CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

125476Orig1s000

STATISTICAL REVIEW(S)

Statistical Team Leader Memorandum

Submission: BLA 125476/002
Product: Entyvio (vedolizumab) 300 mg for injection, for intravenous (IV) use
Sponsor: Takeda Pharmaceuticals U.S.A., Inc. (Takeda)
Indication: Treatment of patients with moderately to severely active ulcerative colitis (UC) or Crohn's Disease (CD)
Medical Division: Division of Gastroenterology and Inborn Errors Products (DGIEP)

Reference: Statistical Review and Evaluation for UC dated May 15, 2014 and Statistical Review and Evaluation for CD dated May 19, 2014.

The purpose of this memorandum is to summarize conclusions regarding the statistical issues discussed in the primary reviewer's evaluations of this original BLA submission, and to present the Team Leader's perspective on the study results.

Takeda submitted this BLA in support of marketing approval of vedolizumab for the treatment of moderately to severely active ulcerative colitis (UC) and for the treatment of Crohn's Disease (CD). Vedolizumab is a humanized immunoglobulin G1 (IgG1) monoclonal antibody directed against the human lymphocyte integrin $\alpha_4\beta_7$. The inflammatory bowel disease (IBD) is primarily comprised of UC and CD conditions.

The phase 3 vedolizumab studies assessed the safety and efficacy of vedolizumab in conjunction with conventional IBD therapies such as 5-ASAs, immunomodulators, and corticosteroids for the treatment of patients with UC or CD. Four phase 3 studies were conducted and the three efficacy studies were evaluated in the primary review:

- Study C13006: Induction and maintenance; UC
- Study C13007: Induction and maintenance; CD
- Study C13011: Induction; CD
- Study C13008: Long-term safety; UC or CD

Ulcerative Colitis (UC)

Study C13006 was designed and conducted to assess the safety and efficacy of vedolizumab for the induction and maintenance of clinical response and remission in patients with moderately to severely active UC. This study was a phase 3, multinational, randomized, double-blind, placebo-controlled trial. The induction and maintenance therapies were studied in two separate phases within Study C13006.

Induction

The 6-week Induction Phase contained two sequentially enrolled cohorts of patients. Cohort 1 patients were randomized to induction treatment with two doses of vedolizumab 300 mg intravenous (IV) injection or placebo administered at Weeks 0 and 2; cohort 1 comprised the

primary efficacy population. Cohort 2 patients all received the vedolizumab induction treatment in an open-label manner.

The primary Induction efficacy endpoint was clinical response, while the secondary endpoints were clinical remission and mucosal healing, all evaluated at Week 6. These endpoints were sequentially tested in a pre-specified order. For the primary analyses of these efficacy endpoints, patients who withdrew from the study prematurely were classified as treatment failures. The study results demonstrated statistically significant treatment benefit of vedolizumab compared with placebo for the primary and secondary efficacy endpoints.

The primary reviewer conducted an exploratory analysis using a different definition of clinical remission than that pre-specified in the protocol. The statistical insignificance concluded in the primary review should be viewed with caution due to the exploratory nature of this analysis. Such a result might be expected because the study was not designed or powered to show statistical significance on this endpoint. Moreover, both treatment groups had less than 5% of the patients achieving clinical remission by this stringent definition while more than twice the patients in each treatment group achieved clinical remission by the pre-specified definition. It is arguable that this alternative definition may not be suitable for a clinical trial.

Other definitions of both clinical remission and mucosal healing were explored in the additional analyses by the primary reviewer. Moreover, extensive sensitivity analyses using different imputation methods on the missing data and subgroup analyses based on various demographics and baseline characteristics were requested by the Agency and conducted by the sponsor. All the results showed a favorable treatment effect for vedolizumab compared to placebo. The subgroup analyses showed an expected variability of the treatment effect. The statistical significance stated in the primary review, including the discussion on the 95% confidence interval (CI) coverage, for all these analyses should be viewed with caution due to their exploratory nature. The main point of these analyses should be to present descriptive statistics and point out possible data relationships for further investigation.

Maintenance

Patients who achieved a clinical response to vedolizumab induction therapy in either Cohort 1 or Cohort 2 were randomized to maintenance treatment with 300 mg vedolizumab every four weeks (Q4W), every eight weeks (Q8W), or placebo in the 46-week Maintenance Phase. The study weeks were numbered continuously from the Induction Phase; hence, the Maintenance Phase was from Week 6 to Week 52. The treatment administered at Week 6 was the first dose of the Maintenance Phase therapy.

The primary Maintenance efficacy endpoint was clinical remission, while the secondary endpoints were durable clinical response, mucosal healing, durable clinical remission, and corticosteroid-free remission, all evaluated at Week 52. These endpoints were sequentially tested in a pre-specified order as were the Induction efficacy endpoints. The study results demonstrated statistically significant treatment benefit of vedolizumab compared with placebo for the primary and secondary efficacy endpoints.

For the primary analyses of these primary and secondary efficacy endpoints, patients who withdrew from the study prematurely were classified as treatment failures as was done for the Induction Phase. The primary reviewer noted that the number of patients with missing data in the placebo group was nearly twice that in the vedolizumab group, and that this could bias the treatment effect estimate in favor of vedolizumab. However, the missing data rates observed in this trial may be expected given the length of these trials and are consistent with those from other UC maintenance trials. Moreover, the majority of discontinuation in all treatment groups was due to adverse events (AEs) or lack of efficacy. The sponsor's imputation strategy in the primary analyses is defendable and the proper inferential statistics for the primary comparisons should be based on the pre-specified primary analyses for the primary and secondary efficacy endpoints.

The sponsor's gate-keeping testing procedure with the Hochberg method for the multiple dose comparisons was inadequate because of the non-separable property of the Hochberg method, i.e., it may not properly preserve the study-wise type I error rate. However, all the comparisons of the primary and secondary efficacy endpoints have shown statistical significance and it is no longer a concern.

The Agency requested the sponsor conduct extensive sensitivity analyses using different imputation methods on the missing data, and perform subgroup analyses based on various demographics and baseline characteristics (including the induction cohorts). All these results showed a favorable treatment effect for vedolizumab compared to placebo. The subgroup analyses showed an expected variability of the treatment effect. The statistical significance stated in the primary review, including the discussion on the 95% CI coverage, regarding these analyses should be viewed with caution due to their exploratory nature, and focus should be on the descriptive statistics.

Conclusion

In summary, Study C13006 showed statistically significant benefit of vedolizumab compared to placebo for both induction and maintenance therapy, as demonstrated by the pre-specified primary and secondary efficacy endpoints.

Crohn's Disease (CD)

Study C13007, very similar in design as the UC study (C13006), was conducted to assess the safety and efficacy of vedolizumab for the induction and maintenance of clinical response and remission in patients with moderately to severely active CD.

A second phase 3 study (C13011) was conducted for the induction treatment only in patients with moderately to severely active CD. However, the primary comparison was conducted on the subpopulation consisted of patients who had failed TNFα antagonist therapy, which was roughly 75% of the study population. Study C13011 failed to demonstrate statistically significant

treatment difference for the primary comparison and so all the other comparisons pre-specified in the protocol can only be viewed as exploratory.

Induction

Both studies (C13007 and C13011) contained a 6-week Induction Phase while only Study C13007 engaged two sequentially enrolled cohorts of patients. Study C13011, although as an induction study, had patients administered the vedolizumab injection at Week 6 and explored a longer induction duration at Week 10.

For Study C13007, the original primary Induction efficacy endpoint was clinical response, while the secondary endpoints were originally enhanced clinical remission and change in C-reactive protein (CRP), all evaluated at Week 6. During the study, the sponsor amended the protocol to identify enhanced clinical remission as a so-called "co-primary" endpoint and proposed the Hochberg method to adjust for the two primary comparisons. In the same amendment, the CDAI score for eligibility was amended to lower the upper limit from 480 to 450. The sponsor claimed that both modifications were based on blinded demographic data from the first 50 patients accrued into the study.

The study results demonstrated statistically significant treatment benefit of vedolizumab compared with placebo in the original primary endpoint of clinical response at Week 6 and failed to show statistical significance on both the other primary endpoint of enhanced clinical remission and the secondary efficacy endpoint of CRP at Week 6. The study should be considered successful per the final protocol specifications.

The sponsor's gate-keeping testing procedure with the Hochberg method for the multiple primary comparisons was inadequate as it may not properly preserve the study-wise type I error rate. Although only one of the comparisons of the primary efficacy endpoints showed statistical significance, the secondary efficacy endpoint failed to show statistical significance even with a nominal p-value; and so the inadequacy of the pre-specified multiplicity adjustment method is no longer a concern.

The primary reviewer conducted several exploratory analyses on the patients with baseline CDAI score lower than the pre-specified lower limit of 220 or higher than the higher limit of 450. However, the screening CDAI score rather than the baseline CDAI score should be used for these analyses because the screening records were used for eligibility determination. There were 8 patients, instead of the 20 patients identified in the primary review, evenly distributed between the two treatment groups, who had protocol violations by having a screening CDAI score less than 220. Moreover, the primary reviewer ignored the aforementioned amendment and neglected to specify the patients with protocol violations by having a screening CDAI score greater than 450 or 480 according to their enrollment time. There were 15 patients (8 in the placebo group and 7 in the vedolizumab group), as opposed to the 18 patients identified in the primary review, enrolled with such a violation. Finally, the primary reviewer applied the Fisher's exact test, instead of the pre-specified Cochran-Mantel-Haenszel (CMH) test for these exploratory analyses. With the relatively small treatment effect size, and the discrete nature of

the data, the sensitivity of the p-value to few patients' data or to the use of an exact test of proportions is not an unexpected result nor one that should necessarily have been a significant review issue. Furthermore, the assumptions underlying the sponsor's use of the CMH test statistic for Study C13007 are defensible, and the proper p-value for the primary comparison should be based on that analysis with the pre-specified primary analysis population. With all that said, the statistical insignificance stated in the primary review should be viewed with caution due to the exploratory nature of these exploratory analyses.

Additional analyses using different CMH tests and the Fisher's exact test were also conducted by the primary reviewer. Moreover, an exploratory analysis using an alternative definition of clinical remission, extensive sensitivity analyses using different imputation methods on the missing data, and subgroup analyses based on various demographics and baseline characteristics were requested by the Agency and conducted by the sponsor. All the results showed a favorable treatment effect for vedolizumab. The subgroup analyses showed an expected variability of the treatment effect. The statistical significance stated in the primary review, including the discussion on the 95% CI coverage, should be viewed with caution due to the exploratory nature of these analyses. The main objective of these exploratory analyses should be to present descriptive statistics of interesting data relationships that may be important for future investigation.

Maintenance

The primary Maintenance efficacy endpoint for Study C13007 was clinical remission, while the secondary endpoints were enhanced clinical response, corticosteroid-free remission, and durable clinical remission, all evaluated at Week 52. These endpoints were sequentially tested in a pre-specified order and a gate-keeping testing procedure with the Hochberg method was proposed by the sponsor for the multiple dose comparisons. The study results demonstrated a statistically significant treatment benefit for vedolizumab compared with placebo in the primary and the first two secondary efficacy endpoints, while the results failed to show statistical significance for the last secondary efficacy endpoint of durable clinical remission.

The sponsor's gate-keeping testing procedure with the Hochberg method for the multiple dose comparisons was inadequate as it may not properly preserve the study-wise type I error rate. However, only the last secondary efficacy endpoint has failed to show statistical significance and it is no longer a concern.

Induction cohort 2 comprised the majority of the maintenance patients. The primary reviewer emphasized the differences between the two induction cohorts and that a larger treatment effect was observed in the cohort 2 patients. One should note that because of the limitation of enrolling approximately 50% of patients who had failed TNFα antagonist therapies in cohort 1 and the sequential enrollment of cohort 2, different presentations of the patient populations for the two induction cohorts were inevitable. Moreover, some variability in the treatment effect across subgroups was to be expected. Once again, the statistical significance stated in the primary review, including the discussion on the 95% CI coverage, should be viewed with caution due to the exploratory nature of this subgroup analysis.

The primary reviewer noted the relatively high missing data rate; however, similar missing data patterns were observed across the three treatment groups for this study. The missing data rates observed in this trial would not be unexpected given the length of this type of trial and they seem to be consistent with missing data patterns observed in previous CD trials. Moreover, the majority of discontinuation in all treatment groups was due to adverse events (AEs) or lack of efficacy. With similar missing data rates, treating missing values as treatment failures should not exaggerate the treatment effect size. Hence, the imputation strategy in the primary analyses was justified, and the proper inferential statistics for the primary comparisons should be based on the pre-specified primary analyses for the primary and secondary efficacy endpoints.

The Agency requested the sponsor conduct extensive sensitivity analyses using different imputation methods for the missing data, and perform subgroup analyses based on various demographics and baseline characteristics (including the induction cohorts). All the results showed a favorable treatment effect for vedolizumab compared to placebo. The subgroup analyses showed an expected variability of the treatment effect. The statistical significance stated by the primary reviewer, including the discussion on the 95% CI coverage, for all these analyses should be viewed with caution due to their exploratory nature, and focus should be on the descriptive statistics.

Conclusion

In summary, Study C13007 showed statistically significant benefit of vedolizumab compared to placebo for treatment of CD, as demonstrated by one of the Induction primary efficacy endpoints, the Maintenance primary efficacy endpoint and two of the three Maintenance secondary efficacy endpoints. However, treatment effect sizes observed were relatively small and an induction period longer than six weeks may be needed for some patients to achieve clinical response.

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

FREDA COONER
05/19/2014

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

BLA #:	125-476
Drug Name:	Entyvio (vedolizumab) (MLN002)
Indication:	For injection for the treatment of patients with moderately to severely active Crohn's Disease
Applicant:	Takeda Pharmaceuticals U.S.A. Inc.
Date(s):	Received June 20, 2013 PDUFA: February 18, 2014 (extended to May 20, 2014)
Review Priority:	Priority
Biometrics Division:	Division of Biometrics III (DB III)
Statistical Reviewer:	Milton C. Fan, Ph.D. , DB III
Concurring Reviewers:	Freda Cooner, Ph.D., Team Leader, DB III
Medical Division:	Gastroenterology and Inborn Errors Products (DGIEP)
Clinical Team:	Klaus Gottlieb, M.D., Anil Rajpal, M.D. (TL) (DGIEP)
Project Manager:	Kevin Bugin (DGIEP)

Keywords: clinical studies, BLA review, placebo controlled

Table of Contents

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

The applicant has submitted the results of two phase 3 studies (C13007 and C13011) to support the indication of Crohn's disease (CD).

Both studies were conducted to evaluate vedolizumab as an induction therapy for moderate to severe CD. Study C13007 evaluated vedolizumab 300 mg in CD patients, of which 50% patients were naïve to TNFα antagonists and 50% with previous TNFα antagonists. Study C13011 included approximately 75% patients who had previously failed TNFα antagonist therapy and approximate 25% patients who were naïve to TNFα antagonist therapy.

Two primary efficacy endpoints, clinical remission and enhanced clinical response at Week 6, were pre-specified for Study C13007. Based on the Induction Study ITT Population, study C13007 showed that a statistically significantly greater proportion of vedolizumab-treated patients achieved clinical remission at Week 6 compared with patients who received placebo. The treatment difference from placebo was 7.8% (95% CI: 1.2, 14.3; p = 0.0206).

However, the treatment comparisons on both the other primary endpoint of enhanced clinical response at Week 6 and the secondary efficacy endpoint of changes from baseline in CRP at Week 6, failed to achieve statistical significance.

For subjects who failed TNFα antagonist therapy, Study C13011 failed to demonstrate a statistically significant difference between the vedolizumab and placebo groups for the proportion of patients in clinical remission at Week 6 with treatment difference of 3.0% (95% CI: -4.5, 10.5; p=0.4332). Study C13007 revealed treatment difference of 6.2% (95% CI: -9.1, 21.3) in this subpopulation favoring vedolizumab.

Only Study C13007 was performed to evaluate vedolizumab as a maintenance therapy for moderate to severe CD. The results from Study C13007 Maintenance Study showed statistically significant difference on the primary endpoint of clinical remission at Week 52 for the every eight weeks (Q8W) regimen. Statistically significant treatment differences were also observed for two of the three pre-specified key secondary efficacy endpoints.

However, for the maintenance phase, results from Study C13007 by the Induction Phase Cohort were notable different between dosing regimens (Q4W vs. Q8W) and between the Induction Cohorts. The results for vedolizumab Q8W against placebo from the overall analysis may be dominated by thoset of Cohort 2**.**

1.2 Brief Overview of Clinical Studies

1.2.1 Study C13007

This study was a pivotal, phase 3, multicenter, multinational, randomized, double-blind, placebo-controlled trial conducted to evaluate the efficacy and safety of both induction and maintenance treatment with vedolizumab in patients with moderately to severely active CD, defined as a Crohn's Disease Activity Index (CDAI) score of 220 to 450 points. This multinational study was conducted at 285 sites. This trial was designed to support the registration of vedolizumab for induction and maintenance treatment of a broad population of patients who have failed one or more standard therapies for CD, including immunomodulators (azathioprine, 6-MP, or methotrexate) and TNFα antagonists. For study centers outside of the US, patients could have also failed treatment with corticosteroids. To ensure that the efficacy of vedolizumab could be evaluated in patients who are naïve to TNFα antagonists, enrollment of patients with previous TNFα antagonist exposure was to be limited to no more than 50% of the overall study population

This study was designed to comprise two randomized, double-blind, placebo-controlled studies conducted under one protocol which, operationally, consisted of two phases:

- The Induction Phase, designed to establish the efficacy and safety of vedolizumab for the induction of clinical response and clinical remission, and

- The Maintenance Phase, designed to establish the efficacy and safety of vedolizumab for the maintenance of clinical response and clinical remission.

1.2.2 Study C13011

This study was a phase 3, multinational, randomized, double-blind, placebo-controlled trial conducted to evaluate the efficacy and safety of vedolizumab for the induction of clinical response and remission in patients with moderately to severely active CD.

The inclusion and exclusion criteria were similar to those for Study C13007 Induction Phase with exception of CDAI score. In this study, moderately to severely active CD was determined by a CDAI of 220 to 400 instead of 200 to 480 used in Study C13007.

After a 21-day screening period, patients were randomized in a 1:1 ratio to receive either 300 mg vedolizumab or placebo at Weeks 0, 2, and 6. Enrollment of patients was monitored by the interactive voice response system (IVRS) to ensure that approximately 75% of the overall population had previously failed TNFα antagonist therapy and approximately 25% were naïve to TNFα antagonist therapy.

1.3 Statistical Issues and Findings

1.3.1 Induction Studies

Two studies (C13007 and C13011) were conducted to evaluate vedolizumab as induction therapy for moderate to severe CD.

During the Study C13007 Induction Phase, the applicant elevated the first key secondary endpoint, enhanced clinical response (decrease in CDAI of ≥ 100 points), to a "co-primary" endpoint. The applicant further specified that the primary objective of the study would be met by achieving statistical significance for either of the co-primary endpoints, and the Hochberg method would be used to adjust for multiple comparisons.

There was disagreement between the applicant and FDA regarding definition of co-primary endpoints. The following statements were conveyed to the applicant in the meeting held on September 10, 2009.

- The term of co-primary endpoint that you have defined for Study C13007is not commonly used for regulatory purposes.

- Two or more primary endpoints are called co-primary if each must be show statistically significant treatment benefit at a pre-specified significance level α (e.g., α=0.025, by one-sided tests).

The applicant performed analyses of clinical remission at Week 6 and enhance clinical response at Week 6 using the Cochran-Mantel-Haenszel (CMH) test, with stratification according to:

1) concomitant use of oral corticosteroids (yes/no);
2) previous exposure to TNFα antagonists (yes/no);
3) concomitant immunomodulator use (yes/no).

Based on the Induction Study ITT Population, a statistically significantly greater proportion of vedolizumab-treated patients achieved clinical remission at Week 6 compared with patients who received placebo. The treatment difference was 7.8% (95% CI: 1.2, 14.3; p = 0.0206).

The difference between the vedolizumab and placebo groups was not statistically significant for the other primary endpoint of enhanced clinical response at Week 6. The difference was 5.7% (95% CI: -3.6, 15.0; p = 0.2322).

This reviewer found that 20 ITT patients (10 patients in each group) who had a baseline CDAI score of less than 220 and 2 patients (1 patient in each group) with baseline CDAI missing were enrolled in this study.

Among these 20 patients with a baseline CDAI < 220, a greater proportion of vedolizumab-treated patients achieved clinical remission at Week 6 as compared with patients who received placebo [70% (7/10) vs. 20% (2/10)].

This reviewer also found 18 patients (9 in each treatment group) who had a baseline CDAI score of greater than 450. Among these 18 patients with a baseline CDAI > 450, proportions of vedolizumab-treated and placebo patients achieved clinical remission at Week 6 were both zeros.

According to the inclusion criteria (a CDAI score of 220 to 450), these 40 patients should not be enrolled in the study, this reviewer performed post-hoc sensitivity analysis by excluding these 40 patients. The resulting treatment different would be 6.25% with nominal p-value of 0.0893 (Fisher's Exact test). If these 40 patients were considered as "non-responders", the resulting treatment difference would be 5.95% with a nominal p-value of 0.0622 (Fisher's Exact test).

This reviewer also found that 16 Per-Protocol (PP) patients (8 patients in each group) who had a baseline CDAI score of less than 220 were included in the PP analysis in this study.

Among these 16 patients with baseline CDAI < 220, a greater proportion of vedolizumab-treated patients achieved clinical remission at Week 6 as compared with patients who received placebo [62.5% (5/8) vs. 12.5% (1/8)].

According to the final inclusion criteria (a CDAI score of 220 to 450), these 16 patients should not be enrolled in the study. This reviewer performed a post-hoc sensitivity analysis by excluding these 16 patients. The treatment different would be 6.67% with a nominal p-value of 0.0606 based on the Fisher's Exact test. If these 16 patients were considered as "non-responders", the treatment difference would be 6.53% with a nominal p-value of 0.0611 based on the Fisher's Exact test.

This reviewer performed a post-hoc Breslow-Day test to evaluate the homogeneity of subgroup by the baseline CDAI (≤330 vs. >330). The p-value from the Breslow-Day test yielded 0.0636 which is smaller than 0.10, the usual level of significance used testing for interaction. It was suggested that vedolizumab might be more effective for patients with baseline CDAI ≤ 330. But, this finding can only be confirmed by the new study.

Per FDA's requested, for the Induction Study ITT Population for Studies C13007, the applicant performed a post hoc analysis using the following alternative definition of clinical remission:

- Total number of liquid/very soft stools of ≤ 10 per day in the relevant week; and
- Abdominal pain rated as 0 or 1 for each day in the relevant week.

Based on alternative definition from FDA, vedolizumab-treated patients failed to achieve statistical significance for clinical remission at Week 6 for vedolizumab group with the difference of 4.8% (95 CI: -0.7, 10.3; p = 0.0848).

For subjects who failed TNFα antagonist therapy, Study C13011 failed to demonstrate a statistically significant difference between the vedolizumab and placebo groups for the

proportion of patients in clinical remission at Week 6 with treatment difference of 3.0% (95% CI: -4.5, 10.5; p=0.4332). Study C13007 revealed treatment difference of 6.2% (95% CI: -9.1, 21.3) in this subpopulation favoring vedolizumab.

1.3.2 Maintenance Study

One study (C13007) was conducted to evaluate vedolizumab as maintenance therapy for moderate to severe CD.

