99)	(-20.23, 9.16)	**<0.001**	35 (81.4%)	(69.8%, 93.0%)
Week 52/FV	36	-20.09 (17.74)	(-26.09, 14.09)	**<0.001**	32 (88.9%)	(78.6%, 99.2%)

CI: Confidence interval; FV: Final visit.

Supplementary table 6 Inflammatory Bowel Disease Disability Index: Sensitivity to change at week 8 and week 52/Final visit – ITT and completers populations

Outcome	Week 8 (*N* = 94)		Week 52/FV (*N* = 73)	
	ITT	Completers	ITT	Completers
IBD-DI: Sensitivity to change				
ES	-0.61	-0.75	-0.77	-1.08
SRM	-0.62	-0.82	-0.62	-1.13

ES: Effect size; FV: Final visit; IBD-DI: Inflammatory Bowel Disease Index, ITT: Intent to treat; SRM: Standardized response mean. ES = mean change divided by the standard deviation of the baseline scores. SRM = mean change divided by the standard deviation of change scores.

Supplementary table 7 Association between clinical outcomes and change in Patient Health Questionnaire – 9 Items total scores – Intent-to-treat and completers analyses

Clinical outcome	Value	Unadjusted LS means (95% CI)	Adjusted LS means (95% CI)
Intent-to-Treat			
Clinical response	Yes	-3.76 (-5.00, -2.53)	-2.50 (-4.93, -0.07)
	No	-1.05 (-2.27, 0.17)	0.80 (-1.47, 3.08)
	Difference (Yes – No)	-2.72 (-4.16, -1.27)	-3.30 (-4.77, -1.84)
	P-value	**<0.001**	**<0.001**
Clinical remission	Yes	-3.27 (-4.41, -2.13)	-2.08 (-3.56, -0.60)
	No	-0.66 (-2.10, 0.77)	1.24 (-0.46, 2.93)
	Difference (Yes – No)	-2.61 (-4.10, -1.12)	-3.32 (-4.86, -1.78)
	P-value	**<0.001**	**<0.001**
Completers			
Clinical response	Yes	-3.53 (-4.93, -2.13)	-3.82 (-5.27, -2.37)
	No	-2.59 (-4.28, -0.91)	-2.97 (-4.68, -1.26)
	Difference (Yes – No)	-0.93 (-2.73, 0.86)	-0.85 (-2.59, 0.90)
	P-value	0.304	0.339
Clinical remission	Yes	-3.42 (-4.76, -2.08)	-2.41 (-4.26, -0.55)
	No	-2.29 (-4.33, -0.24)	-0.70 (-3.07, 1.66)
	Difference (Yes – No)	-1.13 (-3.15, 0.89)	-1.71 (-3.73, 0.32)
	P-value	0.269	0.098

CI: Confidence interval; LS: Least squares.

Supplementary table 8 Change from baseline in other patient-reported outcomes at week 8 – Completers population

PRO measure	Baseline	Change from baseline		*P* value
	mean ± SD	*N*	Week 8 (mean ± SD)	
EQ-5D-5L	0.76 (0.21)	44	0.08 (0.19)	0.010
SIBDQ				
Total score	4.45 (1.17)	44	0.73 (1.14)	< 0.001
Social function	4.67 (2.03)	43	1.08 (1.94)	< 0.001
Emotional function	4.40 (0.76)	44	0.24 (0.76)	0.042
Bowel symptoms	4.41 (1.44)	44	0.87 (1.58)	< 0.001
Systemic symptoms	4.49 (1.69)	44	0.74 (1.35)	< 0.001
FACIT-F				
Fatigue subscale	31.92 (14.70)	44	4.49 (12.41)	0.021
Physical fatigue	18.08 (7.71)	44	3.70 (6.68)	< 0.001
Social impact of fatigue	21.15 (4.02)	44	0.42 (3.77)	0.461
Emotional fatigue	15.71 (4.98)	44	1.66 (4.15)	0.011
Functional fatigue	16.69 (6.07)	44	2.38 (5.54)	0.007
Trial outcome index	66.69 (27.09)	44	10.58 (23.23)	0.004
FACT-G total score	71.63 (18.79)	44	8.17 (16.37)	0.002
FACIT-F total score	103.55 (31.97)	44	12.66 (27.35)	0.004
MOS Sleep				
Sleep problems index I	36.60 (20.60)	44	-4.03 (18.00)	0.145
Sleep problems index II	38.61 (21.33)	44	-2.49 (18.40)	0.374

Supplementary Table 9 Change from baseline in other patient-reported outcomes at week 52–completers population

PRO measure	Baseline	Change from baseline		*P* value
	mean ± SD	*N*	Week 52 (mean ± SD)	
EQ-5D-5L	0.76 (0.21)	44	0.10 (0.22)	0.005
SIBDQ				
Total score	4.45 (1.17)	44	1.01 (1.16)	< 0.001
Social function	4.67 (2.03)	43	1.53 (2.01)	< 0.001
Emotional function	4.40 (0.76)	44	0.35 (0.94)	0.019
Bowel symptoms	4.41 (1.44)	44	1.31 (1.51)	< 0.001
Systemic symptoms	4.49 (1.69)	44	0.88 (1.57)	< 0.001
FACIT-F				
Fatigue subscale	31.92 (14.70)	44	7.10 (14.04)	0.002
Physical fatigue	18.08 (7.71)	44	5.27 (7.17)	< 0.001
Social impact of fatigue	21.15 (4.02)	44	2.30 (3.70)	< 0.001
Emotional fatigue	15.71 (4.98)	44	3.07 (5.07)	< 0.001
Functional fatigue	16.69 (6.07)	44	4.30 (6.04)	< 0.001
Trial outcome index	66.69 (27.09)	44	16.68 (25.09)	< 0.001
FACT-G Total Score	71.63 (18.79)	44	14.94 (17.61)	< 0.001
FACIT-F Total Score	103.55 (31.97)	44	22.05 (30.07)	< 0.001
MOS Sleep				
Sleep problems index I	36.60 (20.60)	44	-8.26 (16.48)	0.002
Sleep problems index II	38.61 (21.33)	44	-6.85 (17.40)	0.012