The Maintenance Phase included three groups of patients who were assigned to treatment groups based on their Induction Phase treatment assignments and responses to the study therapy. Vedolizumab-treated patients from both Induction Cohort 1 (double-blind) and Cohort 2 (open-label) who demonstrated a clinical response according to protocol-specified criteria, as assessed by the investigator, were randomized in a 1:1:1 ratio to double-blind treatment with vedolizumab administered every 4 weeks (Q4W), vedolizumab administered every 8 weeks (Q8W), or placebo. Randomization was stratified by three factors:

- Enrollment in Cohort 1 or Cohort 2 in the Induction Phase
- Concomitant use of oral corticosteroids
- Previous exposure to TNFα antagonists or concomitant immunomodulator use

These patients who were randomized into the Maintenance Phase comprised the Maintenance Study ITT Population, the primary efficacy population.

Vedolizumab-treated patients, who did not demonstrate response at Week 6 of the Induction Phase continued treatment with open-label vedolizumab, administered Q4W. Patients who had been treated with double-blind placebo in the Induction Study continued on double-blind placebo during the Maintenance Phase, regardless of the treatment response during induction. The Maintenance Phase began at Week 6, included study drug dosing at Week 6 and Q4W or Q8W thereafter, and concluded with Week 52 assessments.

In the Induction Phase, a total of 220 vedolizumab patients were enrolled into Cohort 1; a total of 747 additional patients were enrolled into Cohort 2. Among 220 vedolizumab patients in Cohort 1, 96 patients (43.6%) were Week 6 responders. Among 747 additional patients in Cohort 2, 365 patients (48.8%) were Week 6 responders.

A majority of patients (79%) who were randomized into the Maintenance Phase were from Cohort 2. Compared to Cohort 1, Cohort 2 had a greater proportion of patients who had prior TNFα antagonist use (68%) and failure (63%) vs. 50% and 48%, respectively. Cohort 2 also had more patients at sites in Western/Northern Europe and fewer patients entering at sites in Asia/Australia/Africa and Eastern Europe than was observed for the Cohort 1.

The applicant noted that an imbalance across the treatment groups in the proportion of patients who had achieved clinical remission at Week 6 was observed due to the randomization at Week 6 was not stratified by the remission status. Clinical remission at Week 6 was achieved by 27.9% of the patients in the vedolizumab Q4W group and 33.8% of the patients in the vedolizumab Q8W group compared with 36.6% of the patients in the placebo group.

This imbalance may have had an impact on the analyses of the clinical remission-based endpoints in favor of placebo, especially for the vedolizumab Q4W group versus the placebo group because the clinical remission rate at Week 6 for the vedolizumab Q4W group was about 9% lower than that of the placebo .

For the Maintenance Study, the applicant used a Hochberg and sequential testing procedure for the primary and secondary endpoints in order to maintain the overall Type I error rate of 0.05. This multiplicity adjustment method may not be able to properly control the study-wise Type I error rate.

Hence, results from secondary efficacy endpoints were difficult to interpret from statistical perspective.

There were more than 58% of the primary endpoint data missing for placebo, more than 53% of the data missing for vedolizumab Q8W and more than 47% of the data missing for vedolizumab Q4W. Although the percentage of missing data is consistent with that observed from the CD clinical trials with similar designs, it may introduce difficulties in the interpretation of the study results.

2. INTRODUCTION

2.1 Overview

Vedolizumab is a humanized monoclonal antibody that binds to the α_4 β_7 integrin, which is expressed on discrete populations of leukocytes involved in gut mucosal immunity. The mechanism of action of vedolizumab (MLN0002) reduces pathological bowel inflammation, thus providing a potential therapeutic option for patients with inflammatory bowel disease (IBD).

The applicant seeks marketing approval for vedolizumab as an injection for the treatment of patients with moderately to severely active ulcerative colitis (UC) and Crohn's disease (CD).

2.2 Data Sources

The applicant has submitted three phase 3 studies (C13006, C13007, and C13011) for the proposed indication of injection for the treatment of patients with moderately to severely active (UC) or (CD).

This review will focus on the studies (C13007 and C13011) for the CD indication.

These three studies were entitled as follows:

- Clinical Protocol C13006: A Phase 3, Randomized, Placebo-Controlled, Blinded, Multicenter Study of the Induction and Maintenance of Clinical Response and Remission by Vedolizumab (MLN0002) in Patients with Moderate to Severe Ulcerative Colitis

- Clinical Protocol C13007: A Phase 3, Randomized, Placebo-Controlled, Blinded, Multicenter Study of the Induction and Maintenance of Clinical Response and Remission by Vedolizumab (MLN0002) in Patients with Moderate to Severe Crohn's Disease

- Clinical Protocol C13011: "A Phase 3, Randomized, Placebo-controlled, Blinded, Multicenter Study of the Induction of Clinical Response and Remission by Vedolizumab (MLN0002) in Patients with Moderate to Severe Crohn's Diseases".

This original submission was submitted in an eCTD format dated June 20, 2013.

The electronic submission is located at \\cdsesub1\bla\eCTD_Submissions\STN125476\0002\.

The applicant submitted a response, dated September 9, 2013, to this reviewer's Information Request dated August 19, 2013.

The applicant submitted a response, dated October 4, 2013, to this reviewer's Information Request dated September 20, 2013.

The applicant submitted a response, dated October 21, 2013, to this reviewer's Information Request dated October 7, 2013.

The applicant submitted a response, dated January 31, 2014, to the medical reviewer's Information Request dated January 17, 2014.

The applicant submitted a response, dated February 25, 2014, to this reviewer's Information Request dated February 19, 2014.

The applicant submitted a correction on April 2, 2014 to response to request to the medical reviewer's Information Request dated January 17, 2014.

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

3.1.1 Study C13007

3.1.1.1 Study Design

This study was a pivotal, phase 3, multicenter, multinational, randomized, double-blind, placebo-controlled trial conducted to evaluate the efficacy and safety of both induction and maintenance treatment with vedolizumab in patients with moderately to severely active CD, defined as a Crohn's Disease Activity Index (CDAI) score of 220 to 450 points. This multinational study was conducted at 285 sites. This trial was designed to support the registration of vedolizumab for induction and maintenance treatment of a broad population of patients who have failed one or more standard therapies for CD, including immunomodulators (azathioprine, 6-MP, or

methotrexate) and TNFα antagonists. For study centers outside of the US, patients could have also failed treatment with corticosteroids. The applicant stated that to ensure the efficacy of vedolizumab could be evaluated in patients who are naïve to TNFα antagonists, enrollment of patients with previous TNFα antagonist exposure was to be limited to no more than 50% of the overall study population

This study was designed to comprise two randomized, double-blind, placebo-controlled studies conducted under one protocol which, operationally, consisted of two phases:

- The Induction Phase, designed to establish the efficacy and safety of vedolizumab for the induction of clinical response and clinical remission, and

- The Maintenance Phase, designed to establish the efficacy and safety of vedolizumab for the maintenance of clinical response and clinical remission.

Patients in the Induction Phase were to continue in the Maintenance Phase according to protocol-defined criteria. Although conducted under one protocol for operational efficiency, the two phases described above included two separate sequential double-blind, placebo-controlled efficacy studies. Patients who met the protocol-specified criteria for clinical response during the induction phase were eligible for randomization into the maintenance efficacy study. Each study has distinct endpoints, randomization schema, pre-specified populations, and analysis plans.

Primary and secondary efficacy assessments for the Induction and Maintenance Phases were based on CDAI scores and CRP levels. A CDAI score was to be obtained during screening, using patient diary entries within 14 days prior to the enrollment, and hematocrit results within 7 days prior to enrollment. A CDAI score was also to be derived at Weeks 0, 2, 4, 6, 10, 14, 18, 22, 26, 30, 34, 38, 42, 46, 50 and 52 (or early termination [ET] visit) and at any unscheduled visit(s) due to a disease exacerbation. On all dosing days except Week 6, the CDAI score components were to be assessed prior to dosing; the total CDAI score was to be calculated once results were available for all components.

The Week 6 total CDAI score was to be calculated prior to dosing by the investigator or designee and recorded in the patient's source documents. This assessment determined whether the patient had achieved clinical response at Week 6, and therefore determined treatment assignment in the Maintenance Phase.

Extraintestinal manifestations of CD were to be collected as part of the CDAI. The data collected for CDAI were also to be used to derive the Harvey Bradshaw Index (HBI) score at the time points listed above for exploratory efficacy analyses.

Blood samples were obtained at Weeks 0, 6, 22, 38, and 52 (or ET visit) for determination of CRP levels.

Data conventions for the primary and key secondary endpoints for the Induction and Maintenance Studies deemed patients who prematurely discontinued as had not achieved the endpoint of interest at all the time points after the discontinuation.

For the demographic and baseline characteristics, summary tabulations are presented by treatment group and displayed with the number of observations, mean, standard deviation (std. dev.), median, minimum, and maximum for continuous variables, and the number and percent per category for categorical data. Data from both the Induction Phase and the Maintenance Phase were unblinded after all patients either completed their Week 52 visit or discontinued the trial. A formal statistical analysis plan for each study (Induction Study and Maintenance Study) was developed and finalized by the applicant prior to the unblinding of treatment assignment. These plans defined the analysis population, outlined all data handling conventions, and specified all statistical methods to be used for safety and efficacy data analysis.

Demographic and baseline (Week 0) disease characteristics were summarized for Induction and Maintenance by treatment group and overall, using the respective ITT populations.

3.1.1.2 Applicant's Analyses

Approximately 1059 patients were planned to be enrolled into this study from approximately 500 sites worldwide. Enrollment was defined as the point in time at which the patient was assigned a treatment in the Induction Phase. An initial cohort (Cohort 1) of 370 patients was to be randomized in the Induction Phase, based on the sample size requirements for the Induction Phase. Approximately 689 patients were then to be enrolled in Cohort 2. The number of patients to be enrolled in Cohort 2 was determined by the sample size requirements for the Maintenance Phase. The protocol allowed for up to 100 additional patients to be enrolled into Cohort 2 (increasing the total number of study participants to 1159), depending on the observed overall response rate in the combined cohorts, to ensure that at least 501 patients with clinical response at Week 6 to vedolizumab treatment were randomized in the Maintenance Phase. The Data safety Monitoring Board (DSMB) was to monitor the overall response and attrition rate (i.e., patients who were not willing to participate in the Maintenance Phase) to determine if additional enrollment of patients would be required to ensure that the sample size for the Maintenance Study could be achieved.

3.1.1.2.1 Induction Phase

The 6-week Induction Phase contained two cohorts of patients: Cohort 1 patients were randomized and treated with the study drug in a double-blind manner, and Cohort 2 patients were treated with vedolizumab in an open-label manner. Patients with a history of prior TNFα antagonist exposure were permitted to enroll into Cohort 2 if Cohort 1 enrollment had reached the limit of approximately 50% for that subpopulation. Prior to Amendment 5/6, the cohorts in the Induction Phase were enrolled sequentially, i.e., patients were enrolled in Cohort 2 after enrollment in Cohort 1 was complete. The eligibility criteria for both cohorts were identical. In Cohort 1, eligible CD patients who met entry criteria were randomized to the study treatment, in a double-blind manner, with vedolizumab 300 mg or placebo in a 3:2 ratio. The randomization was to be stratified by the presence or absence of two factors, which were considered markers of disease severity:

- Concomitant use of oral corticosteroids

- Previous exposure to TNFα antagonists or concomitant immunomodulator (6-MP, azathioprine, or methotrexate) use

Randomized patients were treated in a double-blind manner with infusions of study drug at Weeks 0 and 2. These patients comprised the population evaluated for efficacy and were referred to as the Induction Study ITT Population. Randomization occurred via an IVRS. Treatment assignment was obtained from the system by the (unblinded) site pharmacist, who prepared study drug and provided it to the site personnel (who remained blinded) in masked infusion bags.

The second cohort of patients was enrolled into the Induction Phase to ensure that the sample size of Induction Phase responders randomized into the Maintenance Study would provide sufficient power for the Maintenance Study primary efficacy analysis. These patients did not contribute to the efficacy analyses performed for the Induction Study. All patients in Cohort 2 were treated with open-label vedolizumab 300 mg, administered at Week 0 and Week 2. Patients in both cohorts were to be assessed for treatment response at Week 6.

Disease activity for entry into this study and for efficacy assessments throughout the study was measured by the CDAI.

The main inclusion criteria were:

1. Diagnosis of CD established at least three months prior to enrollment by clinical and endoscopic evidence and corroborated by a histopathology report. Cases of CD established at least six months prior to enrollment for which a histopathology report was not available and would be considered based on the weight of the evidence supporting the diagnosis and excluding other potential diagnoses, and must have been discussed with the applicant on a case-by-case basis prior to enrollment. (Prior to Amendment 5/6, the diagnosis of CD was to have been established for at least six months prior to enrollment.)

2. Moderately to severely active CD as determined by a CDAI score of 220 to 450 (prior to Amendment 5/6, the CDAI maximum for enrollment was 480) within seven days prior to the first dose of study drug and one of the following:

    - CRP level > 2.87 mg/L during the Screening period *OR*
    - Ileocolonoscopy with photographic documentation of a minimum of 3 nonanastomotic ulcerations (each > 0.5 cm in diameter) or 10 aphthous ulcerations (involving a minimum of 10 contiguous cm of intestine) consistent with CD, within four months prior to randomization *OR*
    - Fecal calprotectin > 250 mcg/g stool during the Screening period in conjunction with computed tomography (CT) enterography, magnetic resonance (MR) enterography, contrast-enhanced small bowel radiography, or wireless capsule endoscopy revealing Crohn's ulcerations (aphthae not sufficient), within four months prior to screening. (Patients with evidence of fixed stenosis or small bowel stenosis with prestenotic dilation should not be included.)

3. CD involvement of the ileum and/or colon, at a minimum

4. Demonstrated, over the previous 5-year period, an inadequate response to, loss of response to, or intolerance of at least one of the following agents as defined below:

- Immunomodulators

  - o Signs and symptoms of persistently active disease despite a history of at least one 8-week regimen of oral azathioprine (≥ 1.5 mg/kg) or 6-MP (≥ 0.75 mg/kg) OR
  - o Signs and symptoms of persistently active disease despite a history of at least one 8-week regimen of methotrexate (≥ 12.5 mg/week) OR
  - o History of intolerance of at least one immunomodulator (including, but not limited to nausea/vomiting, abdominal pain, pancreatitis, liver function test abnormalities, lymphopenia, TPMT genetic mutation, infection)

- TNFα antagonists

  - o Signs and symptoms of persistently active disease despite a history of at least one 4-week induction regimen of one of the following agents

    - Infliximab: 5 mg/kg IV, two doses at least two weeks apart
    - Adalimumab: one 80 mg SC dose followed by one 40 mg dose at least two weeks apart
    - Certolizumab pegol: 400 mg SC, two doses at least two weeks apart or

  - o Recurrence of symptoms during scheduled maintenance dosing following prior clinical benefit (discontinuation despite clinical benefit does not qualify) or
  - o History of intolerance of at least one TNFα antagonist (including, but not limited to infusion-related reaction, demyelination, congestive heart failure, infection)

ONLY APPLICABLE TO PATIENTS OUTSIDE THE US (who may have been enrolled on the basis of corticosteroid treatment history):

- Corticosteroids
    - Signs and symptoms of persistently active disease despite a history of at least one 4-week induction regimen that included a dose equivalent to prednisone 30 mg daily orally for 2 weeks or IV for 1 week, ***OR***
    - Two failed attempts to taper corticosteroids to below a dose equivalent to prednisone 10 mg daily orally on 2 separate occasions, ***OR***
    - History of intolerance of corticosteroids (including, but not limited to, Cushing's syndrome, osteopenia/osteoporosis, hyperglycemia, insomnia, and infection).

The primary objectives were:

- To determine the effect of vedolizumab induction treatment on clinical remission at 6 weeks Clinical remission was defined as a CDAI score ≤ 150 points.
- To determine the effect of vedolizumab induction treatment on enhanced clinical response at 6 weeks Enhanced clinical response was defined as a ≥ 100-point decrease in the CDAI score from baseline (Week 0).

The secondary objective was:

- To determine the effect of vedolizumab induction treatment on serum C-reactive protein (CRP) levels at 6 weeks in patients with elevated CRP levels at baseline.

The exploratory objectives were:

- To analyze key endpoints in the subgroup of patients with previous exposure to TNFα antagonist therapy and in the subgroup of patients defined as having failed TNFα antagonist therapy.
- To analyze key endpoints in the subgroups of patients on concomitant therapies.
- To correlate CDAI scores with Harvey-Bradshaw

Index (HBI) scores

The primary endpoints were:

- Proportion of patients in clinical remission at Week 6.
- Proportion of patients with enhanced clinical response at Week 6.

The secondary endpoint was:

- Change in serum CRP levels at Week 6.

The exploratory endpoints were:

- Key endpoints in the subgroup of patients with previous exposure to TNFα antagonist therapy and in the subgroup of patients defined as having failed TNFα antagonist therapy.
- Key endpoints in the subgroups of patients on concomitant therapies.

There were five induction populations in this study: the Intent-to-Treat (ITT) Population, the Modified ITT Population, the Per-Protocol Population, the Safety Population, and the Completers (Observed Case) population.

For the induction efficacy analyses, the ITT Population consisted of all randomized patients in Cohort 1 who received any amount of blinded study drug. This population was used for the primary efficacy analysis and all proportional-based endpoints, such as remission and enhanced response. Patients in this population were analyzed according to the treatment they were randomized to receive, regardless of any errors in study drug dosing.

The Modified ITT Population for the induction analyses consisted of all randomized patients in Cohort 1 who received any amount of blinded study drug and had a baseline (Week 0) and at least one measurement post-randomization for the endpoint under consideration (e.g., CDAI score).

This population was used for change from baseline (Week 0) analyses, such as analyses of the CDAI score. Patients in this population were analyzed according to the treatment they were randomized to receive, regardless of any errors of dosing.

Patients were included in the induction Per-Protocol Population if they met the following criteria according to the specified hierarchy:

- Confirmed diagnosis of CD of at least six months' duration and with an enrolling CDAI score between 210 and 490 (inclusive) at baseline (220 to 480 prior to Amendment 4)
- Received the correct study medication as assigned
- Met one or more of the following criteria for treatment failure prior to Day 43:

    - Failed, as assessed by the investigator
    - Received any non-study drug due to lack of efficacy
    - Had surgery due to lack of efficacy
    - Had a drug-related adverse event (AE) leading to discontinuation
- Received both doses of study drug, as assigned
- Did not receive concomitant corticosteroids or other potentially effective medications (except as permitted per protocol) for an unrelated comorbid condition (e.g., prednisone for idiopathic thrombocytopenic purpura)
- Had a valid Day 43 assessment (the window for eDiary between 36 and 56 days, inclusive, and a hematocrit measurement between 29 and 56 days)

The Induction Phase Safety Population was defined as all patients, in both Cohort 1 and Cohort 2, who received any amount of study drug in the Induction Phase (Weeks 0-6), according to the actual study drug received. The Induction Phase Safety Population was used for all safety analyses at Week 6.

Additionally, selected safety tables were provided for patients in Cohort 1 because this represented the double-blind safety experience during the Induction Phase, and parallel with the induction efficacy analyses.

The Induction Study Completers (Observed Case) Population was defined as all randomized patients in Cohort 1 who received any amount of blinded study drug, who had a baseline (Week 0) and Week 6 assessment for the endpoint under consideration (e.g., CDAI score).

3.1.1.2.1.1 Pre-specified Analyses

The primary induction study efficacy assessments were on for differences in the proportion of patients with vedolizumab treatment regimen versus placebo who were in clinical remission or had achieved enhanced clinical response at Week 6. The primary comparison of the Induction Phase was tested using the Cochran-Mantel-Haenszel (CMH) chi-square test at a two-sided 5% significance level, with stratification according to the Induction Phase randomization stratification factors (concomitant use of oral corticosteroids and previous exposure to TNFα antagonists and/or concomitant immunomodulator [6-MP, azathioprine, or methotrexate] use). The CMH chi-square p-value and the risk difference along with its 95% two-sided confidence interval (CI) were provided.

The Hochberg method was applied to control the overall Type I error rate at a two-sided 5% significance level for the multiple comparisons of the primary endpoints. If both p-values were ≤ 0.05, both primary endpoints were to be declared significant. If one of the p-values for the primary endpoints was > 0.05, the other p-value was to be tested at the 0.025 level and declared significant only if the p-value was ≤ 0.025. If neither primary endpoint was declared significant, no further testing on the secondary endpoint was to be conducted. If at least one of the primary endpoints was significant, the sequential procedure was to be used to test the secondary endpoint for significance.

In addition to the primary endpoint assessments, there was one secondary assessment of clinical efficacy (mean CRP levels), which compared the treatment difference between vedolizumab and placebo. The applicant stated that to further maintain the overall Type I error rate at 5%, the secondary endpoint was to be tested only if at least one of the primary comparisons was significant.

Changes in CRP level were assessed at Week 6. The change from baseline in CRP level was presented by treatment arm.

Changes in the IBDQ, SF-36, and EQ-5D scores were assessed at Week 6. The mean change from baseline in IBDQ, SF-36, and EQ-5D scores were presented by treatment arm along with 95% two-sided CIs for the differences in mean changes from baseline based on an analysis of covariance (ANCOVA) model.

A total sample size of 1059 was planned for the Induction Phase. An initial cohort of 370 patients (Cohort 1) was to be randomized in a 3:2 ratio to receive vedolizumab (n = 222) or placebo (n = 148) in a double-blind manner. Following the randomization of this first cohort of 370 patients, 689 patients were to be enrolled into Cohort 2, and were to receive open-label vedolizumab induction dosing. Cohort 2 was necessary to provide sufficient power for the Maintenance analyses and was not included in the formal efficacy analysis of the Induction Study.

Power estimates for the primary and key secondary efficacy endpoints for the Induction Phase are based on a total sample size of 370 patients at a 5% significance level and were provided in the table below. The response rate assumptions on which the sample size was based were derived from the phase 2 data.

Table 1 Power Estimated for Primary and Key Secondary Efficacy Analyses in the Induction Phase (Cohort 1) Study C13007

Objective	Endpoint at Week 6 vedolizumab vs. placebo	Assumed Response Rates	Sample Size per Group [b]	Power
Primary[a]	Clinical Remission	Placebo = 21% Vedolizumab = 37%	Placebo = 148 Vedolizumab = 222	91%
Primary	Enhanced Clinical Response	Placebo = 31% Vedolizumab = 46%	Placebo = 148 Vedolizumab = 222	82%
Secondary	Mean Serum CRP levels (mg/L)	Placebo = 21% Vedolizumab = 12% (Std Dev = 31)	Placebo = 148 Vedolizumab = 222	77%

Copied from Table 4, page 80 CSR..

3.1.1.2.1.2 Patient Disposition

A total of 1920 patients were screened for enrollment in the study (data obtained from IVRS). Of these, 804 patients failed screening due to the following reasons: did not meet enrollment criteria (628 patients); withdrew consent (43 patients); applicant discretion (7 patients); serious adverse events (SAEs) (21 patients); and other reasons (105 patients). Thus, 1116 patients were enrolled in the study.

Among the 1116 patients, 368 were enrolled into Cohort 1 and 748 were enrolled into Cohort 2.

The detailed patient disposition is given in table below.

Table 2 Patient Dispositions- Induction Phase
Study C13007

	Induction Cohort 1 ITT Population[a]		Induction Cohort 2[b] Open-label	VDZ Combined	Total
	PLA N = 148	VDZ N = 220	VDZ N = 747	N = 967	N = 1115
Randomized/ assigned	148	220	748[c]	968	1116
Safety Population[d]	148 (100)	220 (100)	747 (100)	967 (100)	1115 (100)
ITT Population[e]	148 (100)	220 (100)		220 (23)	368 (33)
Per-Protocol Population[f]	141 (95)	205 (93)		205 (21)	346 (31)
Completed Induction Phase[g]	137 (93)	199 (90)	674 (90)	873 (90)	1010 (91)
Discontinued (reason)	11 (7)	21 (10)	73 (10)	94 (10)	105 (9)
Adverse event[h]	7 (5)	9 (4)	24 (3)	33 (3)	40 (4)
Protocol violation(s)	0	0	1 (< 1)	1 (< 1)	1 (< 1)
Lack of efficacy	1 (< 1)	3 (1)	28 (4)	31 (3)	32 (3)
Study terminated by sponsor	0	0	0	0	0
Withdrawal of consent	3 (2)	9 (4)	15 (2)	24 (2)	27 (2)
Lost to follow-up	0	0	3 (< 1)	3 (< 1)	3 (< 1)
Other	0	0	2 (< 1)	2 (< 1)	2 (< 1)

Source: Table 14.1.1.2BP.

Abbreviations: ITT = Intent-to-Treat; PLA = placebo; VDZ = vedolizumab.

a All patients enrolled in Cohort 1 who were randomized to blinded induction treatment with vedolizumab or placebo.

b All patients enrolled in Cohort 2 who received open-label vedolizumab induction treatment.

c One patient enrolled in Cohort 2 withdrew from the study prior to dosing and is excluded from all analyses.

d Safety Population consists of all patients who received any amount of study drug during the Induction Phase based on what they actually received.

e ITT Population consists of all randomized patients who received any amount of blinded study drug during the Induction Phase based on what they were randomized to receive.

f Per-Protocol Population consists of all randomized patients who met prespecified criteria (Section 10.2.2-I).

g Defined as completed dosing at Weeks 0 and 2 and completed the predose assessments at Week 6.

h One additional ITT placebo patient is presented in Table 33 as discontinuing due to an AE; this patient is not counted here as the AE that led to discontinuation was not treatment emergent.

Copied from Table 6, page 115 CSR.

Inclusion and exclusion criteria not met at the Induction Phase entry are summarized for the Induction Study ITT population in the table below.

Table 3 Inclusion and Exclusion Criteria Not Met at Induction Phase Entry
Induction Study ITT Population
Study C13007

Type of Unmet Criteria[a], n (%)	PLA N = 148	VDZ N = 220	Total N = 368
Patients with at least 1 unmet entry criterion	8	15	23
Inclusion criteria			
CDAI score of 220 to 450 (prior to Amendment 5/6: maximum was 480) within 7 days prior to first dose of study drug and either a) CRP > 2.87 mg/L during screening, b) at least 3 non-anastomotic ulcerations or 10 aphthous ulcerations (per Amendment 5/6) within 4 months prior to randomization, or c) fecal calprotectin > 250 mcg/g with appropriate imaging during screening (per Amendment 5/6)[b]	3 (2)	4 (2)	7 (2)
CD ≥ 3 months' duration prior to enrollment corroborated with histopathology report or ≥ 6 months' duration if report not available (prior to Amendment 5/6: CD ≥ 6 months' duration)[b]	0	4 (2)	4 (1)
Demonstrated over the previous 5-year period an inadequate or lost response/intolerance of steroids, immunomodulators, and/or TNFα antagonists	1 (< 1)	3 (1)	4 (1)
Documented evidence of colonoscopy within 12 months of enrollment for patients with long-standing disease	0	1 (< 1)	1 (< 1)
Initial steroid dose stable for 4 weeks prior to enrollment, or for the 2 weeks prior to enrollment if tapering	1 (< 1)	0	1 (< 1)
Gastrointestinal exclusion criteria			
C. difficile infection or other intestinal pathogen within 28 days prior to enrollment (prior to Amendment 5/6: *C. difficile* infection within 60 days or other intestinal pathogen within 30 days prior to enrollment)[b]	1 (< 1)	0	1 (< 1)
Infectious disease exclusion criteria			
Missing baseline tuberculin test	0	2 (< 1)	2 (< 1)
General exclusion criteria			
Positive PML subjective symptom checklist prior to first dose of study drug	1 (< 1)	1 (< 1)	2 (< 1)
Hemoglobin level < 8 g/dL during screening	1 (< 1)	1 (< 1)	2 (< 1)
Lymphocyte count < 0.5×10^9/L during screening	0	1 (< 1)	1 (< 1)

Copied from Table 7, page 116-117 CSR.

As seen from the table above, in the Induction Study ITT Population, a total of 23 patients (8 in placebo; 15 in vedolizumab) had at least one unmet entry criterion. The most common deviations were failure to meet the inclusion criterion for a baseline CDAI score of 220 to 450 associated with either a) a CRP level > 2.87 mg/L, a minimum of three nonanastomotic ulcerations, b) 10 aphthous ulcerations consistent with CD, or c) a fecal calprotectin > 250 µg/g with appropriate imaging (3 in placebo; 4 in vedolizumab); CD diagnosis of at least three months confirmed by histology or of at least six months based on other supporting evidence if histology report was not available (0 in placebo; 4 in vedolizumab); and inadequate or lost response/intolerance of steroids, immunomodulators, and/or TNFα antagonists (1 in placebo; 3 in vedolizumab). An additional 63 patients in the open-label vedolizumab group had violations of inclusion/exclusion criteria; which were primarily failures to meet the inclusion criterion for baseline CDAI score. No notable trends were observed for the treatment groups with respect to inclusion/exclusion criteria deviations.

Criteria that led to exclusion from the Induction Study Per-Protocol Population are summarized for the Induction Study ITT Population in the table below.

Table 4 Criteria Leading to Exclusion from the Per-Protocol Population
Induction Study ITT Population
Study C13007

Criterion[a], n (%)	PLA N = 148	VDZ N = 220	Total N = 368
Number of patients excluded from the Per-Protocol Population	7	15	22
Screening and baseline CDAI score < 210 OR > 490 or CD duration < 3 months	3 (2)	3 (1)	6 (2)
Received incorrect study medication as assigned at any study visit	0	0	0
Received < 2 doses of study medication, unless patient met 1 of the criteria for failure	3 (2)	4 (2)	7 (2)
Received concomitant corticosteroids or other potentially effective medications for unrelated comorbid condition	0	0	0
Invalid Day 43 assessment[b]	6 (4)	12 (5)	18 (5)
Patients who had blind broken	0	0	0

Copied from Table 8, page 118 CSR.

As seen from the table above, in the Induction Study ITT Population, a total of 22 patients (7 in placebo; 15 in vedolizumab) met at least one criterion that led to exclusion from the Per-Protocol Population. An invalid Day 43 (Week 6) assessment was the most common reason for exclusion in each treatment group (6 in placebo; 12 in vedolizumab). Seven patients (3 in placebo; 4 in vedolizumab) received less than two doses of study medication due to either the patient had elected to withdraw from the study (2 in each treatment group), the occurrence of an AE (2 in vedolizumab), or missing the second dose of study drug (1 in placebo).Three patients in each treatment group with baseline CDAI scores that were either missing (1 patient in each treatment

group) or < 210 (2 patients in each treatment group; baseline CDAI score range: 132 to 208) were excluded from the Per-Protocol Population.

3.1.1.2.1.3 Treatment Group Comparability

Baseline demographic characteristics of the Induction Phase Safety Population are summarized by treatment group in Appendix Table 1.

As seen from Appendix Table 1, overall, baseline demographic characteristics were similar between the treatment groups in the Induction Study ITT Population. In the overall population, there was higher proportion of female patients than male patients (53% vs. 47%). Most patients were White (89%) and non-Hispanic (96%). The median age was 34.0 years; most patients were < 35 years of age (52%) while only a few patients were ≥ 65 years (2%). The median body weight was 66.2 kg and the median body mass index (BMI) was 22.9 kg/m^2. With respect to geographic distribution, 36% were enrolled at sites in the North America, including 24% from sites in the US, and 64% were enrolled at sites outside of North America, including 23% at Western/Northern European sites, 19% at Central European sites, 14% at sites located in Asia, Australia, and Africa, and 8% at Eastern European sites.

The demographic characteristics of the open-label vedolizumab group were generally similar to those observed in the Induction Study ITT Population, except that the open-label vedolizumab group had more patients enrolling at sites in Western/Northern Europe and fewer patients at sites in Asia/Australia/Africa and Eastern Europe than that was observed for the Induction Study ITT Population.

Appendix Table 2 presents a comparison of selected baseline demographic characteristics of patients randomized to placebo versus patients randomized to vedolizumab in the Induction Study ITT Population.

As seen from Appendix Table 2, no statistically significant differences were noted between the treatment groups for selected baseline demographic characteristics including gender, race, age, body weight, and geographic region.

Baseline (Week 0) CD characteristics of the Induction Phase Safety Population are summarized by treatment group in Appendix Table 3.

As seen from Appendix Table 3, consistent with the study's inclusion criteria, patients with moderately to severely active CD were enrolled, as demonstrated by the baseline disease characteristics of the treatment groups. The mean duration of disease was 9.0 years (median 7.0 years) and the mean baseline disease activity, as assessed by the baseline CDAI score, was 323.6. Baseline CDAI scores were > 330 in 44% of the patients. A majority of the patients had a baseline CRP > 10 mg/L (53%), a baseline fecal calprotectin > 500 μg/g (56%), and disease involvement of both the ileum and colon (55%). A history of prior surgery for CD was reported for 42% of the patients. A majority of the patients had no history of fistulizing disease (63%); 15% of the patients had a draining fistula at baseline. Extraintestinal manifestations of the

disease were present at baseline in 62% of the patients; 82% of the patients had a history of extraintestinal manifestations. Most patients had never smoked or were former smokers (73%).

The baseline disease characteristics of the treatment groups in the Induction Study ITT Population were generally comparable, although the vedolizumab group had greater proportions of patients with CD duration of ≥ 7 years (50%) and with a history of prior surgery for CD (45%) compared to the placebo group (43% and 36%, respectively). The baseline disease characteristics of the open-label vedolizumab group were generally similar to those observed in the Induction Study ITT Population.

The prior use of TNFα antagonists and treatment failure to CD therapies are summarized for the Induction Phase Safety Population in Appendix Table 4.

As noted, information regarding prior use of CD medications, previous treatment failure, and concomitant medications was captured at both screening and baseline (Week 0), and during the study.

Therefore, the numbers of patients in this table and subsequent summaries of baseline and concomitant medication use might vary based on how the data were collected (IVRS vs. eCRF).

As seen from Appendix Table 4, approximately half of the patients in the Induction Study ITT Population (placebo 49%; vedolizumab 50%) reported prior TNFα antagonist use. The proportions of patients who had previously failed TNFα antagonist therapy or were naïve to TNFα antagonist therapy were similar between the treatment groups. In addition, the treatment groups were similar with respect to the number of TNFα antagonist therapies patients had previously failed.

A hierarchical approach was used to categorize treatment failure to TNFα antagonists, immunomodulators, and corticosteroids ("worst treatment failure"). TNFα antagonist failure was prioritized over failure to immunomodulators, which was prioritized over failure of corticosteroids. Within each treatment category, patients were categorized by type of failure to a particular agent, For TNFα antagonists, patients were categorized as having had an inadequate response (persistently active disease despite induction treatment), loss of response (recurrence of symptoms during maintenance treatment following prior clinical benefit), or intolerance (treatment-related toxicity). For immunomodulators and corticosteroids, treatment failure was categorized as either inadequate response (persistently active disease despite a 4-week regimen of corticosteroids or an 8-week regimen of immunomodulators) or intolerance, using similar definitions. As patients may have had more than one definition of treatment failure, only one category was assigned to each patient. Worst treatment failure was assigned using a hierarchical approach, with inadequate response considered worse than loss of response, and loss of response worse than intolerance.

Using this approach, for patients with any prior TNFα antagonist failure, the proportions of patients in the Induction Study ITT Population in each prior failure category were comparable between the treatment groups, with a majority of patients in each treatment group having shown

inadequate response (primary failure: placebo 59%; vedolizumab 53%) or loss of response (secondary failure: placebo 31%; vedolizumab 38%) to prior TNFα antagonist therapy. Similar proportions of patients in each treatment group had previously failed immunomodulators, without TNFα antagonist failure (placebo 34%; vedolizumab 35%); fewer patients had failed corticosteroids alone (placebo 18%; vedolizumab 17%).

Compared to the Induction Study ITT Population, the open-label vedolizumab group had a greater proportion of patients who had prior TNFα antagonist use (68%) and failure (63%), with most patients having shown inadequate response (primary failure: 47%) or loss of response (secondary failure: 40%).

Prior therapy with other CD treatments is presented for the Induction Study ITT Population in the Appendix Table 5.

As seen from Appendix Table 5, exposure to systemic corticosteroids was reported by 92% of the patients and exposure to immunomodulators was reported by 78% of the patients. Exposure to TNFα antagonists was reported by 52% of the patients.

Baseline CD therapy, as recorded in the IVRS, was summarized for the Induction Phase Safety Population in Appendix Table 6.

As seen from Appendix Table 6, in the Induction Study ITT Population, CD therapy use at baseline was similar between the treatment groups. Corticosteroid use was reported by 48% of the patients in each treatment group; 30% of patients were treated with corticosteroids alone. Approximately one-third of the patients in each treatment group reported immunomodulator use at baseline; 17% of patients were treated with immunomodulators alone.

Baseline CD therapy use in the open-label vedolizumab group was generally similar to that was observed in the ITT Population.

Appendix Table 7 presents a comparison of selected baseline CD characteristics and medication use of patients randomized to placebo versus patients randomized to vedolizumab in the Induction Study ITT Population.

As seen from Appendix Table 7, no statistically significant differences were noted between the treatment groups for selected baseline CD characteristics, including mean duration of CD, mean disease activity, corticosteroid use at randomization, immunomodulator use at randomization, prior TNFα antagonist use, and prior failure to TNFα antagonist therapy.

3.1.1.2.1.4 Analysis Population

Table below summarizes the analysis populations within the Induction Study ITT Population.

Table 5 Summary of Analysis Proportions for Induction Phase – Cohort 1 Study C13007

Data Set, n (%)	PLA N = 148	VDZ N = 220
Randomized patients	148	220
Safety Population[a]	148 (100)	220 (100)
ITT Population[b]	148 (100)	220 (100)
Modified ITT Population[c]	143 (97)	214 (97)
Per-Protocol Population[d]	141 (95)	205 (93)
Completers (Observed Case) Population[e]	136 (92)	200 (91)

Copied from Table 18, page 133 CSR.

3.1.1.2.1.5 Applicant's Analyses of the Primary Efficacy Endpoints

The primary endpoints for the Induction Study were the proportions of patients who achieved clinical remission at Week 6 and the proportions of patients who achieved an enhanced clinical response at Week 6 in the Induction Study ITT Population.

Clinical remission is defined as CDAI score ≤ 150 points. Enhanced clinical response is defined as a ≥ 100 point reduction from baseline in CDAI score.

The CMH chi-square test was performed with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to TNFα antagonists and/or concomitant immunomodulator use (yes/no).

The results from the analysis of primary efficacy endpoints of clinical remission and enhanced clinical response at Week 6 for Induction Study ITT Population are given below.

Table 6 Primary Efficacy Endpoints of Clinical Remission and Enhanced Clinical Response at Week 6 – Induction Study ITT Population Study C13007

	Clinical Remission[a]		Enhanced Clinical Response[b]	
	PLA N = 148	VDZ N = 220	PLA N = 148	VDZ N = 220
Number (%) achieving endpoint	10 (6.8)	32 (14.5)	38 (25.7)	69 (31.4)
95% CI	(2.7, 10.8)	(9.9, 19.2)	(18.6, 32.7)	(25.2, 37.5)
Difference from placebo[c]		7.8		5.7
95% CI for difference from placebo		(1.2, 14.3)		(-3.6, 15.0)
P-value for difference from placebo[d]		0.0206		0.2322
Relative risk[e]		2.1		1.2
95% CI for relative risk		(1.1, 4.2)		(0.9, 1.7)

Copied from Table 19, page 135 CSR.

As seen from the table above, in the Induction Study ITT Population, a statistically significant greater proportion of vedolizumab-treated patients achieved clinical remission at Week 6 compared with patients who received placebo. The treatment difference between vedolizumab and placebo was 7.8% (95% CI 1.2, 14.3; p = 0.0206).

Although a trend in favor of vedolizumab was observed for the other primary endpoint of enhanced clinical response at Week 6, the difference between the vedolizumab and placebo groups was not statistically significant. The treatment difference between vedolizumab and placebo was 5.7% (95% CI -3.6, 15.0; p = 0.2322),

The pre-specified Hochberg method was applied to control the overall Type I error rate at a 5% significance level for the multiple comparisons of the primary endpoints. Since the p-value for the endpoint of enhanced clinical response at Week 6 was > 0.05, the p-value for the endpoint of clinical remission at Week 6 was tested at the 0.025 level of significance. As the p-value for clinical remission at Week 6 was < 0.025 (p = 0.0206), the study was considered to have met the primary endpoint of clinical remission at Week 6.

The proportions of patients who achieved clinical remission at Week 6 are presented in Appendix Table 8 for the Induction Study Per-Protocol Population and in Appendix Table 9 for the Induction Study Completers (Observed Case) Population.

As seen from Appendix Tables 8 and 9, the results of these analyses were similar to those observed in the Induction Study ITT Population, with statistically significantly greater proportions of vedolizumab-treated patients achieving clinical remission at Week 6 compared to patients treated with placebo.

The proportions of patients who achieved enhanced clinical response at Week 6 are presented in Appendix Table 10 for the Induction Study Per-Protocol Population and in Appendix Table 11 for the Induction Study Completers (Observed Case) Population.

A seen from Appendix Table 10 and 11, the results of these analyses were similar to those observed in the Induction Study ITT Population, with a trend in favor of vedolizumab, but no statistically significant difference compared to placebo.

3.1.1.2.1.5.1 Subgroup Analyses

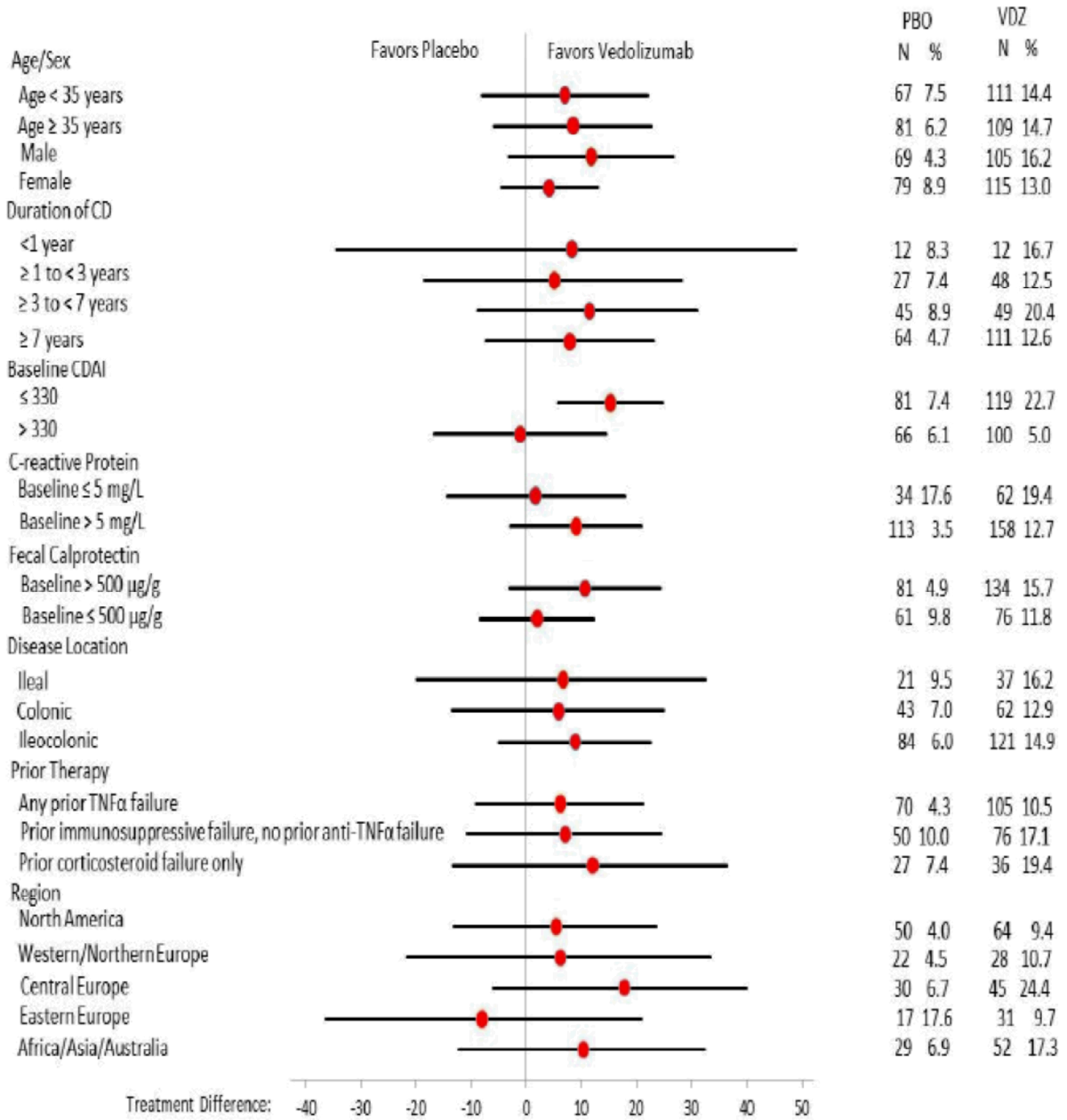
Subgroup analyses for clinical response at Week 6 in the Induction Study ITT population were provided based on: age (age < 35, age ≥ 35 years), gender, race, duration from UC diagnosis to first dose, baseline CDAI (≤330 , >330), baseline C-reactive Protein (≤5 mg/L, > 5 mg/L), geographic region, baseline fecal calprotectin (≤ 500 µg/g, > 500 µg/g), prior therapy, and disease localization.

2.1.1.2.1.5.1.1 Clinical Remission at Week 6

The figure below summarizes the risk differences (percentages) between the treatment groups for the primary endpoint of clinical remission at Week 6 in patient subgroups according to the

aforementioned demographic characteristics and measure of disease activity in the Induction Study ITT Population.

Figure 1 Treatment Difference in Percentage Points for Clinical Remission at Week 6 with the 95% Confidence Interval by Baseline Subgroups- Induction ITT Population
Study C13007



Copied from Figure 7-17, Applicant's Advisory Committee Briefing Document.