CI: Confidence interval; EQ-5D-5L: EuroQol 5-Dimensions, 5 Levels; FACIT-F: Functional Assessment Chronic Illness Therapy-Fatigue; MOS: Medical Outcomes Study; PRO: Patient-reported outcome; SIBDQ: Short Quality of Life in Inflammatory Bowel Disease Questionnaire. Observed changes in scores were tested by paired sample t-test.

Supplementary Table 10 Changes in patient-reported outcome measures from baseline by clinical outcomes–intent-to-treat population

PRO measure		LS means (95% CI)	
		Clinical response	**Clinical remission**
IBD Disability Index	Clinical outcome = Yes	-18.11 (-21.80, -14.41)	-21.54 (-25.64, -17.44)
	Clinical outcome = No	-2.24 (-6.98, 2.50)	-5.34 (-9.39, -1.30)
	Difference (Yes – No)	-15.87 (-20.73, -11.00)	-16.20 (-21.06, -11.34)
	P value	**< 0.001**	**< 0.001**
EQ-5D-5L	Clinical outcome = Yes	0.06 (0.03, 0.10)	0.09 (0.05, 0.12)
	Clinical outcome = No	-0.01 (-0.06, 0.03)	-0.00 (-0.04, 0.03)
	Difference (Yes – No)	0.08 (0.03, 0.12)	0.09 (0.05, 0.13)
	P value	**0.002**	**< 0.001**
SIBDQ total score	Clinical outcome = Yes	0.93 (0.73, 1.13)	1.17 (0.96, 1.37)
	Clinical outcome = No	0.14 (-0.11, 0.40)	0.22 (0.02, 0.42)
	Difference (Yes – No)	0.78 (0.52, 1.05)	0.95 (0.70, 1.19)
	P value	**< 0.001**	**< 0.001**
FACIT-F fatigue subscale	Clinical outcome = Yes	6.90 (4.57, 9.23)	8.28 (5.77, 10.78)
	Clinical outcome = No	1.12 (-1.80, 4.03)	1.82 (-0.65, 4.29)
	Difference (Yes – No)	5.79 (2.79, 8.78)	6.46 (3.59, 9.33)
	P value	**< 0.001**	**< 0.001**
MOS Sleep – Sleep problems index I	Clinical outcome = Yes	-6.83 (-10.18, -3.48)	-6.25 (-9.99, -2.50)
	Clinical outcome = No	-2.65 (-6.90, 1.60)	-4.22 (-7.92, -0.52)
	Difference (Yes – No)	-4.18 (-8.60, 0.24)	-2.03 (-6.33, 2.27)

	P value	0.064	0.354
MOS Sleep – Sleep problems index II	Clinical outcome = Yes	-6.39 (-9.78, -2.99)	-6.55 (-10.27, -2.83)
	Clinical outcome = No	-2.84 (-7.06, 1.38)	-3.48 (-7.15, 0.19)
	Difference (Yes – No)	-3.55 (-7.84, 0.75)	-3.07 (-7.23, 1.09)
	P value	0.105	0.147

CI: Confidence interval; EQ-5D-5L: EuroQol 5-Dimensions, 5 Levels; FACIT-F: Functional Assessment Chronic Illness Therapy-Fatigue; IBD: Inflammatory bowel disease; MOS: Medical Outcomes Study; PRO: Patient-reported outcome; SIBDQ: Short Quality of Life in Inflammatory Bowel Disease Questionnaire.

Clinical response based on SCCAI: Decrease from baseline of ≥ 2. Clinical remission: SCCAI ≤ 2.

Supplementary Table 11 Changes from baseline in productivity at week 52: Work Productivity and Activity Impairment and Valuation of Lost Productivity–completers population

Outcome	N	Baseline (% or Mean)	Week 52/FV (% or Mean)	Change from baseline Mean	95%CI
WPAI					
Percent work time missed due to health	37	19.7%	10.3%	-9.4%	(-17.5%, -1.5%)
Percent work impairment while working due to health	38	40.0%	25.5%	-14.5%	(-22.6%, -6.4%)
Percent overall work impairment due to health	35	44.8%	30.2%	-14.5%	(-22.2%, -6.8%)
Percent activity impairment due to health	64	43.1%	26.4%	-16.7%	(-24.5%, -9.2%)
VOLP					
Any paid work productivity loss in the past x months (%)	27	74.1%	37.0%	-37.0%	(-54.8%, -19.4%)
Paid work productivity loss in the past x months (h)	27	82.9	19.3	-63.6	(-100.1, -31.7)
Any unpaid work productivity loss in the past days (%)	731	25.8%	19.4%	-6.5%	(-22.6%, 9.7%)
Unpaid work productivity loss in the past days (h)	731	4.9	2.3	-2.6	(-9.3, 3.2)
Any costs of lost productivity in the past month (%)	x27	74.1%	48.2%	-25.9%	(-45.2%, -9.7%)
Total costs of lost productivity in the past months ($)	x27	4692.1	1326.4	-3365.7	(-6019.1, -1159.8)

CI: Confidence interval; FV: Final visit; VOLP: Valuation of Lost Productivity; WPAI: Work Productivity and Activity Impairment.