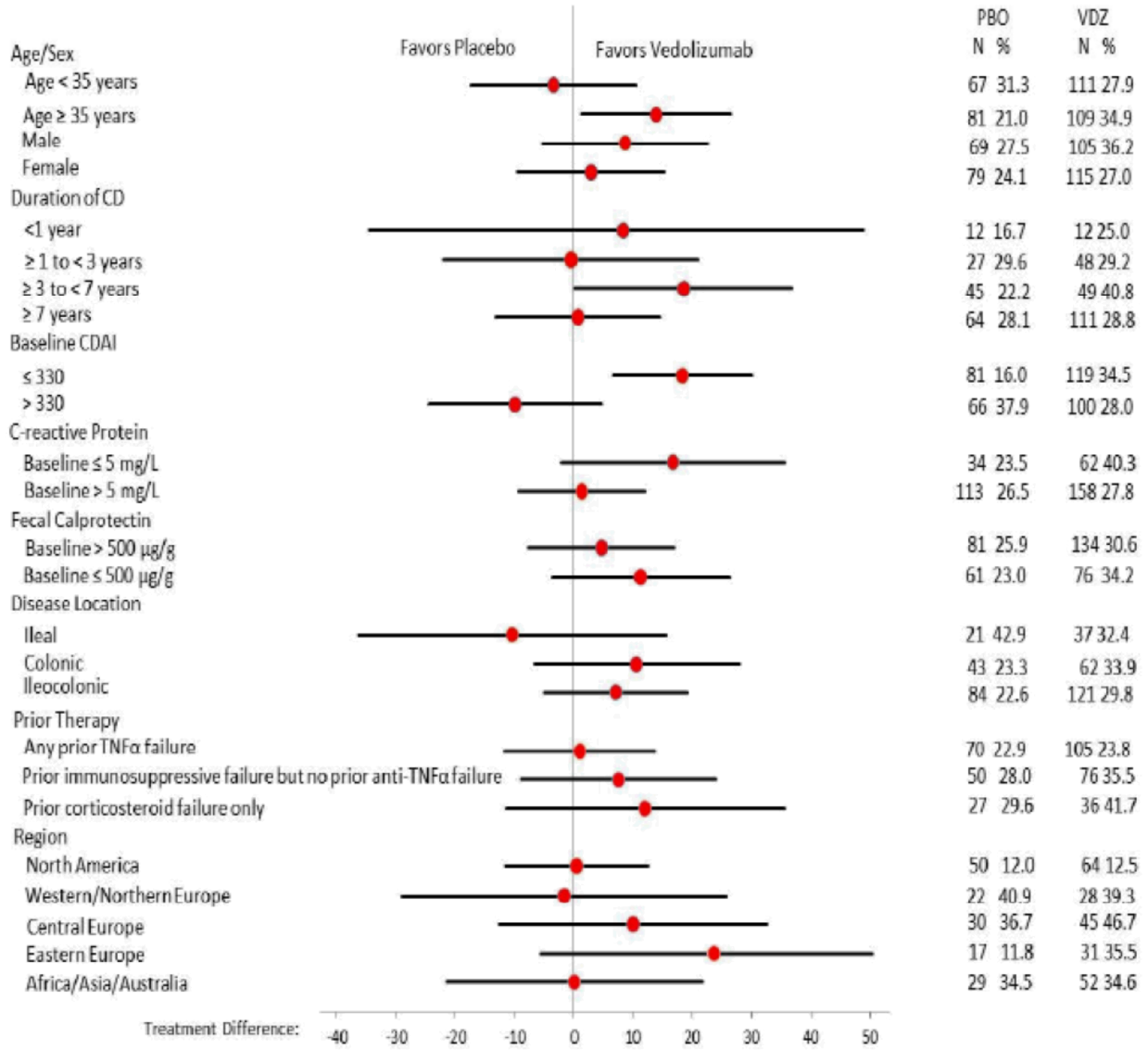
As seen from the figure above, for clinical remission at Week 6, the risk difference between treatment groups favored vedolizumab only in the subgroup of patients who had baseline CDAI

≤ 330 point. There was a notable greater variability and the 95% CIs for the treatment differences often included zero for most of subgroups.

2.1.1.2.1.5.1.2 Enhanced Clinical Response at Week 6

The figure below summarizes the risk differences (percentages) between the treatment groups for the primary endpoint of enhanced clinical response at Week 6 in patient subgroups according to the aforementioned demographic characteristics and measure of disease activity in the Induction Study ITT Population.

Figure 2 Treatment Difference in Percentage Points for Enhanced Clinical Response (CDAI-100 Response) at Week 6 with 95% Confidence Interval by Baseline Subgroups-Induction ITT Population
Study C13007



Copied from Applicant's Advisory Committee Briefing Document.

As seen from the figure above, for the enhanced clinical response (CDAI-100 response) at Week 6, the risk difference between treatment groups favored vedolizumab in the subgroups of patients aged ≥ 35 and of patients who had baseline CDAI ≤ 330 point. There was notable greater variability and the 95% CIs for the differences from placebo often included zero for most of subgroups.

3.1.1.2.1.6 Applicant's Analysis of the Secondary Efficacy Endpoint

The secondary efficacy endpoint was the change from baseline in serum CRP levels at Week 6.

Changes from Baseline in CRP at Week 6 in the Induction Study ITT Population are presented in given in the table below.

Table 7 Changes from Baseline in CRP at Week 6
Induction Study ITT Population
Study C13007

CRP Level (mg/L)	PLA N = 148	VDZ N = 220
Week 6[a]		
n	147	220
Baseline[b] mean (Std Dev)	23.6 (27.85)	24.1 (27.23)
Week 6 mean (Std Dev)	19.9 (30.05)	21.1 (26.92)
Change from baseline mean (Std Dev)	-3.6 (30.04)	-2.9 (16.28)
Change from baseline median	-0.5	-0.9
10th and 90th percentile	(-27.6, 12.1)	(-20.6, 10.3)
Wilcoxon[c] P-value		0.9288

Source: Table 14.3.1.6A.

CRP = C-reactive protein; ITT = intent-to-treat; PLA = placebo; Std Dev = standard deviation; VDZ = vedolizumab.

a If CRP was missing at Week 6, the last observation carried forward (LOCF) was used to carry out imputation.

b Baseline CRP was derived as the CRP value collected on Day 1 prior to dose; if missing, the screening CRP value was used.

c Wilcoxon Rank Sum test on the CRP change from baseline values (two-sided).

Copied from Table 20, page 139 CSR.

As seen from the table above, among patients in the Induction Study ITT Population, no treatment difference was observed for changes from baseline in CRP. The median change from baseline at Week 6 in CRP was -0.5 mg/L in the placebo group and -0.9 mg/L in the vedolizumab group.

3.1.1.2.1.7 Applicant's Analyses of the Exploratory Endpoints – Induction Phase

3.1.1.2.1.7.1 Key Induction Endpoints in Patients by Prior TNFα Antagonist Use

The proportions of patients in the Induction Study ITT Population without prior TNFα antagonist exposure (TNFα antagonist naïve patients) and those with prior TNFα antagonist failure who achieved clinical remission at Week 6 are summarized by treatment group in the table below.

Table 8 Clinical Remission at Week 6 by Prior TNFα Antagonist Use or Failure
Induction Study ITT Population
Study C13007

	Patients Without Prior TNFα Antagonist Use (Naïve)[a]		Patients with Prior TNFα Antagonist Failure[b]	
	PLA N = 76	VDZ N = 109	PLA N = 70	VDZ N = 105
Number (%) achieving clinical remission[c]	7 (9.2)	19 (17.4)	3 (4.3)	11 (10.5)
95% CI	(2.7, 15.7)	(10.3, 24.6)	(0.9, 12.0)	(5.3, 18.0)
Difference from placebo[d]		8.2		6.2
95% CI for difference from placebo		(-1.4, 17.9)		(-9.1, 21.3)

Copied from Table 21, Page 140 CSR.

As seen from Table above, a trend was observed in the TNFα antagonist naïve patients, with a greater proportion of vedolizumab-treated patients achieving clinical remission at Week 6 compared with patients who received placebo. The treatment difference from placebo was 8.2% (95% CI: -1.4, 17.9).

Similar results were also observed for patients who had previously failed TNFα antagonist therapy, with a greater proportion of vedolizumab-treated patients achieving clinical remission at Week 6 compared with patients who received placebo The treatment difference from placebo was 6.2% (95% CI: -9.1, 21.3).

3.1.1.2.2 Maintenance Phase

The Maintenance Phase included three groups of patients who were to be assigned to treatment groups based on their Induction Phase treatment assignment and response to the study therapy. Vedolizumab-treated patients from both Cohort 1 and Cohort 2 who demonstrated a clinical response according to protocol-specified criteria, as assessed by the investigator, were to be randomized, in a double-blind manner, in a 1:1:1 ratio to treatment with vedolizumab administered every 4 weeks (Q4W), vedolizumab administered every 8 weeks (Q8W), or placebo. Randomization was to be stratified by three factors:

- Enrollment in Cohort 1 or Cohort 2 in the Induction Phase;
- Concomitant use of oral corticosteroids;
- Previous exposure to TNFα antagonists or concomitant immunomodulator use.

As in the Induction Phase, the unblinded study pharmacist obtained the Maintenance Phase treatment assignment based on information provided by the investigator; the investigator remained blinded to the Induction Phase treatment (for those patients in Cohort 1 who had received double-blind treatment) and there was no interruption of treatment between the two phases. These patients comprised the Maintenance Study ITT Population, the primary efficacy population.

Vedolizumab-treated patients who did not demonstrate responses at Week 6 of the Induction Phase were to continue treatment with open-label vedolizumab, administered Q4W. Patients who had been treated with double-blind placebo in the Induction Study were to continue on double-blind placebo during the Maintenance Phase, regardless of treatment response during the induction phase. The Maintenance Phase began at Week 6, included study drug dosing at Week 6 and Q4W or Q8W thereafter, and concluded with Week 52 assessments.

Beginning at Week 6, patients receiving oral corticosteroids who had achieved a clinical response were to begin a corticosteroid tapering regimen. In addition, at Week 6, patients in Cohort 1 participating in the US who were taking concomitant azathioprine, 6-MP, or methotrexate during the Induction Phase were required to discontinue these medications.

After the Week 52 assessments, patients had been eligible to enroll in Study C13008 (Long-term Safety Study) to receive open-label vedolizumab treatment. Patients who withdrew early (prior to Week 52) due to sustained nonresponse, disease worsening, or the need for rescue medications had been eligible to enroll in Study C13008. Patients who did not enroll into Study C13008 were instructed to complete a final on-study safety assessment at Week 66 (or Final Safety visit 16 weeks after the last dose) in the Maintenance Phase of Study C13007. In addition, after the end of the study, all patients who did not enroll in Study C13008 were instructed to participate in a follow-up period in which they were contacted by telephone every six months for two years.

The follow-up questionnaire administered at each time point collected information on events such as infections resulting in hospitalization (at six months only), pregnancy, colorectal dysplasia, cancer, IBD-related surgeries, and the development of progressive multifocal leukoencephalopathy (PML).

The primary objective was:

- To determine the effect of vedolizumab maintenance treatment on clinical remission at 52 weeks.

The secondary objectives were:

- To determine the effect of vedolizumab maintenance treatment on enhanced clinical response at 52 weeks;
- To determine the effect of vedolizumab maintenance treatment on corticosteroid-free remission at 52 weeks;

- To determine the effect of vedolizumab maintenance treatment on durability of clinical remission. Durable clinical remission was defined as clinical remission at ≥80% of study visits including final visit (Week 52).

The exploratory objectives were:

- To examine the effect of maintenance vedolizumab treatment on clinical response, durability of clinical response, and durability of enhanced clinical response

- To examine the effect of maintenance vedolizumab treatment on

    - Time to disease worsening
    - Closure of draining fistulae
    - Serum CRP level in patients with an elevated CRP level at baseline
    - Extraintestinal manifestations of CD
    - Reduction of oral corticosteroid use
    - The proportion of patients at Week 52 who have corticosteroid-free remission for 90 days
    - The proportion of patients at Week 52 who have corticosteroid-free remission for 180 days
    - Enhanced clinical response and remission by Week 14

- To correlate CD-associated genetic polymorphisms and serum biomarkers with therapeutic response to vedolizumab

- To analyze key endpoints in the subgroup of patients with previous exposure to TNFα antagonist therapy and in the subgroup of patients defined as having failed TNFα antagonist therapy

The primary endpoint was:

- Proportion of patients in clinical remission at Week 52.

The secondary endpoints were:

- Proportion of patients with enhanced clinical response at Week 52;
- Proportion of patients using oral corticosteroids at baseline who have discontinued corticosteroids and are in clinical remission at Week 52;
- Proportion of patients with durable clinical remission.

The exploratory endpoints were:

- Time to disease worsening;
- Closure of draining fistulae;
- Reduction in serum CRP level in patients with an elevated CRP level at baseline;
- Improvements in extraintestinal manifestations of CD;

- Reduction in oral corticosteroid use;
- Proportion of patients at Week 52 who have corticosteroid-free remission for 90 days;
- Proportion of patients at Week 52 who have corticosteroid-free remission for 180 days;
- Protein biomarkers associated with CD activity in serum and stool samples;
- Proportions of patients with enhanced clinical response and remission by Week 14;
- Genomic DNA analyzed for polymorphisms associated with therapeutic response to Vedolizumab;
- Key endpoints in the subgroup of patients with previous exposure to TNFα;
- antagonist therapy and in the subgroup of patients defined as having failed TNFα antagonist therapy;
- Key endpoints in the subgroups of patients on concomitant therapies.

There were eight maintenance populations in this study: the ITT Population, the Modified ITT Population, the Per-Protocol Population, the Safety Population, the Delayed-Response Population, the PK population, the PD population, and the Completers (Observed Case) Population. Of note, the efficacy analyses populations for maintenance were separate for each of the maintenance dose regimens.

For the maintenance efficacy analyses, the ITT Population was defined as all randomized patients who received vedolizumab during the Induction Phase and met the protocol definition of clinical response at Week 6, as assessed by the investigator, were randomized, and received any amount of double-blind study drug in the Maintenance Phase.

This population was used for the primary efficacy analysis and all proportional-based endpoints, such as remission, response and corticosteroid-free remission. Patients in this population were analyzed according to the treatment they were randomized to receive, regardless of any errors in study drug dosing.

The modified ITT Population for maintenance analyses included all patients randomized as Week 6 responders who received vedolizumab during the Induction Phase, met the protocol definition of clinical response at Week 6, and then received any amount of study drug and had a baseline (Week 0) and at least one post-Week 6 measurement in the Maintenance Phase for the endpoint under consideration.

This population was used for the change from baseline (Week 0) analyses such as analyses of CDAI score. Patients in this population were analyzed according to the treatment they were randomized to receive, regardless of any errors of dosing.

Patients were included in the maintenance Per-Protocol Population if they met the following criteria according to the specified hierarchy:

- Confirmed diagnosis of CD of at least six months' duration with a CDAI between 210 and 490 score (inclusive);
- Received the correct study medication as assigned;
- Did not have treatment assignment unblinded by the investigator;
- Met one or more of the following criteria for treatment failure prior to Week 52:

    - Failed, as assessed by the investigator
    - Received any non-study drug due to lack of efficacy
    - Had surgery due to lack of efficacy
    - Had a drug-related AE leading to discontinuation
- Received 80% of doses of study drug, as assigned;
- Did not receive concomitant corticosteroids or other potentially effective medications (except as permitted per protocol) for an unrelated comorbid condition (e.g., prednisone for idiopathic thrombocytopenic purpura);
- Had a valid Week 52 or ET assessment for CDAI.

Analyses using the Per-Protocol Population were conducted as sensitivity analyses.

The Maintenance Phase Safety Population for the maintenance analyses was defined as all patients who received any amount of the study drug (Weeks 0-66), according to the actual study drug received. The Maintenance Phase Safety Population was used for all safety analyses at Week 52 and at Week 66.

The Maintenance Study Completers (Observed Case) Population was defined as all randomized patients designated as responders through IVRS in the Induction Study, who received any amount of blinded study drug during the Maintenance Study, and who had a baseline (Week 0) and Week 52 assessment for the endpoint under consideration (e.g., CDAI).

The Delayed-Response Population included all vedolizumab-treated patients who did not meet the protocol definition of calculated clinical response (as assessed by the investigator) and were classified as non-responders by IVRS at Week 6. The Delayed-Response Population was used for all analyses conducted at Weeks 10 and 14.

3.1.1.2.2.1 Pre-specified Analyses

The primary efficacy assessment was conducted on the differences in the proportion of patients who were in clinical remission at Week 52 in vedolizumab Q4W versus placebo and vedolizumab Q8W versus placebo groups. For the two comparisons of the primary endpoint of clinical remission at 52 weeks, the Hochberg method was applied to control the overall Type I error rate at a 5% significance level. The applicant stated that if both p-values were ≤ 0.05, both dose regimens were to be declared significant. If one of the p-values for the two dose comparisons was > 0.05, the other p-value was to be tested at the 0.025 level and declared significant only if the p-value was ≤ 0.025. If neither dose was declared significant for the primary endpoint, no further testing was to be conducted. If at least one of the dose regimens was significant, the sequential procedure was to be used to test the secondary endpoints for significance.

For both assessments of the primary endpoint, the CMH chi-square test was used to compare the two treatment groups at the 5% level of significance with stratification according to the maintenance stratification factors (induction treatment cohort assignment, concomitant use of oral corticosteroids, and previous exposure to TNFα antagonists and/or concomitant

immunomodulator [6-MP, azathioprine, or methotrexate] use). The CMH chi-square p-value and the risk difference along with its 95% two-sided CI were provided.

In addition to the primary comparisons, there were three key secondary assessments of clinical efficacy (enhanced clinical response, corticosteroid-free remission, and durability of clinical remission) in maintenance phase, which compared treatment differences through proposed closed testing procedures. To control the overall Type I error rate at 5% for the multiple-dose comparisons in each key secondary endpoint, the Hochberg method was used. To further maintain the overall Type I error rate at 5%, the key secondary assessments were also performed sequentially. The first key secondary endpoint was to be tested only if one or both of the primary comparisons were significant and the next key secondary endpoint was to be tested only if the previous key secondary endpoint was significant for at least one dose. The order of the key secondary objectives was specified in the statistical analysis plan before clinical database lock. The differences in the proportion of patients in enhanced clinical response, in corticosteroid-free remission, and in durable clinical remission by Week 52 were analyzed in the same fashion as the primary endpoint.

Changes in health-related quality to life (HRQOL) over time were assessed using the IBDQ score, Short Form (SF)-36, and EQ-5D questionnaire.

The mean changes from baseline in IBDQ score, SF-36, and EQ-5D were presented by treatment arm along with 95% two-sided CIs for the differences in mean changes from baseline based on an ANCOVA model.

A sample size of 501 was required to power the Maintenance Study primary and secondary efficacy endpoints. Assuming an induction response rate of 55% among patients receiving vedolizumab (either in Cohort 1 or 2), there would be approximately 501 patients on vedolizumab in the Induction Phase who achieved clinical response at Week 6. During induction, the overall response rate and attrition rate for patients who were not willing to participate in the Maintenance Phase were evaluated periodically by the DSMB to assess study assumptions. This monitoring allowed for necessary adjustments to the number of patients enrolled into the second cohort to achieve the target Maintenance Study sample size of approximately 501 patients. A maximum of 100 additional patients may have been enrolled into the second cohort of the Induction Phase to achieve the target Maintenance Study sample size.

At Week 6, all patients who had received vedolizumab induction treatment and had achieved clinical response, as assessed by the investigator were to be randomized 1:1:1 to vedolizumab Q4W, vedolizumab Q8W, or placebo during the Maintenance Phase.

The sample size calculation for the Maintenance Study was based on the number of patients who received vedolizumab (in either Cohort 1 or Cohort 2) in the Induction Phase and achieved clinical response at Week 6. Power estimates based on a total sample size of 501 patients (167 per arm) and two-sided 5% significance level are provided in Table below

Table 9 Power Estimates for the Primary and Key Secondary Efficacy Analysis in the Maintenance Phase Study 13007

Objective	Endpoint at Week 52	Assumed Response Rates	Sample Size per Group [a]	Power
Primary	Clinical remission for vedolizumab vs. placebo	Placebo = 22% vedolizumab = 38%	167	89%
Key Secondary	Enhanced clinical response for vedolizumab vs. placebo	Placebo = 24% vedolizumab = 40%	167	88%
	Corticosteroid-free remission for vedolizumab vs. placebo	Placebo = 11% vedolizumab = 30%	83[b]	86%
	Durable clinical remission for vedolizumab vs. placebo	Placebo = 12% vedolizumab = 24%	167	81%

a Sample sizes included patients from Cohort 1 and Cohort 2. Patients receiving placebo for induction treatment and patients from all treatment arms who were not in clinical response at Week 6 were to be excluded from these analyses.

b It was expected that 50% to 55% (statistical analysis plan) of the 167 patients per group would have been on corticosteroids at baseline (ie, at least 83 patients per group would contribute to this endpoint).

Copied from Table 5, page 81 CSR.

The total maintenance sample size was expected to be 1059 patients, consisting of the following patient subgroups:

- 501 patients who received vedolizumab in the Induction Phase and who achieved clinical response at Week 6. These patients were randomized to placebo or one of the two dose regimens of vedolizumab as described previously.

- 410 patients who received vedolizumab in the Induction Phase but who did not achieve clinical response at Week 6. These patients were not randomized and were to be assigned to vedolizumab Q4W to provide additional safety, efficacy, and exposure data.

- 148 placebo patients from Cohort 1 who were not randomized were to continue to receive placebo during the Maintenance Phase to serve as a control group.

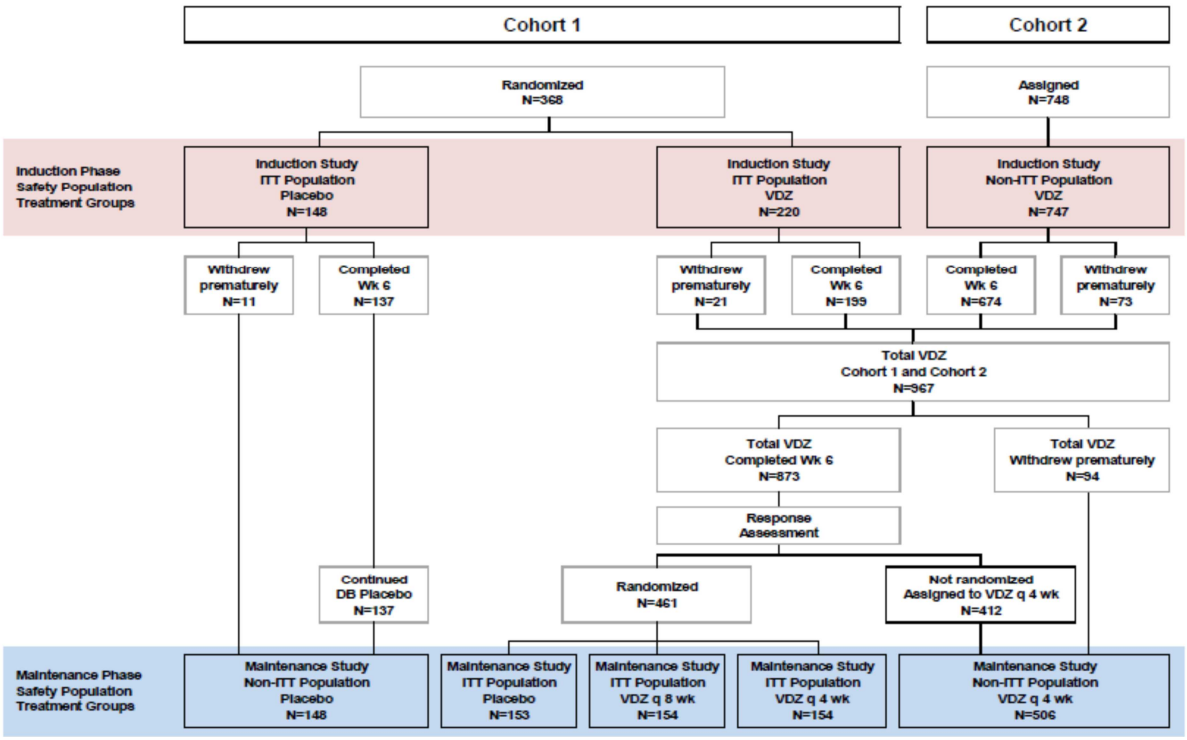
The first group (patients on vedolizumab in the Induction Phase who achieved clinical response at Week 6) was included in the Maintenance Study primary and secondary efficacy analyses. The latter groups (patients on vedolizumab in the Induction Phase who did not achieve clinical response at Week 6, and all placebo patients from Cohort 1) were excluded from the Maintenance Study primary and secondary efficacy analyses but contributed to safety analyses and some exploratory efficacy analyses.

3.1.1.2.2.2 Applicant's Analyses

All patients who completed the Induction Phase entered the Maintenance Phase. Treatment assignments were based on the Induction Phase treatment and the investigator-assessed treatment response.

The figure below summarizes the flow of patients from the Induction Phase into the Maintenance Phase treatment groups and summarizes the composition of the Maintenance Phase Safety Population treatment groups.

Figure 3 Overview of Treatment Groups in Induction Phase and Maintenance Phase Safety Population
Study C13007



Copied from Figure 9, page 195 CSR.

The Maintenance Study ITT Population included vedolizumab-treated patients who had a clinical response at Week 6; at the start of the Maintenance Phase, these patients were randomized to one of the two vedolizumab IV dosing regimens (300 mg Q4W or Q8W) or placebo.

The Maintenance Non-ITT Population included two additional treatment groups: placebo and vedolizumab administered Q4W. The non-ITT placebo group comprised those patients who were randomized to placebo in the Induction Phase; these patients remained on placebo in the Maintenance Phase, per the study design. The non-ITT vedolizumab group comprised those

patients who received vedolizumab in the Induction Phase and were assessed by the investigator as not having achieved clinical response at Week 6; these patients received vedolizumab 300 mg Q4W for the duration of the study. These patients contributed to the safety analyses in the Maintenance Phase, and exploratory efficacy analyses were done for this population.

It should be noted that the safety analyses of patients in the Maintenance Phase included assessments from their participation during the entire study, starting at Week 0 of the Induction Phase. As such, information presented for the Non-ITT Population treatment groups included safety assessments from patients who withdrew from treatment during the Induction Phase. Also of note, all patients in the Maintenance ITT Population randomized to the placebo treatment group in Maintenance Phase were treated with vedolizumab during the Induction Phase. Maintenance safety data for this group include safety assessments made while on vedolizumab treatment during the Induction Phase and on placebo during the Maintenance Phase.

A total of 461 patients responded to vedolizumab therapy during the Induction Phase of the study and were randomized to treatment in the Maintenance Study. Of these, 153 were randomized to placebo, 154 were randomized to vedolizumab Q8W, and 154 were randomized to vedolizumab Q4W. Within each of these treatment groups, all patients received at least one dose of blinded study drug and were included in the Safety and ITT Populations.

3.1.1.2.2.2 Patient Disposition

Patient disposition data for all patients in the Maintenance Phase Safety Population are summarized by treatment groups in the table below.

Table 10 Patient Disposition– Maintenance Phase Safety Population Study C13007

	Maintenance ITT[a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Maintenance Non-ITT		Combined	
	PLA N = 153	VDZ Q8W N = 154	VDZ Q4W N = 154	PLA[b] (from Week 0) N = 148	VDZ Q4W[c] (Week 6 Nonresponders) N = 506	PLA N = 301	VDZ N = 814
Completed induction treatment	153 (100)	154 (100)	154 (100)	137 (93)	412 (81)	290 (96)	720 (88)
Randomized/assigned	153 (100)	154 (100)	154 (100)	148 (100)	507 (100)	301 (100)	815 (100)
Randomized but not dosed	0	0	0	0	1 (< 1)	0	1 (< 1)
Safety Population[d]	153 (100)	154 (100)	154 (100)	148 (100)	506 (100)	301 (100)	814 (100)
ITT Population[e]	153 (100)	154 (100)	154 (100)			153 (51)	308 (38)
Per-Protocol Population[f]	147 (96)	149 (97)	144 (94)			147 (49)	293 (36)
Completed study[g]	64 (42)	73 (47)	82 (53)	42 (28)	163 (32)	106 (35)	318 (39)
Discontinued (reason)	89 (58)	81 (53)	72 (47)	106 (72)	343 (68)	195 (65)	496 (61)
Adverse event	15 (10)	12 (8)	9 (6)	14 (9)	71 (14)	29 (10)	92 (11)
Protocol violation(s)	1 (< 1)	2 (1)	3 (2)	0	5 (< 1)	1 (< 1)	10 (1)
Lack of efficacy	64 (42)	58 (38)	48 (31)	80 (54)	208 (41)	144 (48)	314 (39)
Study terminated by sponsor	0	0	0	0	0	0	0
Withdrawal of consent	7 (5)	6 (4)	9 (6)	10 (7)	48 (9)	17 (6)	63 (8)
Lost to follow-up	1 (< 1)	3 (2)	2 (1)	2 (1)	8 (2)	3 (< 1)	13 (2)
Other	1 (< 1)	0	1 (< 1)	0	3 (< 1)	1 (< 1)	4 (< 1)
Enrolled into C13008	127 (83)	126 (82)	122 (79)	107 (72)	244 (48)	234 (78)	492 (60)

Copied from Table 44, page 196 CSR.

As seen from the table above, in the Maintenance Study ITT Population, slightly greater proportions of vedolizumab treated patients completed Week 52 assessments (47% vedolizumab Q8W; 53% vedolizumab Q4W) compared with placebo patients (42%). Premature discontinuation due to lack of efficacy was highest among placebo patients (42%), with 38% of the vedolizumab Q8W and 31% of the vedolizumab Q4W group. In addition, premature discontinuation due to AEs was highest among placebo patients (10%), followed by vedolizumab Q8W (8%) and vedolizumab Q4W (6%) patients.

The 814 patients in the all vedolizumab combined group consisted of 308 vedolizumab patients from the Maintenance Study ITT Population and 506 vedolizumab patients from the non-ITT Population; one patient withdrew consent prior to dosing during the Induction Phase.

Among these patients, 39% completed Week 52 assessments, with 61% prematurely discontinuing study, primarily due to lack of efficacy (39%) and AEs (11%). Among the 148 patients who received placebo throughout the Induction and Maintenance Phases (non-ITT Placebo Population), 72% prematurely discontinued the study, primarily due to lack of efficacy (54%) and AEs (9%).

Inclusion/exclusion criteria not met prior to study entry (i.e., Induction Phase) are summarized for the Maintenance Phase Safety Population (with ITT and non-ITT populations presented separately) in the table below.

Table 11 Inclusion and Exclusion Criteria Not Met at Induction Phase Entry
Maintenance Study Safety Population
Study C13007

Type of Unmet Criteria[b], n (%)	Maintenance ITT[a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Maintenance Non-ITT		Combined	
	PLA N = 153	VDZ Q8W N = 154	VDZ Q4W N = 154	PLA[c] (from Week 0) N = 148	VDZ Q4W[d] (Week 6 Nonresponders) N = 506	PLA N = 301	VDZ N = 814
Patients with at least one unmet entry criterion	8	12	12	8	45	16	69
Inclusion criteria							
CDAI score of 220 to 450 (prior to Amendment 5/6: 480) within 7 days prior to first dose of study drug and either a) CRP > 2.87 mg/L during screening, b) at least 3 non-anastomotic ulcerations or 10 aphthous ulcerations (per Amendment 5/6) within 4 months prior to randomization, or c) fecal calprotectin > 250 mcg/g with appropriate imaging during screening (per Amendment 5/6)[e]	3 (2)	6 (4)	3 (2)	3 (2)	11 (2)	6 (2)	20 (2)
Inadequate or lost response/intolerance of steroids, immunomodulators, and/or TNFα antagonists	1 (< 1)	2 (1)	4 (3)	1 (< 1)	9 (2)	2 (< 1)	15 (2)
Initial steroid dose stable for 4 weeks prior to enrollment, or for the 2 weeks prior to enrollment if tapering	0	0	2 (1)	1 (< 1)	5 (< 1)	1 (< 1)	7 (< 1)
CD ≥ 3 months' duration prior to enrollment corroborated with histopathology report or ≥ 6 months' duration if report not available (prior to Amendment 5/6: CD ≥ 6 months' duration)[e]	0	2 (1)	1 (< 1)	0	4 (< 1)	0	7 (< 1)
Stable doses of AZA, 6-MP, or methotrexate for 8 weeks prior to enrollment	1 (< 1)	1 (< 1)	0	0	3 (< 1)	1 (< 1)	4 (< 1)
Documented evidence of colonoscopy within 12 months of enrollment for patients with long-standing disease	1 (< 1)	0	0	0	2 (< 1)	1 (< 1)	2 (< 1)
Age 18 to 80	0	0	0	0	1 (< 1)	0	1 (< 1)
Colon cancer screening up-to-date for patients at increased risk	0	0	0	0	1 (< 1)	0	1 (< 1)
Gastrointestinal exclusion criteria							
C. difficile infection or other intestinal pathogen within 28 days prior to enrollment (prior to Amendment 5/6: *C. difficile* infection within 60 days or other intestinal pathogen within 30 days prior to enrollment)[e]	2 (1)	2 (1)	4 (3)	1 (< 1)	3 (< 1)	3 (< 1)	9 (1)
Extensive colonic resections, subtotal or total colectomy	0	0	1 (< 1)	0	0	0	1 (< 1)
History of > 3 small bowel resections or diagnosis of short bowel syndrome	0	0	0	0	1 (< 1)	0	1 (< 1)
Ileostomy, colostomy, or symptomatic stenosis	0	0	0	0	1 (< 1)	0	1 (< 1)
5-ASA or steroid enemas/suppositories within 2 weeks of first dose of study drug	0	0	0	0	1 (< 1)	0	1 (< 1)
Colonic mucosal dysplasia	0	0	0	0	1 (< 1)	0	1 (< 1)
Infectious disease exclusion criteria							
Missing baseline tuberculin test	0	2 (1)	2 (1)	0	2 (< 1)	0	6 (< 1)
Tuberculosis on chest x-ray within 3 months prior to enrollment	1 (< 1)	1 (< 1)	2 (1)	0	0	1 (< 1)	3 (< 1)
General exclusion criteria							
Positive PML subjective symptom checklist prior to first dose of study drug	0	0	1 (< 1)	1 (< 1)	1 (< 1)	1 (< 1)	2 (< 1)
Hemoglobin < 8 g/dL during screening	0	1 (< 1)	0	1 (< 1)	0	1 (< 1)	1 (< 1)
Lymphocyte < 0.5×10^9/L during screening	0	0	0	0	2 (< 1)	0	2 (< 1)
History of malignancy	0	0	0	0	1 (< 1)	0	1 (< 1)

Copied from Table 45, page 199-200 CSR.

As seen from the table above, in the Maintenance Study ITT Population, a total of 32 patients (8 patients in placebo, 12patients in vedolizumab Q8W, and 12 patients in vedolizumab Q4W) failed to meet at least one study entry criterion. The most common deviations across the treatment groups were failure to meet the inclusion criteria for baseline CDAI score of 220 to 450, with either a CRP level > 2.87 mg/L, a minimum of three non-anastomotic ulcerations or 10

aphthous ulcerations consistent with CD, or a fecal calprotectin > 250 µg/g with appropriate imaging (placebo 2%; vedolizumab Q8W 4%; vedolizumab Q4W 2%); inadequate or lost response/intolerance of steroids, immunomodulators, and/or TNFα antagonists (placebo < 1%; vedolizumab Q8W 1%; vedolizumab Q4W 3%); and met the exclusion criterion of *C. difficile* infection or other intestinal pathogen within 28 days of study entry (placebo 1%; vedolizumab Q8W 1%; vedolizumab Q4W 3%).

All of the inclusion or exclusion criteria deviations occurred in ≤ 2% of the all vedolizumab combined group, as well as the non-ITT placebo group.

Criteria that led to exclusion from the Maintenance Study Per-Protocol Population are summarized for the Maintenance Study ITT Population in the table below.

Table 12 Criteria Leading to Exclusion from the Per-Protocol Population Maintenance Study ITT Population Study C13007

Criteria[a], n (%)	PLA N = 153	VDZ Q8W N = 154	VDZ Q4W N = 154
Number of patients excluded from the Per-Protocol Population	6	5	10
Screening and baseline CDAI score < 210 OR > 490 or CD duration < 3 months	1 (< 1)	2 (1)	4 (3)
Received incorrect study medication as assigned at any study visit	0	2 (1)	0
Received < 80% of doses of study medication, unless patient met 1 of the criteria for failure	2 (1)	0	3 (2)
Received concomitant corticosteroids or other potentially effective medications for unrelated comorbid condition	2 (1)	0	0
Invalid Week 52/ET assessment[b]	1 (< 1)	0	3 (2)
Patients had blind broken	0	1 (< 1)	0

Copied from Table 46, page 202 CSR.

As seen from the table above, in the Maintenance Study ITT Population, a total of 21 patients (six placebo; five vedolizumab Q8W; ten vedolizumab Q4W) met at least one criterion that led to exclusion from the Per-Protocol Population. Among the treatment groups, baseline CDAI scores < 210 or > 490 or CD duration < 3 months was the most common reason for exclusion (1 placebo; 2 vedolizumab Q8W; 4 vedolizumab Q4W).

The blind was broken for one patient (Patient C13007-58018-701) in the vedolizumab Q8W group after the patient was hospitalized with an AE of edema peripheral.

3.1.1.2.2.3 Treatment Group Comparability

Baseline (i.e., Week 0) demographic characteristics of the Maintenance Phase Safety Population are summarized by treatment group in Appendix Table 12.

As seen from Appendix Table 12, in the Maintenance Study ITT Population, the demographic characteristics were generally similar among the treatment groups, except for the geographic region. With respect to geographic distribution, greater proportions of patients in the vedolizumab Q8W and Q4W groups were enrolled at sites in North America (38% and 31%, respectively) compared with the placebo group (24%), whereas a greater proportion of placebo patients were enrolled at sites in Western/Northern Europe (35%) compared with the vedolizumab Q8W and Q4W groups (19% and 25%, respectively).

The demographic characteristics of the all vedolizumab combined group were generally consistent with those observed in the Maintenance Study ITT Population, including the greatest proportion of patients enrolling from sites in North America (39%). In addition, the demographic characteristics of the non-ITT vedolizumab patients (Week 6 non-responders) were consistent with those of the Maintenance Study ITT Population (Week 6 responders).

Appendix Table 13 presents the comparison of selected baseline (i.e., Week 0) demographic characteristics among the treatment groups in the Maintenance Study ITT Population.

As seen from Appendix Table 13, no significant differences were noted between the treatment groups for selected baseline demographic characteristics including gender, race, age, and body weight.

Baseline (i.e., Week 0) CD disease characteristics of the Maintenance Phase Safety Population are summarized in Appendix Table 14.

As seen from Appendix Table 14, the baseline disease characteristics were generally similar among the treatment groups in the Maintenance Study ITT Population and indicated the moderately to severely active CD present in this population. Although the majority of patients in each of the treatment groups had baseline CDAI scores ≤ 330, the incidence was highest in the vedolizumab Q8W group (62%), followed by the placebo (56%) and the vedolizumab Q4W (51%) groups. The proportions of patients who had both ileum and colonic involvement was highest in the vedolizumab Q8W (64%) group, followed by the placebo (59%) and the vedolizumab Q4W (47%) groups.

The disease characteristics at baseline for the all vedolizumab combined group and the non-ITT placebo group were generally comparable to those of the Maintenance Study ITT Population, with the exception of higher mean baseline values for CRP. Disease characteristics of the non-ITT vedolizumab patients (Week 6 non-responders) were consistent with greater disease severity including longer disease duration and history of prior CD surgery, and greater disease activity with increased CRP and an increased proportion of patients who had previously failed TNFα antagonist therapy, when compared with the Maintenance Study ITT Population.

Treatment failure to CD therapies is summarized by treatment group for the Maintenance Phase Safety Population and presented in Appendix Table 15.

Information regarding prior use of CD medications, previous treatment failure, and concomitant medications was captured at both screening and Week 0, and during the study. Therefore, the numbers of patients in this table and subsequent summaries of baseline and concomitant medication use may vary based on how the data were collected (IVRS vs. eCRF).

As seen from Appendix Table 15, of the 461 patients in the Maintenance Study ITT Population, 51% had previously failed TNFα antagonist therapy per the eCRF. The proportions of patients who were previously exposed to TNFα antagonist therapy or were naïve to TNFα antagonist therapy were similar between the treatment groups. In addition, the treatment groups were similar with respect to the number of TNFα antagonist therapies patients had previously failed.

Treatment failure to CD therapies was categorized using the hierarchical approach. The proportions of patients who had previously failed TNFα antagonists, immunomodulators, and corticosteroids were similar among the treatment groups in the Maintenance Study ITT Population. Most patients who had previously failed TNFα antagonists had either an inadequate response or lost response to these agents. Notably, 45% of the patients in the placebo and vedolizumab Q8W treatment groups and 40% of the patients in the vedolizumab Q8W group who had previously failed TNFα antagonist therapy had an inadequate response (primary failure).

The proportion of patients with prior TNFα antagonist failure was higher in the all vedolizumab combined group (62%), which includes the non-ITT vedolizumab patients (67%) who had failed to respond during the Induction Phase.

Prior CD therapies of the Maintenance Phase Safety Population are summarized by treatment group and presented in Appendix Table 16.

As seen from Appendix Table 16, in the Overall ITT Population, all patients in both treatment groups reported prior therapy with other CD treatments. Prior exposure to systemic corticosteroids was reported by 91% of the combined vedolizumab patients and by 94% of the combined placebo patients. Prior exposure to immunomodulators was reported by 84% of the combined vedolizumab patients and by 80% of the combined placebo patients.

Baseline (i.e., Week 0) CD therapy, as recorded by the IVRS, is summarized for the Maintenance Phase Safety Population in Appendix Table 17.

As seen from Appendix Table 17, in the Maintenance Study ITT Population, CD therapy use at baseline was similar among the treatment groups.

Appendix Table 18 presents the comparison of selected baseline (Week 0) CD characteristics and medication use among the treatment groups in the Maintenance Study ITT Population.

As seen from Appendix Table 18, no significant differences were noted among the treatment groups for selected baseline CD characteristics; including mean duration of CD, mean disease activity, corticosteroid use at randomization, immunomodulator use at randomization, prior TNFα antagonist use, and prior failure to TNFα antagonist therapy.

3.1.1.2.2.4 Analysis Populations

All patients randomized into the Maintenance Study ITT Population were treated with vedolizumab during the Induction Phase and achieved clinical response. Patients in the placebo treatment group for the Maintenance Study ITT Population received their first dose of placebo at Week 6.

The table below summarizes the efficacy and safety analysis populations for the Maintenance Study ITT Population.

Table 13 Summary of Efficacy and Safety Analysis Populations for Maintenance Phase Study C13007

Data Set, n (%)	PLA N = 153	VDZ Q8W N = 154	VDZ Q4W N = 154
Randomized patients	153	154	154
Safety Population[a]	153 (100)	154 (100)	154 (100)
Intent-to-Treat Population[b]	153 (100)	154 (100)	154 (100)
Modified ITT Population[c]	153 (100)	150 (97)	151 (98)
Per-Protocol Population[d]	147 (96)	149 (97)	144 (94)
Completers (Observed Case) Population[e]	63 (41)	72 (47)	81 (53)

Copied from Table 56, page 226 CSR.

Three additional analysis populations were detailed for the Maintenance Phase of the study. The Delayed-Response Population included all vedolizumab-treated patients who did not meet the protocol definition of calculated clinical response (as assessed by the investigator) and were classified as non-responders by IVRS at Week 6. The Delayed-Response Population was used for all analyses conducted at Weeks 10 and 14.

An imbalance across the treatment groups in the proportion of patients who had achieved clinical remission at Week 6 was observed because randomization at Week 6 was not stratified by remission status. Only 27.9% of vedolizumab Q4W patients and 33.8% of vedolizumab Q8W patients had achieved clinical remission at Week 6 compared to 36.6% of placebo patients. The applicant stated that this imbalance may have had an impact on the analyses of the clinical remission-based primary endpoint as well as the secondary and exploratory endpoints, especially for the vedolizumab Q4W group versus the placebo group.

3.1.1.2.2.5 Applicant's Analyses of the Primary Efficacy Endpoint

The primary efficacy endpoint for the Maintenance Study was the proportion of patients in clinical remission at Week 52.

The CMH chi-square test was performed with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to TNFα antagonists and/or concomitant immunomodulator use (yes/no); 3) enrollment in Cohort 1 or Cohort 2 in the Induction Phase.

Results are summarized by treatment groups for the Maintenance Study ITT population in the table below.

Table 14 Clinical Remission at Week 52
Maintenance Study ITT Population
Study C13007

Clinical Remission[a]	PLA N = 153	VDZ Q8W N = 154	VDZ Q4W N = 154
Number (%) achieving endpoint	33 (21.6)	60 (39.0)	56 (36.4)
95% CI	(15.1, 28.1)	(31.3, 46.7)	(28.8, 44.0)
Difference from placebo[b]		17.4	14.7
95% CI for difference from placebo		(7.3, 27.5)	(4.6, 24.7)
P-value for difference from placebo[c]		0.0007	0.0042
Relative risk[d]		1.8	1.7
95% CI for relative risk		(1.3, 2.6)	(1.2, 2.4)

Copied from Table 57, page 228 CSR.

As seen from the table above, statistically significantly greater proportions of vedolizumab-treated patients in the Q8W (39.0%) and Q4W (36.4%) treatment groups achieved clinical remission at Week 52 compared with patients who received placebo (21.6%; p = 0.0007 and p = 0.0042, respectively). In the vedolizumab Q8W group, the treatment difference from placebo was 17.4% (95% CI: 7.3, 27.5). In the vedolizumab Q4W group, the treatment difference from placebo was 14.7% (95% CI: 4.6, 24.7).

The proportions of patients who achieved clinical remission at Week 52 are presented in Appendix Table 19 for the Per-Protocol Population and in Appendix Table 20 for the Completers (Observed Case) Population.

As seen from Appendix Tables 19 and 20, results of these analyses were similar to those observed in the Maintenance Study ITT Population, with statistically significantly greater proportions of vedolizumab-treated patients in the Q8W and Q4W treatment groups achieving clinical remission at Week 52 compared to the placebo group.

The applicant found that there were a total of 107 patients whose response status at Week 6 was miss-categorized as reported by the investigator (Appendix Table 21, post hoc analysis). Forty-

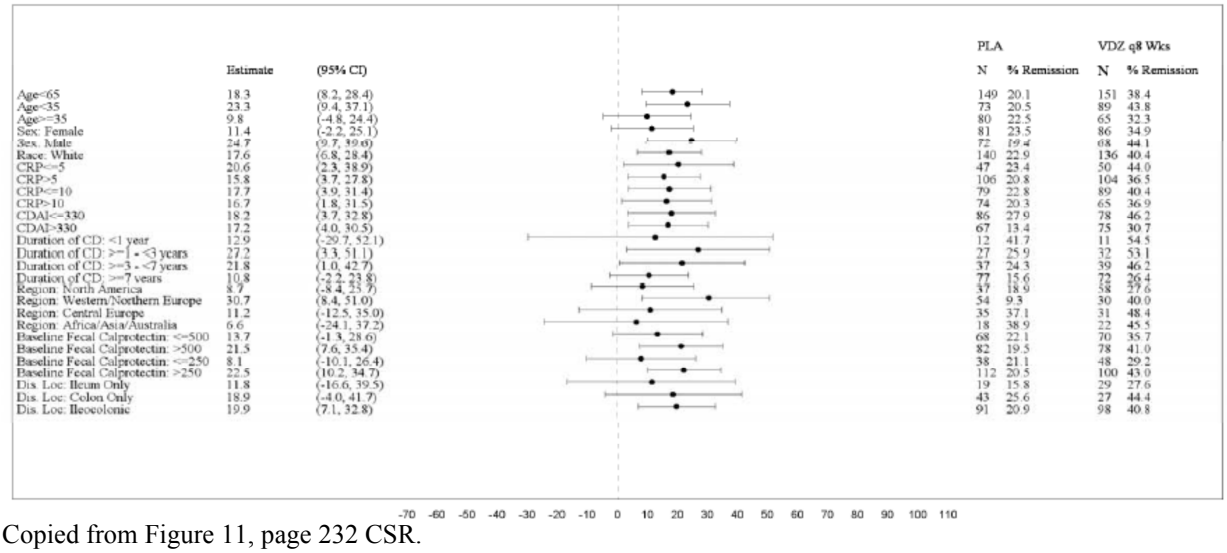
one (41) patients were reported by the investigator as "non-responders" but were "responders" as calculated by the applicant based on IVRS-reported patient subscores (number of liquid or very soft stools, abdominal pain, and general well-being), investigator assessments (extra-intestinal manifestations, abdominal mass), and other CDAI components (anti-diarrheal, hematocrit, and body weight); 66 patients were categorized as "responders" by the investigator but were "non-responders" as calculated by the applicant. Of these miss-categorized patients, 15 patients were in the non-ITT groups, and 51 were in the maintenance ITT groups. Of the 51 patients who were re-randomized into the Maintenance ITT Population, 16 (10%) were in the placebo group, 19 (12%) in the vedolizumab Q8W group and 16 (10%) in the vedolizumab Q4W group.

Post hoc sensitivity analyses were performed by the applicant to assess the impact of the inclusion of these patients in the Maintenance Study; the primary endpoint of clinical remission and the secondary endpoint of enhanced clinical response were assessed for all patients in the ITT population who met the protocol definition of clinical response at Week 6. The results of these analyses were similar to that of the primary efficacy analysis, with treatment differences for the vedolizumab Q8W and Q4W groups of 17.7% (95% CI: 6.8, 28.6, p = 0.0014) and 16.6% (95% CI: 5.8, 27.4; p = 0.0027), respectively (Appendix Table 22, post hoc analysis) for the endpoint of clinical remission,.

3.1.1.2.2.5.1 Subgroup Analyses

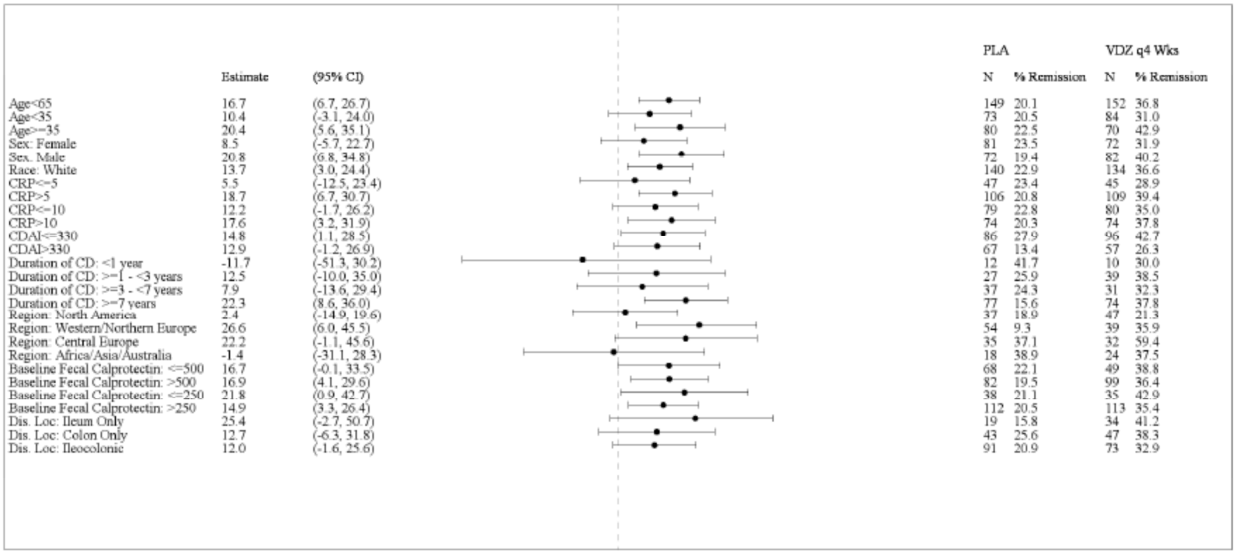
The figures below summarize the risk differences (percentages) from placebo for the vedolizumab Q8W and Q4W treatment groups, respectively, for the primary endpoint of clinical remission at Week 52 in patient subgroups according to demographic characteristics and measures of disease severity in the Maintenance Study ITT Population.

Figure 4 Risk Difference (Percentage) and the 95% Confidence Intervals for Subgroup Analyses of Clinical Remission at Week 52 for Vedolizumab Q8W vs. Placebo Maintenance Study ITT Population Study C13007



Copied from Figure 11, page 232 CSR.

Figure 5 Risk Difference (Percentage) and the 95% Confidence Intervals for Subgroup Analyses of Clinical Remission at Week 52 for Vedolizumab Q4W vs. Placebo Maintenance Study ITT Population
Study C13007



Copied from Figure 13, page 234

As seem from the figures above, the treatment benefit of vedolizumab for the maintenance of clinical remission at Week 52 in the Maintenance Study ITT Population was preserved in patient subgroups according to demographic variables and disease characteristics. In both vedolizumab groups, the treatment effect was observed in a majority of the patient subgroups by age, gender, race, and geographic region, although not all of the treatment difference 95% CIs exc luded zero. Both males and females had a positive response to treatment, but the treatment differences from placebo were greater in males compared with females in both vedolizumab Q8W and Q4W treatment groups. With respect to age, the treatment difference in the vedolizumab Q8W treatment group was greater for patients < 35 years of age than for patients 35 years of age or older. Conversely, in the vedolizumab Q4W treatment group, patients 35 years of age or older had a greater treatment difference from placebo compared to patients < 35 years of age.

Similar results were also observed for the subgroups according to disease activity and severity, including CDAI, baseline CRP, baseline fecal calprotectin, and disease location. Consistent with the results observed for age, treatment differences from placebo were greater among patients in the vedolizumab Q8W treatment group with a disease duration ≥ 1 to < 3 years and ≤ 3 to < 7 years, compared to those with a disease duration ≥ 7 years, whereas the converse was observed in the vedolizumab Q4W treatment group.

3.1.1.2.2.6 Applicant's Analyses of the Secondary Efficacy Endpoints

The three secondary efficacy endpoints for this study are presented by treatment group as follows:

- Proportion of patients with enhanced clinical response at Week 52 in the

Maintenance Study ITT Population

- Proportion of patients in corticosteroid-free clinical remission at Week 52 in the Maintenance Study ITT Population
- Proportion of patients in durable clinical remission at Week 52 in the Maintenance Study ITT Population

3.1.1.2.2.6.1 Enhanced Clinical Response at Week 52

The number and proportion of patients with enhanced clinical response at Week 52 are summarized by treatment groups for the Maintenance Study ITT population in the table below.

Table 15 Enhanced Clinical Response at Week 52
Maintenance Study ITT Population
Study C13007

Enhanced Clinical Response[a]	PLA N = 153	VDZ Q8W N = 154	VDZ Q4W N = 154
Number (%) achieving endpoint	46 (30.1)	67 (43.5)	70 (45.5)
95% CI	(22.8, 37.3)	(35.7, 51.3)	(37.6, 53.3)
Difference from placebo[b]		13.4	15.3
95% CI for difference from placebo		(2.8, 24.0)	(4.6, 26.0)
P-value for difference from placebo[c]		0.0132	0.0053
Relative risk[d]		1.4	1.5
95% CI for relative risk		(1.1, 1.9)	(1.1, 2.0)

Copied from Table 58, page 237 CSR.

As seen from the table above, statistically significantly greater proportions of vedolizumab-treated patients in the Q8W (43.5%) and Q4W (45.5%) treatment groups achieved enhanced clinical response at Week 52 compared with patients who received placebo (30.1%; p = 0.0132 and p = 0.0053, respectively). In the vedolizumab Q8W group, the treatment difference from placebo was 13.4% (95% CI: 2.8, 24.0). In the vedolizumab Q4W group, the treatment difference from placebo was 15.3% (95% CI: 4.6, 26.0).

Post hoc sensitivity analyses were done by the applicant to assess the impact of the inclusion of these patients whose response status at Week 6 was miss-categorized in the Maintenance Phase. The secondary endpoint of enhanced clinical response was assessed for all patients in the ITT population who met the protocol definition of clinical response at Week 6. The results of this analysis were similar to those of the primary efficacy analysis, with treatment differences for the vedolizumab Q8W and Q4W groups of 15.4% (95% CI: 4.0, 26.9, p = 0.0082) and 18.0% (95% CI: 6.6, 29.5; p = 0.0021), respectively (Appendix Table 23, post hoc analysis) for the endpoint of enhanced clinical response.

3.1.1.2.2.6.2 Corticosteroid-free Clinical Remission at Week 52

Patients receiving corticosteroids at the beginning of the Maintenance Phase were to taper the medications according to the regimen. Slightly more than half of the patients in each treatment group (52% to 54%) were receiving corticosteroid therapy at the start of the Maintenance Phase.

The number and proportion of patients in the Maintenance Study ITT population with corticosteroid-free clinical remission at Week 52 are summarized by treatment group in the table below.

Table 16 Corticosteroid-free Clinical Remission at Week 52
Maintenance Study ITT Population
Study C13007

Corticosteroid-free Clinical Remission[a]	PLA N = 82	VDZ Q8W N = 82	VDZ Q4W N = 80
Number (%) achieving endpoint	13 (15.9)	26 (31.7)	23 (28.8)
95% CI	(7.9, 23.8)	(21.6, 41.8)	(18.8, 38.7)
Difference from placebo[b]		15.9	12.9
95% CI for difference from placebo		(3.0, 28.7)	(0.3, 25.5)
P-value for difference from placebo[c]		0.0154	0.0450
Relative risk[d]		2.0	1.8
95% CI for relative risk		(1.1, 3.6)	(1.0, 3.3)

Copied from Table 59, page 245 CSR.

As seen from the table above, among these patients, statistically significantly greater proportions treated with vedolizumab in the Q8W (31.7%) and Q4W (28.8%) treatment groups achieved corticosteroid-free clinical remission at Week 52 compared with patients who received placebo (15.9%; p = 0.0154 and p = 0.0450, respectively). In the vedolizumab Q8W group, the treatment difference from placebo was 15.9% (95% CI: 3.0, 28.7). In the vedolizumab Q4W group, the treatment difference from placebo was 12.9% (95% CI: 0.3, 25.5),

3.1.1.2.2.6.3 Durable Clinical Remission

Durable clinical remission in the Maintenance Study was defined as clinical remission in ≥ 80% of study visits, including the Week 52 visit (11 out of 13 study visits). The number and proportion of patients in the Maintenance Study ITT population with a durable clinical remission are summarized by treatment groups in the table below.

Table 17 Durable Clinical Remission at Week 52
Maintenance Study ITT Population
Study C13007

Durable Clinical Remission[a]	PLA N = 153	VDZ Q8W N = 154	VDZ Q4W N = 154
Number (%) achieving endpoint	22 (14.4)	33 (21.4)	25 (16.2)
95% CI	(8.8, 19.9)	(14.9, 27.9)	(10.4, 22.1)
Difference from placebo[b]		7.2	2.0
95% CI for difference from placebo		(-1.5, 16.0)	(-6.3, 10.2)
P-value for difference from placebo[c]		0.1036	0.6413
Relative risk[d]		1.5	1.1
95% CI for relative risk		(0.9, 2.4)	(0.7, 1.9)

Copied from Table 60 page 247 CSR.

As seen from the table above, no statistically significant differences were observed between either vedolizumab group and the placebo group for the endpoint of durable clinical remission, although a trend of treatment difference was observed in favor of for the vedolizumab Q8W group (7.2%).

3.1.1.2.2.7 Applicant's Exploratory Endpoints at Week 52

3.1.1.2.2.7.1 Key Maintenance Endpoints in Patients by Prior TNFα Antagonist Use or Failure

The number and proportion of patients who achieved clinical remission, enhanced clinical response, and corticosteroid-free clinical remission at Week 52 are summarized by treatment group in the table below for those patients in the Maintenance Study ITT Population without prior TNFα antagonist exposure (TNFα antagonist naïve patients) and for those who had previously failed TNFα antagonist therapy.

Table 18 Key Maintenance Efficacy Endpoints by Prior TNFα Antagonist Use or Failure Maintenance Study ITT Population Study C13007

	Patients Without Prior TNFα Antagonist Use (Naïve)[a]			Patients with Prior TNFα Antagonist Failure[b]		
Response at Week 52	PLA N = 71	VDZ Q8W N = 66	VDZ Q4W N = 71	PLA N = 78	VDZ Q8W N = 82	VDZ Q4W N = 77
Number (%) achieving clinical remission[a]	19 (26.8)	34 (51.5)	33 (46.5)	10 (12.8)	23 (28.0)	21 (27.3)
95% CI	(16.5, 37.1)	(39.5, 63.6)	(34.9, 58.1)	(5.4, 20.2)	(18.3, 37.8)	(17.3, 37.2)
Difference from placebo[b]		24.8	19.7		15.2	14.5
95% CI for difference from placebo		(8.9, 40.6)	(4.2, 35.2)		(3.0, 27.5)	(2.0, 26.9)
Number (%) achieving enhanced clinical response[c]	27 (38.0)	40 (60.6)	38 (53.5)	16 (20.5)	24 (29.3)	29 (37.7)
95% CI	(26.7, 49.3)	(48.8, 72.4)	(41.9, 65.1)	(11.6, 29.5)	(19.4, 39.1)	(26.8, 48.5)
Difference from placebo[b]		22.6	15.5		8.8	17.1
95% CI for difference from placebo		(6.3, 38.9)	(-0.7, 31.7)		(-4.6, 22.1)	(3.1, 31.2)
	N = 40	N = 38	N = 36	N = 38	N = 41	N = 43
Number (%) achieving corticosteroid-free clinical remission[d]	11 (27.5)	15 (39.5)	16 (44.4)	0	10 (24.4)	7 (16.3)
95% CI	(13.7, 41.3)	(23.9, 55.0)	(28.2, 60.7)	(0.0, 9.3)	(12.4, 40.3)	(6.8, 30.7)
Difference from placebo[b]		12.0	16.9		24.4	16.3
95% CI for difference from placebo		(-8.8, 32.8)	(-4.4, 38.3)		(2.4, 45.1)	(-5.7, 37.0)

Copied from Table 61, page 249 CSR.

As seen from the table above, among TNFα antagonist naïve patients, greater proportions of vedolizumab-treated patients in the Q8W and Q4W treatment groups achieved the primary endpoint of clinical remission at Week 52 and the secondary endpoints of enhanced clinical response and corticosteroid-free remission at Week 52 compared with patients who received placebo. The results observed in this subgroup of patients for clinical remission in both vedolizumab groups and for enhanced clinical response in the vedolizumab Q8W group were consistent with the statistically significant treatment differences observed in the overall Maintenance Study ITT Population. The other treatment differences also favored the vedolizumab groups, but the 95% CIs included zero.

Among the patients who had previously failed TNFα antagonist therapy, greater proportions of vedolizumab-treated patients in the Q8W and Q4W treatment groups achieved the primary endpoint of clinical remission at Week 52 and the secondary endpoints of enhanced clinical response and corticosteroid-free remission at Week 52 compared with the patients who received placebo. The results observed in this subgroup of patients for clinical remission in both vedolizumab groups, for enhanced clinical response in the vedolizumab Q4W group, and for corticosteroid-free clinical remission in the vedolizumab Q8W group were consistent with the statistically significant treatment differences observed in the overall Maintenance Study ITT Population. The other treatment differences also favored the vedolizumab groups, but the 95% CIs included zero.

Treatment differences from placebo for the endpoints of clinical remission and enhanced clinical response at Week 52 for the vedolizumab Q8W treatment group were higher for TNFα antagonist naïve patients compared to those who have previously failed TNFα antagonist therapy; treatment differences from placebo for the vedolizumab Q4W treatment group were similar between the TNFα antagonist naïve and failure subgroups. For the endpoint of

corticosteroid-free clinical remission at Week 52, the treatment difference from placebo for the vedolizumab Q8W treatment group was higher among patients who had previously failed TNFα antagonist therapy compared with those who were naïve to TNFα antagonist therapy; treatment differences for the vedolizumab Q4W treatment group were similar between the subgroups.

3.1.1.3 Reviewer's Comments and Evaluation

3.1.1.3.1 Induction Phase

3.1.1.3.1.1 Treatment Group Comparability in Cohort 1 and Cohort 2

The major differences between Cohort 1 and Cohort 2 are given below:

- The open-label vedolizumab group had more patients enrolling at sites in Western/Northern Europe and fewer patients entering at sites in Asia/Australia/Africa and Eastern Europe than was observed for the Induction Study ITT Population.

- The vedolizumab group had greater proportions of patients with CD duration of ≥ 7 years (50%) and with a history of prior surgery for CD (45%) compared to the placebo group (43% and 36%, respectively).

- For prior use of TNFα antagonists and treatment failure to CD therapies, when compared to the Induction Study ITT Population, the open-label vedolizumab group had a greater proportion of patients who had prior TNFα antagonist use (68%) and failure (63%), with most patients having shown inadequate response (primary failure: 47%) or loss of response (secondary failure: 40%).

3.1.1.3.1.2 Primary Efficacy Endpoints

In the Type C Meeting held on September 10, 2009, the applicant stated that there was an unanticipated population shift and potential negative impact of this shift on the primary endpoint, clinical remission (CDAI ≤150).

The applicant proposed to elevate the first key secondary endpoint, enhanced clinical response (decrease in CDAI of ≥ 100 points) , to a co-primary endpoint The applicant further defined that co-primary means that the primary objective of the study would be met by achieving statistical significance for either of the co-primary endpoints. Hochberg method was used to adjust for the multiplicity comparisons on the two primary endpoints.

In the response to applicant's question 2, this reviewer stated the following:

- The term of co-primary endpoint defined by the applicant for Study C13007 is not commonly used for regulatory purpose.

- Two or more primary endpoints are called co-primary if each must be shown statistically significant treatment benefit at a pre-specified significance level α (e.g., α=0.025, by one-sided tests).

3.1.1.3.1.2.1 The Difference between Clinical Remission at Week 6 and Enhanced Clinical Response at Week 6

The medical reviewer, Klaus Gottlieb, M.D. found seven patients in clinical remission (CDAI ≤ 150) did not meet the criteria for enhanced clinical response (decrease in CDAI of ≥ 100 points from baseline) at Week 6 for this study.

The table below is the list of these seven patients.

Table 19 Patients Who Achieved Clinical Remission at Week 6 but Failed to Achieve Clinical Response Study C13007

Obs	USUBJID	IARM	BASECDAI	CDAI6	CDAICHG6	WK6CR	WK6ECR	WK6RM
1	C13007-22009-701	PLA	155	97	-58	N	N	Y
2	C13007-24005-702	VDZ	132	54	-78	Y	N	Y
3	C13007-42004-701	VDZ	192	146	-46	N	N	Y
4	C13007-55006-703	PLA	191	130	-61	N	N	Y
5	C13007-58093-705	VDZ	218	150	-68	N	N	Y
6	C13007-58108-702	VDZ	142	128	-14	N	N	Y
7	C13007-58132-701	VDZ	213	128	-85	Y	N	Y

Complied by this reviewer.

As seen from the table above, all seven patients (five in vedolizumab; two in placebo) did not meet baseline enrollment criteria. This reviewer performed post-hoc sensitivity analysis by excluding these 7 patients. The resulting treatment different would be 7.08% with nominal p-value of 0.0293 (Fisher's Exact test). If these seven patients were considered as "non-responders", the resulting treatment difference would be 6.87% with nominal p-value of 0.0299 (Fisher's Exact test).

3.1.1.3.1.2.2 Comments on Applicant's ITT Analysis of Clinical Remission at Week 6

This reviewer found 20 patients (10 patients in each treatment group) who had baseline CDAI score of less than 220 and 2 patients (1 patient in each treatment group) with baseline CDAI missing were enrolled in this study. The listing of these patients and their outcome is given below.

Table 20 Clinical Outcomes for Patients Who Had Baseline CDAI < 220 Study C13007

Obs	USUBJID	IARM	BASECDAI	CDAI6	CDAICHG6	WK6CR	WK6ECR	WK6RM
1	C13007-03002-703	PLA	196	155	-41	N	N	N
2	C13007-06004-703	PLA	160	201	41	N	N	N
3	C13007-07019-704	PLA	213	177	-36	N	N	N
4	C13007-07032-707	PLA	177	244	67	N	N	N
5	C13007-07148-704	PLA	210	165	-45	N	N	N
6	C13007-12009-706	VDZ	215	189	-26	N	N	N
7	C13007-12019-706	VDZ	213	90	-123	Y	Y	Y
8	C13007-17002-702	PLA	216	324	108	N	N	N
9	C13007-21001-702	PLA	214	208	-6	N	N	N
10	C13007-22009-701	PLA	155	97	-58	N	N	Y
11	C13007-24001-702	PLA	208	.	.			
12	C13007-24005-702	VDZ	132	54	-78	Y	N	Y
13	C13007-29003-703	VDZ	204	187	-17	N	N	N
14	C13007-42004-701	VDZ	192	146	-46	N	N	Y
15	C13007-46009-701	VDZ	200	220	20	N	N	N
16	C13007-49002-702	VDZ	141	41	-100	Y	Y	Y
17	C13007-55005-701	PLA	.	.	.			
18	C13007-55005-702	VDZ	.	261	.			N
19	C13007-55006-703	PLA	191	130	-61	N	N	Y
20	C13007-58093-705	VDZ	218	150	-68	N	N	Y
21	C13007-58108-702	VDZ	142	128	-14	N	N	Y
22	C13007-58132-701	VDZ	213	128	-85	Y	N	Y

Copied by this reviewer.

As seen from the table above, among these 20 patients with a baseline CDAI < 220, a greater proportion of vedolizumab-treated patients achieved clinical remission at Week 6 as compared with patients who received placebo [70% (7/10) vs. 20% (2/10)].

This reviewer also found 18 patients (9 in each treatment group) who had a baseline CDAI score of greater than 450. Among these 18 patients with a baseline CDAI > 450, proportions of vedolizumab-treated and placebo patients achieved clinical remission at Week 6 were zeros.

According to the inclusion criteria (a CDAI score of 220 to 450), these 40 patients should not be enrolled in the study, this reviewer performed post-hoc sensitivity analysis by excluding these 40 patients. The resulting treatment different would be 6.25% with nominal p-value of 0.0893 (Fisher's Exact test). If these 40 patients were considered as "non-responders", the resulting treatment difference would be 5.95% with nominal p-value of 0.0622 (Fisher's Exact test).

3.1.1.3.1.2.3 Comments on Applicant's Per Protocol Analysis of Clinical Remission at Week 6

The applicant stated that three patients in each treatment group with baseline CDAI scores that were either missing (1 patient in each treatment group) or < 210 (2 patients in each treatment group; baseline CDAI score range: 132 to 208) were excluded from the Per-Protocol Population.

However, this reviewer found that 16 patients (8 in each treatment group) who had baseline CDAI score of less than 220 were enrolled in this study. The listing of these patients and their outcome is given in the table below.

Table 21 Clinical Outcomes for Patients who had baseline CDAI < 220 Study C13007

Obs	USUBJID	IARM	BASECDAI	CDAI6	CDAICHG6	WK6CR	WK6ECR	WK6RM
1	C13007-03002-703	PLA	196	155	-41	N	N	N
2	C13007-06004-703	PLA	160	201	41	N	N	N
3	C13007-07019-704	PLA	213	177	-36	N	N	N
4	C13007-07032-707	PLA	177	244	67	N	N	N
5	C13007-07148-704	PLA	210	165	-45	N	N	N
6	C13007-12009-706	VDZ	215	189	-26	N	N	N
7	C13007-12019-706	VDZ	213	90	-123	Y	Y	Y
8	C13007-17002-702	PLA	216	324	108	N	N	N
9	C13007-21001-702	PLA	214	208	-6	N	N	N
10	C13007-22009-701	PLA	155	97	-58	N	N	Y
11	C13007-29003-703	VDZ	204	187	-17	N	N	N
12	C13007-42004-701	VDZ	192	146	-46	N	N	Y
13	C13007-46009-701	VDZ	200	220	20	N	N	N
14	C13007-58093-705	VDZ	218	150	-68	N	N	Y
15	C13007-58108-702	VDZ	142	128	-14	N	N	Y

Obs	USUBJID	IARM	BASECDAI	CDAI6	CDAICHG6	WK6CR	WK6ECR	WK6RM
16	C13007-58132-701	VDZ	213	128	-85	Y	N	Y

As seen from the able above, among these 16 patients with baseline CDAI < 220, a greater proportion of vedolizumab-treated patients achieved clinical remission at Week 6 as compared with the patients who received placebo [62.5% (5/8) vs. 12.5% (1/8)].
According to the inclusion criteria (a CDAI score of 220 to 450), these 16 patients should not be enrolled in the study, this reviewer performed post-hoc sensitivity analysis by excluding these 16 patients. The resulting treatment different would be 6.67% with nominal p-value of 0.0606 (Fisher's Exact test). If these 16 patients were considered as "non-responders", the resulting treatment difference would be 6.53% with nominal p-value of 0.0611 (Fisher's Exact test).

3.1.1.3.1.2.4 Additional Comments on Applicant's ITT Analysis of Clinical Remission at Week 6

This reviewer performed a post-hoc unadjusted analysis of clinical remission at Week 6 using Fisher's exact test to see whether the applicant's result was robust and method independent.

The result p-value from Fisher's exact test yielded 0.0287 which is greater than 0.025, level of significance. So, the applicant's analysis of clinical remission at Week 6 might not be robust and was method dependent.

3.1.1.3.1.3 Additional Subgroup Analyses

3.1.1.3.1.3.1 Subgroup Analyses of Clinical Remission at Week 6 by baseline CDAI (CDAI ≤ 330 vs. CDAI > 330), Inflammatory (High vs. Low) and by Geographic Regions (North America, Western North Europe, Central Europe, Eastern Europe, Africa, Asian, and Australia.

A summary of subgroup analyses of clinical remission at Week 6 for baseline CDAI, geographic regions, inflammatory, inflammatory by baseline CDAI are given in the table below.

Table 22 Subgroup Analyses for Clinical Remission At Week 6
Study C13007

Category	Placebo	vedolizumab	Difference	95% CI
Country				
North America	2/50 (4.0%)	6/64 (9.4%)	5.4%	(-5.2, 15.7)
Western North Europe	1/22 (4.5%)	3/28 (10.7%)	6.2%	(-13.1, 23.9)
Central Europe	2/30 (6.7%)	11/45 (24.4%)	17.8%	(0.0, 34.2)
Eastern Europe	3/17 (17.6%)	3/31 (9.9%)	-7.9%	(-34.3, 12.9)
Africa, Asia, Australia	2/29 (6.9%)	9/52 (17.3%)	10.4%	(-6.6, 24.9)
Baseline CDAI				
≤ 330	6/81 (7.4%)	27/119 (22.7%)	15.3%	(4.8, 24.9)
>330	4/66 (6.1%)	5/100 (5.0%)	-1.1%	(-10.1, 6.3)
Inflammatory				
High	3/76 (3.9%)	18/124 (14.5%)	10.6%	(2.2, 19.7)
Low	7/65 (10.8%)	12/86 (14.0%)	3.2%	(-8.3, 14.1)
Inflammatory by Baseline CDAI				
High; ≤330	1/41 (2.4%)	15/59 (25.4%)	23.0%	(9.7, 36.4)
High; >330	2/34 (5.9%)	3/64 (4.7%)	-1.2%	(-15.0, 8.8)
Low; ≤330	5/35 (14.3%)	11/56 (19.6%)	5.4%	(-12.2, 20.9)
low; >330	2/30 (6.7%)	1/30 (3.3%)	-3.3%	(-20.0, 11.1)

Complied by this reviewer.

As seen from the table above, proportion of patients in clinical remission at Week 6 was consistent for geographic region except Eastern Europe.

The proportion of patients in clinical remission at Week 6 was inconsistent for baseline CDAI (≤330 vs. >330) and inflammatory by baseline CDAI.

The 95% confidence intervals do not include zero for Central Europe, baseline CDAI ≤330, and high inflammatory, and high inflammatory and baseline CDAI ≤ 330.

Furthermore, the treatment group differed with respect to the proportion of patients enrolled by geographic site with p-value of 0.0610. This reviewer performed a post-hoc analysis of clinical remission at Week 6 adjusted for geographic site using the CMH chi-square test to see whether the applicant's result was robust and method independent.

The resulting p-value from the CMH chi-square test yielded 0.0279 which is greater than 0.025, level of significance. So, the applicant's analysis of clinical remission at Week 6 might not be robust and was method dependent.

This reviewer also performed a post-hoc Breslow-Day test for homogeneity of odd ratios for baseline CDAI (≤330 vs. >330). The resulting p-value from the Breslow-Day test yielded 0.0636 which is smaller than 0.10, level of significance. There may be heterogeneity of odd ratios between baseline CDAI (≤330 vs. >330) subgroups.

Vedolizumab might be effective for patients with baseline CDAI ≤ 330, high inflammatory, and high inflammatory and CDAI ≤ 330.But, they need to be confirmed by the other study.

6.1.1.3.1.3.2 Subgroup Analyses of Clinical Remission at Week 6 by Baseline CDAI

This reviewer performed additional subgroup analysis of clinical remission at Week 6 by various categorization of baseline CDAI.

The summary of subgroup analyses of clinical remission at Week 6 for various categorization of baseline CDAI are given below.

Table 23 Subgroup Analyses of Clinical Remission At Week 6 for Various Categorization of Baseline CDAI Study C13007

Category	Placebo	vedolizumab	Difference	95% CI
≤ 200	2/5 (40.0%)	4/6 (66.7%)	26 7%	(-39.0, 76.4)
<200 - ≤250	0/17 (0.0%)	7/20 (35.0%)	35.0%	(12.3, 59.2)
<250 - ≤450	8/105 (7.6%)	21/166 (12.7%)	5.1%	(-2.8, 12.3)
>450	0/9 (0.0%)	0/9 (0.0%)	0.0%	(-33.6, 33.6)
≤ 200	2/5 (40.0%)	4/6 (66.7%)	26 7%	(-39.0, 76.4)
<200 - ≤220	0/4 (0.0%)	3/5 (60.0%)	60.0%	(-8.9, 94.7)
<220 - ≤330	4/66 (6.1%)	20/104 (19.2%)	13.1%	(2.9, 23.0)
<330 - ≤450	4/52 (7.7%)	5/77 (6.5%)	-1.2%	(-12.0, 8.3)
>450	0/9 (0.0%)	0/9 (0.0%)	0.0%	(-33.6, 33.6)
≤220	2/9 (22.2%)	7/11 (63.6%)	41.40%	(-4.6, 75.7)
<220 - ≤330	4/66 (6.1%)	20/104 (19.2%)	13.1%	(2.9, 23.0)
<330 - ≤450	4/52 (7.7%)	5/77 (6.5%)	-1.2%	(-12.0, 8.3)
>450	0/9 (0.0%)	0/9 (0.0%)	0.0%	(-33.6, 33.6)
≤220	2/9 (22.2%)	7/11 (63.6%)	41.40%	(-4.6, 75.7)
<220 - ≤450	8/118 (6.8%)	25/181 (13.8%)	7.0%	(-0.2, 13.9%)
>450	0/9 (0.0%)	0/9 (0.0%)	0.0%	(-33.6, 33.6)

Compiled by this reviewer.

As seen from the table above, the results seemed to be in favor of vedolizumab more in the subgroups of patients who had smaller baseline CDAI (e.g. ≤250).

3.1.1.3.1.4 Sensitivity Analysis

Per FDA's request, to address the missing data issue, the applicant performed the following sensitivity analyses for the primary and secondary endpoints for this study:

- Observed case: exclude subjects from the analysis at a specific time point if the subjects have insufficient data at that time point.

- Complete case: exclude subjects from the analysis at all time points if they have insufficient data at any of the time points of analysis.
- Worst case: (1) subjects with missing observations at any of the time points of analysis are assumed to be non-responders; (2) subjects receiving placebo with missing observations at any of the time points of analysis are assumed to be responders, and subjects receiving treatment with missing observations at any of the time points of analysis are assumed to be non-responders.
- LOCF analysis
- Multiple imputation

The applicant stated clarifying information relating to the five bulleted questions above as follows:

- Observed Case and Complete Case: The observed and complete case analyses are identical to the analyses done without imputation. Only one set of analysis is provided with this response because observed case and complete case are identical analyses. In the observed case, insufficient data at a specific time point implies that there are no data at Week 6 or Week 52 (Study C13007) and no data at Week 6 or Week 10 (Study C13011.) In such case, observed case is equivalent to analyses without imputation. In the complete case, insufficient data at all analyses time points indicate that there are no data at Week 6 or Week 52 (Study 7) and no data at Week 6 and Week 10 (Study C13011).
- Worst Case 2: The requested analyses are provided in this response. Patients receiving placebo who had missing data were assumed to be responders and patients receiving vedolizumab who had missing data were assumed to be non-responders. There are limitations to this analysis, due to an imbalance in missing data between placebo and vedolizumab groups. This is because larger numbers of placebo patients failed treatment earlier and were allowed to enroll in Study C13008. Thus, considering failure as a success for the placebo group will be biased against the vedolizumab group.
- LOCF analyses: The requested analyses are provided in this response. If a subject had missing data at a particular time point, then data from the prior time point was imputed.
- Multiple imputations: The requested analyses are provided in this response. A multiple imputation was performed using SAS PROC MI. The number of iterations was set to 10. For the Induction Phase of Study C13007 and C13011, stratification factors of concomitant use of oral corticosteroids, previous exposure to TNFα antagonist and/or concomitant immunomodulator use were used as adjusting factors. For the Maintenance Phase of Study C13007, stratification factors of concomitant use of oral corticosteroids, previous exposure to TNFα antagonist and/or concomitant immunomodulator use, and participation in Cohort 1 or Cohort 2 were used as adjusting factors.

The summary of the results from these sensitivity analyses for clinical remission at Week 6 are given below.

Table 24 Sensitivity Analyses – Clinical Remission at Week 6
Study C13007

Analysis	Placebo	VDZ	Difference	P-value
Primary	10/148 (6.8%)	32/220 (14.5%)	7.8%	0.0206
Observed Case	10/136 (7.4%)	32/200 (16.0%)	8.6%	0.0174
Per Protocol	9/141 (6.4%)	30/205 (14.6%)	8.3%	0.0153
LOCF	10/148 (6.8%)	32/220 (14.5%)	7.8%	0.0206
Multiple Imputation	10/148 (6.8%)	32/220 (14.5%)	7.8%	0.0206

Compiled from Tables 19, 14.3.1.2C, 14.3.1.2B, 39.5.3.1A, 39.5.4.1A. ,39.5.5.1A.
P-values were based on the CMH chi-square test was performed with 2 stratification 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to TNFα antagonists and/or concomitant immunomodulator use (yes/no)

The summary of results from these sensitivity analyses for enhanced response at Week 6 are given below.

Table 25 Sensitivity Analyses – Enhanced Clinical Response at Week 6
Study C13007

Analysis	Placebo	VDZ	Difference	P-value
Primary	38/148 (25.7%)	69/220 (31.4%)	5.7%	0.2322
Observed Case	38/136 (27.9%)	69/200 (34.5%)	6.6%	0.1871
Per Protocol	38/141 (27.0%)	68/205 (33.2%)	6.2%	0.1972
LOCF	38/148 (25.7%)	70/220 (31.8%)	6.1%	0.1981
Multiple Imputation	39/148 (26.4%)	75/220 (34.1%)	7.7%	0.1098

Compiled from Tables 19, 14.3.1 4C, 14.3.1.4B, 39.5.3.1B, 39.5.4.1B, 39.5.5.1B.
P-values were based on the CMH chi-square test was performed with 2 stratification 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to TNFα antagonists and/or concomitant immunomodulator use (yes/no)

As seen from the tables above, results from the sensitivity analyses were similar to those from the primary analysis for clinical remission at Week 6 and enhanced clinical remission at Week 6.

3.1.1.3.1.5 Alternative Definitions for Clinical Remission

Per FDA's request, the summary tables below for the Induction Study ITT Population are provided by the applicant for analyses using the following alternate definition of clinical remission:

- Total number of liquid/very soft stools of ≤ 10 per day in the relevant week; and
- Abdominal pain rated as 0 or 1 for each day in the relevant week.

Table 26 Clinical Remission at Week 6 – Based on Alternative Definition from FDA
Intent-to-Treat Population
Study C13007

Clinical remission	Placebo (n=148)	VDZ (n=220)
N (%) achieve clinical remission at week 6	7 (4.7)	21 (9.5)
95% CI	(1.3, 8.1)	(5.7,13.4)
Difference from placebo		4.8
95% CI for difference from placebo		(-0.7, 10.3)
p-value for difference from placebo		0.0848

Complied from Table 39.2.1.1 by this reviewer.
P-values were based on the CMH chi-square test was performed with 2 stratifications: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to TNFα antagonists and/or concomitant immunomodulator use (yes/no)

As seen from the table above, based on alternative definition from FDA, vedolizumab-treated patients failed to achieve statistical significance for clinical remission at Week 6.

Per FDA's request, the summary tables below for the Induction Study ITT Population are provided by the applicant for analyses using the above alternate definitions of clinical remission regardless of CDAI score.

The applicant stated d that the analyses provided in the tables below have been updated from those provided in the response to Response to Information Request dated 19 August 2013 in two ways:

- CDAI score is not used for the definition of clinical remission, and
- Criterion regarding the number of liquid/very soft stools is calculated as the total of such stools in the 7 days prior to the study visit, not the daily average over the prior week.

Because of these changes, direct comparisons to the tables in the response to the Response to Information Request dated August19 2013 could not be made.

In order to be consistent with the Reviewers Guide located in Data Tabulation Definition.xml for C13007 and the Reviewers Guide located in Data Tabulation Definition.xml for C13011, the same programming rules that were applied in the calculation of CDAI scores for handling missing data were used for the analyses and are detailed here:

- As requested by the FDA, patient diary data from the 7 days prior to the study visit were used for both liquid/very soft stools and abdominal pain score calculations for determination of clinical remission.

- If patient diary data on stool number or abdominal pain were missing in the seven days prior to the visit, the diary data from up to 10 days prior to the study visit were used, starting with eight days prior, etc.

- If there were less than 7 days but more than 3 days of patient diary data within the prior 10 days of the study visit, imputation was used for the missing stool quantity data. To determine the total number of liquid/very soft stools for a 7-day period, the average number of daily stools was calculated from the available diary data and then multiplied by 7.

- No imputation was done for missing abdominal pain; the available data were used to determine clinical remission.

- If a minimum of four days of patient diary data were not available within the ten days prior to the visit, imputation was not performed and the patient was defined as not being in clinical remission, regardless of the available patient diary data.

Table 27 Clinical Remission at Week 6 – Based on Alternative Definition from FDA
Intent-to-Treat Population
Study C13007

Clinical remission	Placebo (n=148)	VDZ (n=220)
N (%) achieve clinical remission at week 6	6 (4.1)	22 (10.0)
95% CI	(0.9, 7.2)	(6.0,14.0)
Difference from placebo		6.0
95% CI for difference from placebo		(0.5, 11.4)
p-value for difference from placebo		0.0332

Complied from Table 39.2.1.1 by this reviewer.
P-values were based on the CMH chi-square test was performed with 2 stratification 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to TNFα antagonists and/or concomitant immunomodulator use (yes/no)

3.1.1.3.2 Maintenance Phase

3.1.1.3.2.1 Data Discrepancy

It was found that there was a discrepancy in the number of vedolizumab patients who were Week 6 responders in the Induction Phase in Cohort 1given in Table 14.3.1.32A of CSR and in Open Label Cohort 2 given in Table 39.31.1.1, (Response to Agency Questions dated October 18) (99 in Cohort 1 and 355 in Cohort 2) and also in the number of patients who were randomized in Maintenance Phase given in Figure 3-1(C13007 FESA) (96 in Cohort 1 and 365 in Cohort 2).

As requester, the sponsor provided the following detailed explanations:

The discrepancy between the number of patients in the Maintenance ITT population as shown in C13007 FESA Figure 3-1 and the number of patients who achieved CDAI-70 response as shown in Table 14.3.1.32A and Table 39.31.1.1 (sequence 0027), is attributable to differences in patient classification by clinical sites for the purpose of randomization at Week 6, and the Applicant for the purpose of analysis, respectively. Recall that only patients randomized (Cohort 1) or assigned (Cohort 2) to vedolizumab for induction treatment and who were in CDAI-70 clinical response at Week 6, as determined at the clinical sites, were then randomized to the Maintenance Phase. Placebo induction patients continued on placebo regardless of response, and vedolizumab non-responders were continued on vedolizumab. Also, randomization into Maintenance ITT was determined by clinical sites based on patients' Week 6 CDAI scores in the interactive voice response system. On the other hand, CDAI-70 clinical response status was determined by the Applicant based on Week 6 CDAI scores calculated from data in the clinical database.

:

As shown in C13007 FESA Figure 3-1, the clinical sites determined that a total of 461 vedolizumab patients (96 patients from Cohort 1 and 365 patients from Cohort 2), were categorized as responders and randomized these patients into the Maintenance ITT. Following a review by the Applicant, it was determined that 51 patients categorized as responders by the clinical sites, and therefore randomized into the Maintenance ITT, were actually non-responders. These 51 patients (8 patients from Cohort 1 and 43 patients from Cohort 2) were excluded from the number of vedolizumab patients who were categorized as responders at Week 6 for the purpose of analysis for CDAI-70 clinical response.

Conversely, there were 44 vedolizumab-treated patients (11 patients from Cohort 1, and 33 patients from Cohort 2) who were categorized by the clinical sites as non-responders who the Applicant, upon later review, determined to be CDAI-70 responders. These 44 patients were included in the number of vedolizumab patients who were categorized as responders at Week 6 for the purpose of CDAI-70 response analyses. These 2 discrepancies resulted in a difference of 7 vedolizumab-treated patients randomized into the Maintenance ITT (a total of 461 patients) and the total number of vedolizumab patients categorized as Week 6 responders in the CDAI-70 response analyses (99 patients in Cohort 1 and 355 patients in Cohort 2 for a total of 454 patients). Please refer to Table 1.a.

Table 1.a Treatment Group Calculations – Study C13007 FESA Figure 3-1, Study C13007 CSR Table 14.3.1.32A, and Response to Agency Table 39.31.1.1

		Incorrectly Randomized		
Cohort	Week 6 Clinical Response [a]	Randomized as Responders but Classified by Applicant as Nonresponders in Post Hoc Exploratory Analysis of CDAI-70 Endpoint [b]	Randomized as Nonresponders but Analyzed as Responders [c]	Randomized to Maintenance ITT Population [d]
Cohort 1	99	8	11	96
Cohort 2	355	43	33	365
Total	454	51	44	461

Source: Study C13007 CSR Table 14.3.1.32A, and C13007FESA Figure 3-1; 21 October 2014 Response to Agency Table 39.31.1.1 (sequence 0027).

Abbreviations: CSR = clinical study report; ITT = intent-to-treat.

a Study C13007 CSR, Table 14.3.1.32A, and 21 October 2014 Response to Agency Table 39.31.1.1.

b Add to Study C13007 CSR, Table 14.3.1.32A, and 21 October 2014 Response to Agency Table 39.31.1.1.

c Subtract from Study C13007 CSR, Table 14.3.1.32A, and 21 October 2014 Response to Agency Table 39.31.1.1.

d Figure 3-1.

3.1.1.3.2.2 Cohort 2

In the Induction Phase, a total of 220 vedolizumab patients were enrolled into Cohort 1; a total of 747 additional patients were enrolled into Cohort 2. Among the 220 vedolizumab patients in Cohort 1, 96 patients (43.6%) were Week 6 responders. Among 747 additional patients in Cohort 2, 365 patients (48.8%) were Week 6 responders.

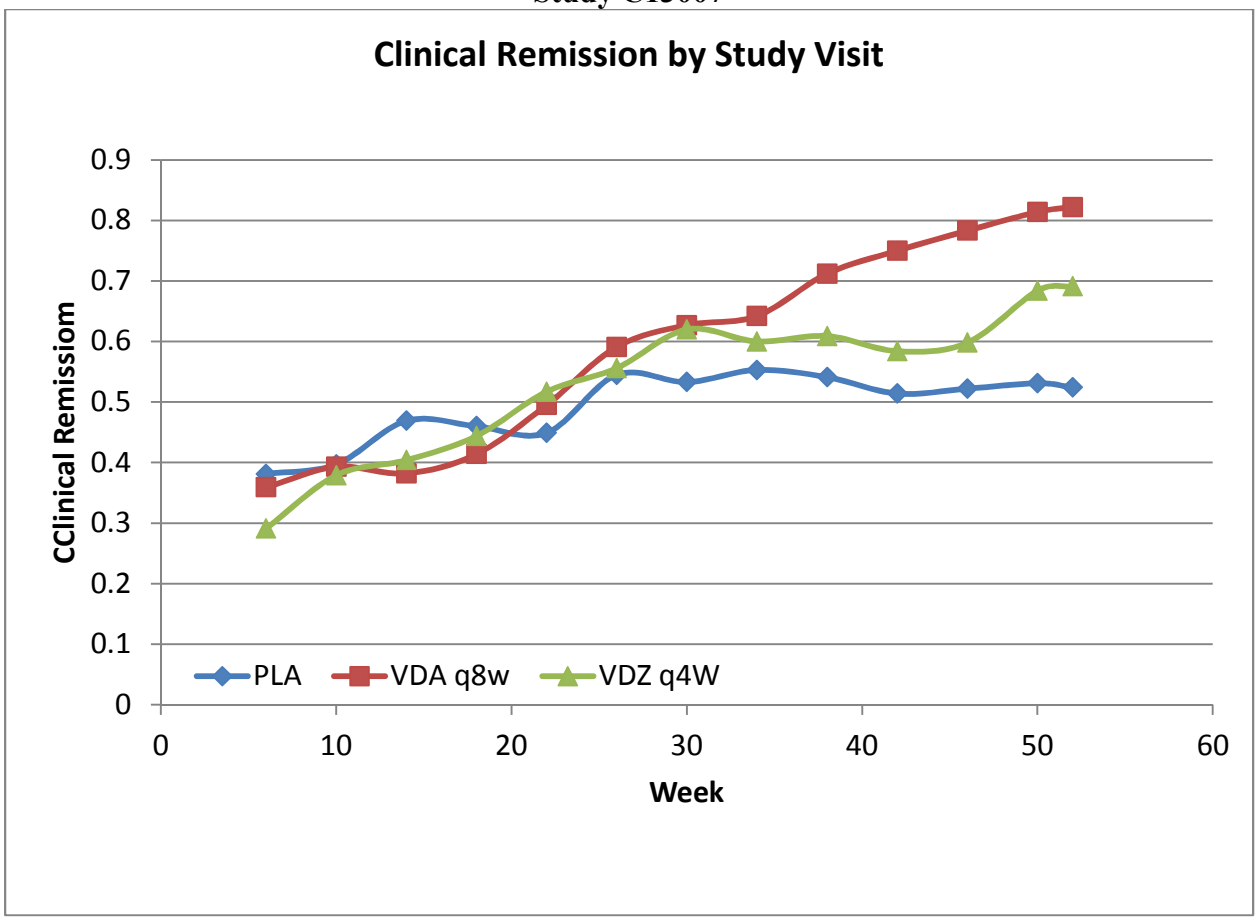
The results for treatment comparisons of overall analysis might be driven by Cohort 2. However, as stated earlier, the cohorts had some notable differences of patient population, which may be the cause of the observed treatment effect difference between the cohorts.

3.1.1.3.2.3 Clinical Remission by Study Visit (Observed)

Per the FDA's request, the sponsor provide summary of the proportion of patients who were observed in clinical remission at all of the assessment time points from Week 6 to Week 52 with no imputation.

This reviewer plotted curves of clinical remission by treatment. Plot is given in the figure below.

Figure 6 Clinical Remission by Study Visit
Study C13007



Complied from Table 39.31.6.1 by this reviewer.

As seen from the figure above, the curves of clinical remission for vedolizumab were notably separated from that of the placebo starting at Week 42.

3.1.1.3.2.4 Hochberg and Sequential Testing Procedure

For the testing of the Maintenance Study's primary and secondary endpoints, the applicant planned to use a Hochberg and sequential testing procedure in order to maintain the overall Type I error rate of 0.05.

This reviewer provided the following comments in the Statistical Review and Evaluation for applicant's IND 9-125 submission S/N 0411 dated January 9, 2012.

> However, the Hochberg procedure is generally not recommended for sequencing testing. It is not assumption free. Furthermore, it is known to provide overall α-control for independent and for certain types of positively correlated endpoints. But its properties for other types of dependent endpoints are not fully known.

We recommend you use a Bonferroni based gatekeeping procedures to test all endpoints in the primary endpoint family and proceed to the secondary family of endpoints only if there has been statistical success in the primary family. When used as a gatekeeping strategy to test the primary family endpoints, the Bonferroni method has an important property of preserving some alpha for testing the secondary endpoint family when at least one of the endpoints in the primary family is statistical significant. The endpoint-specific alpha from each test that successfully rejects the null hypothesis is summed and becomes the alpha available to the secondary endpoint family.

3.1.1.3.2.5 Sensitivity Analyses

The Summary of results from sensitivity analyses for clinical remission at Week 52 are given below.

Table 28 Sensitivity Analyses – Clinical Remission at Week 52
Study C13007

Analysis	Placebo	VDZ q8w	Difference	P-value
Primary	33/153 (21.6%)	60/154 (39.0%)	17.4%	0.0007
Observed Case	33/63 (52.4%)	59/72 (81.9%)	30.5%	<0.0001
Per Protocol	33/147 (22.4%)	56/149 (37.6%)	15.0%	0.0042
LOCF	45/153 (29.4%)	67/154 (43.5)	14,1%	0.0090
Multiple Imputation	66/153 (43.1%)	119/154 (77.3%)	34.2%	<0.0001

Analysis	Placebo	VDZ q4w	Difference	P-value
Primary	33/153 (21.6%)	56/154 (36.4%)	14.7%	0.0042
Observed Case	33/63 (52.4%)	56/81 (69.1%)	17.2%	0.0350
Per Protocol	33/147 (22.4%)	55144 (38.2%)	15.9%	0.0029
LOCF	45/153 (29.4%)	71/154 (46.1)	16.6%	0.0023
Multiple Imputation	66/153 (43.1%)	103/154 (66.9%)	23.7%	<0.0001

Compiled from Tables 57, 14.3.1.2 BM, 39.5.4.1 D, 39.5.5.1 D,
P-values were based on the CMH chi-square test was performed with 2 stratification 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to TNFα antagonists and/or concomitant immunomodulator use (yes/no); 3) enrollment in cohort 1 or cohort 2 in the induction phase.

As seen from Table above, due to the differential of the number of patient with missing data at Week 52, results from the multiple imputation mays tend to over-estimate the treatment effect because the assumption that missing at random might not be true.

Furthermore, at Week 52, with 58% data missing for placebo, 53% data missing for vedolizumab Q8W and 47% data missing for vedolizumab Q4W, the observed treatment effect at Week 52 might not be reliable.

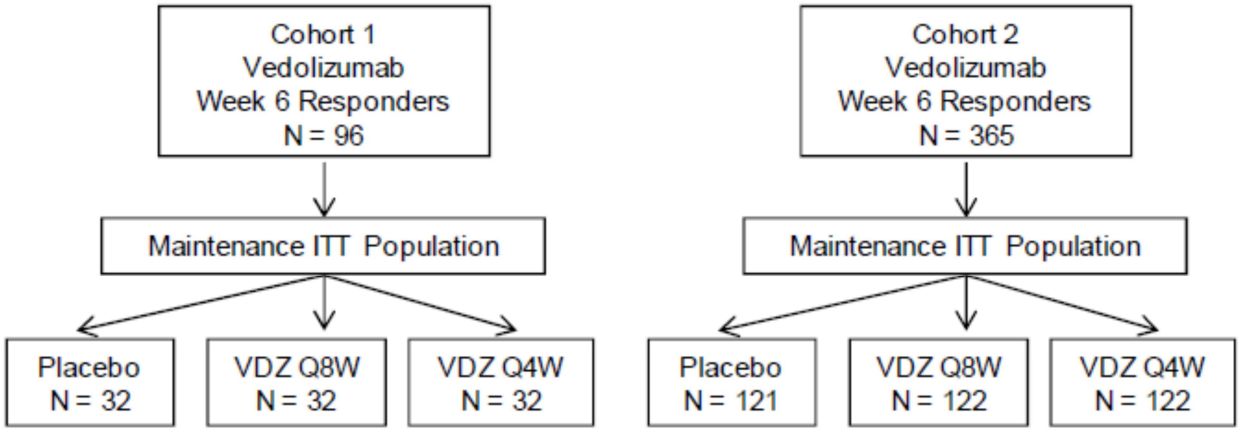
3.1.1.3.2.6 Analysis by Induction Cohort

To assess if the primary and secondary efficacy endpoints at Week 52 were affected by the patient's Induction Phase cohort, additional analyses were requested during a post-phase 3 Type C meeting held on July 24, 2012.

A total of 96 of 220 (44%) patients from Cohort 1 and 365 of 747 (49%) patients from Cohort 2 were randomized to treatment in the Maintenance Study ITT Population.

The figure below displays the distribution of patients from each cohort randomized to the Maintenance Study ITT Population.

Figure 7 Overview of Patients in Cohort 1 and Cohort 2



As seen from the figure above, as patients were randomized to treatment groups based on their Induction Phase cohort and response to therapy, the numbers of patients within each of the treatment groups presented by cohort was balanced.

However, a majority of patients (79%) who were randomized into the Maintenance Phase were from Cohort 2. As discussed earlier, compared to Cohort 1, Cohort 2 had a greater proportion of patients who had prior TNFα antagonist use (68%) and failure (63%), with most patients having shown inadequate response (primary failure: 47%) or loss of response (secondary failure: 40%). Cohort 2 had more patients enrolling at sites in Western/Northern Europe and fewer patients entering at sites in Asia/Australia/Africa and Eastern Europe than was observed for Cohort 1.

The applicant also stated that an imbalance across the treatment groups in the proportion of patients who had achieved clinical remission at Week 6 was observed because randomization at Week 6 was not stratified by remission status. Clinical remission at Week 6 was achieved by 27.9% of patients in the vedolizumab Q4W group and 33.8% of patients in the vedolizumab Q8W group, compared with 36.6% of patients in the placebo group.

This imbalance may have had an impact on the analyses of the clinical remission-based endpoints, especially for the vedolizumab Q4W group versus the placebo group.

3.1.1.3.2.7 Analysis for Primary Endpoint of the Maintenance Study by Cohort.

The number and proportion of patients in the Maintenance Study ITT Population who achieved clinical remission at Week 52 are summarized by induction phase cohort in the table below.

Table 29 Clinical Remission at Week 52 by Induction Phase Cohort
Maintenance Study ITT Population
Study C13007

	Cohort 1			Cohort 2		
Clinical Remission,[a] n (%)	**PLA N = 32**	**VDZ Q8W N = 32**	**VDZ Q4W N = 32**	**PLA N = 121**	**VDZ Q8W N = 122**	**VDZ Q4W N = 122**
Number (%) achieving clinical remission	6 (18.8)	8 (25.0)	13 (40.6)	27 (22.3)	52 (42.6)	43 (35.2)
95% CI	(5.2, 32.3)	(10.0, 40.0)	(23.6, 57.6)	(14.9, 29.7)	(33.8, 51.4)	(26.8, 43.7)
Difference from placebo[b]		6.3	22.0		20.3	12.8
95% CI for difference from placebo		(-12.9, 25.4)	(-0.3, 44.3)		(8.6, 32.0)	(1.5, 24.0)
P-value for difference from placebo[c]		0.5220	0.0535		0.0007	0.0262
Relative risk[d]		1.3	2.1		1.9	1.6
95% CI for relative risk		(0.5, 3.3)	(0.9, 5.0)		(1.3, 2.8)	(1.0, 2.4)

Copied from Table 3-1, page 12, Crohn's Disease Supplemental Efficacy Analyses Report (C13007 FESA).

As seen from the table above, when compared to placebo, treatment differences were observed for the vedolizumab Q4W group (12.8%, p = 0.0262) and the vedolizumab Q8W group (20.3%, p = 0.0007) from Cohort 2, and trends favoring the vedolizumab Q8W and Q4W groups from Cohort 1 were observed.

The results by cohort were inconsistent between dosing regimens (Q4W vs. Q8W) between Cohorts. The results for Q8W against placebo of overall analysis were driven by that of Cohort 2.

In the applicant's response to this reviewer's Information Request (IR) dated August 18, 2013, the applicant combined the Q4W and Q8W maintenance treatment arms to increase the power for treatment comparison. The table below presents clinical remission at Week 52 by Cohort 1 and Cohort 2.

Table 30 Clinical Remission at Week 52 by Induction Phase Cohort
Maintenance Study ITT Population
Study C13007

	Cohort 1		Cohort 2	
Clinical Remission	**Placebo (N = 32)**	**Vedolizumab (N = 64)**	**Placebo (N = 121)**	**Vedolizumab (N = 244)**
Number (%) achieving parameter	6 (19.8)	21 (33)	27 (22.3)	95 (39.0)
Difference from Placebo (%)		13.2		16.7
Relative risk		1.67		1.75

Copied from Table 11a, page 21, Response to Agency Questions (Questions Received August 19, 2013.

As seen from the table above, the effect size (difference between vedolizumab and placebo) for Cohort 2 was greater than that for Cohort 1.

3.1.1.3.2.8. Analysis for Secondary Endpoint of the Maintenance Study by Cohort

3.1.1.3.2.8.1 Enhanced Clinical Response at Week 52 by Induction Phase Cohort

The number and proportion of patients in the Maintenance Study ITT Population who achieved enhanced clinical response at Week 52 are summarized by Induction Phase cohort in the table below.

Table 31 Enhanced Clinical Response at Week 52 by Induction Phase Cohort
Maintenance Study ITT Population
Study C13007

	Cohort 1			Cohort 2		
Enhanced Clinical Response,[a] n (%)	**PLA N = 32**	**VDZ Q8W N = 32**	**VDZ Q4W N = 32**	**PLA N = 121**	**VDZ Q8W N = 122**	**VDZ Q4W N = 122**
Number (%) achieving enhanced clinical response	10 (31.3)	13 (40.6)	15 (46.9)	36 (29.8)	54 (44.3)	55 (45.1)
95% CI	(15.2, 47.3)	(23.6, 57.6)	(29.6, 64.2)	(21.6, 37.9)	(35.4, 53.1)	(36.3, 53.9)
Difference from placebo[b]		9.4	15.9		14.5	15.1
95% CI for difference from placebo		(-12.6, 31.4)	(-7.9, 39.6)		(2.4, 26.6)	(3.1, 27.2)
P-value for difference from placebo[c]		0.4033	0.1904		0.0188	0.0137
Relative risk[d]		1.3	1.5		1.5	1.5
95% CI for relative risk		(0.7, 2.4)	(0.8, 2.9)		(1.1, 2.1)	(1.1, 2.1)

Copied from Table 3-2, page 13, Crohn's Disease Supplemental Efficacy Analyses Report (C13007 FESA).

As seen from the table above, when compared to placebo, treatment differences were observed for the vedolizumab Q4W group (15.1%, p = 0.0137) and the vedolizumab Q8W group (14.5%, p = 0.0188) from Cohort 2, and trends favoring the vedolizumab Q8W and Q4W groups from Cohort 1 were observed.

The results by cohort were inconsistent between cohorts for VDZq8w. The results for Q8W against placebo of overall analysis were driven by that of Cohort 2.

In the applicant's response to this reviewer's Information Request (IR) dated August 18, 2013, the applicant combined the Q4W and Q8W maintenance treatment arms to increase the power for treatment comparison. The table below presents enhanced clinical response at Week 52 by Cohort 1 and Cohort 2.

Table 32 Enhanced Clinical Response at Week 52 by Induction Phase Cohort Maintenance Study ITT Population Study C13007

	Cohort 1		Cohort 2	
Enhanced Clinical Response	**Placebo (N = 32)**	**Vedolizumab (N = 64)**	**Placebo (N = 121)**	**Vedolizumab (N = 244)**
Number (%) achieving parameter	10 (31.3)	28 (43.8)	36 (29.8)	109 (44.7)
Difference from Placebo (%)	12.5		14.9	
Relative risk	1.4		1.5	

Copied from Table 11b, page 22, Response to Agency Questions (Questions Received August 19, 2013.

As seen from the table above, the effect size (difference between vedolizumab and placebo) for Cohort 2 was greater than that for Cohort 1.

3.1.1.3.2.8.2 Corticosteroid –Free Clinical Remission at Week 52 by Induction Phase Cohort

The number and proportion of patients in the Maintenance Study ITT Population who achieved corticosteroid-free clinical remission at Week 52 are summarized by Induction Phase cohort in the table below.

Table 33 Corticosteroid-Free Clinical Remission at Week 52 by Induction Phase Cohort Study C13007

Corticosteroid-Free Clinical Remission,[a] n (%)	Cohort 1			Cohort 2		
	PLA N = 18	VDZ Q8W N = 18	VDZ Q4W N = 17	PLA N = 64	VDZ Q8W N = 64	VDZ Q4W N = 63
Number (%) achieving corticosteroid-free clinical remission	3 (16.7)	2 (11.1)	7 (41.2)	10 (15.6)	24 (37.5)	16 (25.4)
95% CI	(0.0, 33.9)	(0.0, 25.6)	(17.8, 64.6)	(6.7, 24.5)	(25.6, 49.4)	(14.6, 36.1)
Difference from Placebo[b]		-5.6	24.5		21.9	9.7
95% CI for difference from placebo		(-27.8, 16.7)	(-5.4, 54.5)		(6.7, 37.1)	(-4.1, 23.5)
P-value for difference from placebo[c]		0.6246	0.1086		0.0048	0.1685
Relative risk [d]		0.7	2.5		2.4	1.6
95% CI for relative risk		(0.1, 3.5)	(0.7, 8.1)		(1.3, 4.6)	(0.8, 3.3)

Copied from Table 3-3, page 15, Crohn's Disease Supplemental Efficacy Analyses Report (C13007 FESA).

As seen from the table above, among the patients who received treatment in Cohort 2 during the Induction Phase, greater proportions treated with vedolizumab in the Q8W and Q4W treatment groups achieved corticosteroid-free clinical remission at Week 52 compared with those who received placebo. A treatment difference was observed for the vedolizumab Q8W group (21.9%, p = 0.0048); a trend was observed for the vedolizumab Q4W group. Among the patients who received treatment in Cohort 1 during the Induction Phase, a trend favoring the vedolizumab Q4W group was observed.

The results by cohort were inconsistent between dosing regimens (Q4W vs. Q8W) between Cohorts. The results for Q8W against placebo of overall analysis were driven by that of Cohort 2.

In the applicant's response to this reviewer's Information Request (IR) dated August 18, 2013, the applicant combined the Q4W and Q8W maintenance treatment arms to increase the power for treatment comparison. The table below presents corticosteroid-free clinical remission at Week 52 by Cohort 1 and Cohort 2.

Table 34 Corticosteroid Free Clinical Remission at Week 52 by Induction Phase Cohort Maintenance Study ITT Population Study C13007

Corticosteroid-free Clinical Remission	Cohort 1		Cohort 2	
	Placebo (N = 18)	Vedolizumab (N = 35)	Placebo (N = 64)	Vedolizumab (N = 127)
Number (%) achieving parameter	3 (16.7)	9 (25.7)	10 (15.6)	40 (31.5)
Difference from Placebo (%)	9.0%		15.9	
Relative risk	1.5		2.0	

Copied from Table 11b, page 22, Response to Agency Questions (Questions Received August 19, 2013.

As seen from the table above, the effect size (difference between vedolizumab and placebo) for Cohort 2 was greater than that for Cohort 1.

3.1.1.3.2.8.3 Durable Clinical Remission at Week 52 by Induction Phase Cohort

The number and proportion of patients in the Maintenance Study ITT Population who achieved durable clinical remission at Week 52 (defined as CDAI score ≤ 150 points at ≥ 80% of study visits including final visit [Week 52]) are summarized by Induction Phase cohort in the table below.

Table 35 Durable Clinical Remission at Week 52 by Induction Phase Cohort
Maintenance Study ITT Population
Study C13007

	Cohort 1			Cohort 2		
Durable Clinical Remission,[a] n (%)	**PLA N = 32**	**VDZ Q8W N = 32**	**VDZ Q4W N = 32**	**PLA N = 121**	**VDZ Q8W N = 122**	**VDZ Q4W N = 122**
Number (%) achieving durable clinical remission	3 (9.4)	4 (12.5)	5 (15.6)	19 (15.7)	29 (23.8)	20 (16.4)
95% CI	(0.0, 19.5)	(1.0, 24.0)	(3.0, 28.2)	(9.2, 22.2)	(16.2, 31.3)	(9.8, 23.0)
Difference from placebo[b]		3.7	8.1		8.0	0.6
95% CI for difference from placebo		(-13.5, 20.9)	(-10.4, 26.6)		(-1.9, 18.0)	(-8.6, 9.8)
P-value for difference from placebo[c]		0.6731	0.3899		0.1137	0.8996
Relative risk [d]		1.3	1.7		1.5	1.0
95% CI for relative risk		(0.3, 5.1)	(0.4, 6.7)		(0.9, 2.5)	(0.6, 1.8)

Copied from Table 3-4, page 16, Crohn's Disease Supplemental Efficacy Analyses Report (C13007 FESA).

As seen from the table above, among the patients who received treatment in Cohort 2 during the Induction Phase, a trend favoring the vedolizumab Q8W group was observed for the proportion of patients who achieved durable clinical remission at Week 52 compared with those who received placebo.

The results by cohort were inconsistent between dosing regimens (Q4W vs. Q8W) between Cohorts.

In the applicant's response to this reviewer's Information Request (IR) dated August 18, 2013, the applicant combined the Q4W and Q8W maintenance treatment arms to increase the power for treatment comparison. The table below presents durable clinical remission at Week 52 by Cohort 1 and Cohort 2.

Table 36 Durable Clinical Remission at Week 52 by Induction Phase Cohort
Maintenance Study ITT Population
Study C13007

	Cohort 1		Cohort 2	
Corticosteroid-free Clinical Remission	**Placebo (N = 32)**	**Vedolizumab (N = 64)**	**Placebo (N = 121)**	**Vedolizumab (N = 244)**
Number (%) achieving parameter	3 (9.4%)	9 (14.0%)	19 (15.7%)	49 (20.0%)
Difference from Placebo (%)		4.6		4.3
Relative risk		1.5		1.3

Copied from Table 11b, page 23, Response to Agency Questions (Questions Received August 19, 2013.

As seen from the table above, the effect size (difference between vedolizumab and placebo) is similar between Cohort 1 and Cohort 2.

3.1.1.3.2.9. Alternative Definition of Clinical Remission

Per FDA's requested, the summary tables below for the Induction Study ITT Population are provided by the applicant for analyses using the following alternative definition of clinical remission regardless of CDAI score :

- Total number of liquid/very soft stools of ≤ 10 per day in the relevant week; and
- Abdominal pain rated as 0 or 1 for each day in the relevant week.

Table 37 Clinical Remission at Week 52 Based on Alternative Definition from FDA
Intent-to-Treat Population
Study C13007

Clinical remission	Placebo (n=153)	VDZ Q8 wks (n=154)	VDZ Q4 wks (n=154)
N (%) achieve clinical remission at week 6	19 (12.4)	37 (24.0)	31 (20.1)
95% CI	(7.2,17.6)	(17.3,30.8)	(13.8, 26.5)
Difference from placebo		12.0	7.7
95% CI for difference from placebo		(3.2, 20.8)	(-0.5,15.9)
p-value for difference from placebo		0.0078	0.0671

Complied from Table 48.1.1.2A by this reviewer.
P-values were based on the CMH chi-square test was performed with 2 stratification 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to TNFα antagonists and/or concomitant immunomodulator use (yes/no); 3) enrollment in cohort 1 or cohort 2 in the maintenance phase.

A seen from the table above, based on the alternative definition from FDA, vedolizumab-treated patients achieved greater clinical remission at Week 52 for vedolizumab Q8W group compared to the placebo group.

Per FDA's request, the summary tables below for the Induction Study ITT Population are provided by the applicant for analyses using the above alternative definitions of clinical remission regardless of CDAI score.

3.1.2 Study C13011

3.1.2.1 Study Design

This study was a phase 3, multinational, randomized, double-blind, placebo-controlled trial conducted to evaluate the efficacy and safety of vedolizumab for the induction of clinical response and remission in patients with moderately to severely active CD. Of the total patients enrolled, approximately 75% had previously failed TNFα antagonist therapy and approximately 25% had been naïve to TNFα antagonist therapy.

The inclusion and exclusion criteria were similar to those for Study C13007 with exception of the CDAI score. In this study, moderate to severely active CD was determined by a CDAI of 220 to 400 instead of 200 to 480 used in Study C13007.

After a 21-day Screening period, patients were randomized in a 1:1 ratio to receive either 300 mg vedolizumab or placebo at Weeks 0, 2, and 6. The enrollment of patients was monitored by the IVRS to ensure that approximately 75% of the overall population had previously failed TNFα antagonist therapy and approximately 25% were naïve to TNFα antagonist therapy.

The randomization to treatment assignment was stratified by the presence or absence of each of the following, as entered into IVRS at screening:

- Previous failure of TNFα antagonist therapy or naïve to TNFα antagonist therapy
- Concomitant use of oral corticosteroids
- Concomitant use of immunomodulators (6-MP, azathioprine, or methotrexate)

After completing the Week 10 assessments, patients were eligible to enroll in Study C13008 (open-label, long-term safety study) if the study drug was well-tolerated, and no major surgical intervention for CD occurred or was required.

After the 21-day Screening period, enrolled patients were to complete the 10-week induction period, at which time they may have been eligible to receive vedolizumab treatment by enrolling in Study C13008 (an open-label, long-term safety study). Patients, who did not enroll in Study C13008, whether they completed Week 10 or withdrew early from the study, completed the Final Safety visit (Week 22, or 16 weeks after the last dose of study drug). In addition, after the completion of the study, all patients who did not enroll in Study C13008 were to participate in a two-year follow-up survey.

The primary objective was:

- To determine the effect of vedolizumab induction treatment on clinical remission at Week 6 in the subgroup of patients defined as having failed tumor necrosis factor

alpha (TNFα) antagonist therapy (TNFα antagonist failure subpopulation)

The secondary objectives were:

- To determine the effect of vedolizumab induction treatment on clinical remission at Week 6 in the entire study population;
- To determine the effect of vedolizumab induction treatment on clinical remission at Week 10 in the TNFα antagonist failure subpopulation and in the entire study population;
- To determine the effect of vedolizumab induction treatment on sustained clinical remission (i.e., clinical remission at both Week 6 and Week 10) in the TNFα antagonist failure subpopulation and in the entire study population;
- To determine the effect of vedolizumab induction treatment on enhanced clinical response at Week 6 in the TNFα antagonist failure subpopulation.

The safety objectives were:

- To determine the safety profile of vedolizumab induction treatment in the entire study population;
- To determine the safety profile of vedolizumab induction treatment in the TNFα antagonist failure subpopulation.

There were six planned patient populations in this study for analyses: the Overall Intent-to-Treat (ITT) Population, the Overall Modified ITT Population, the Overall Per-Protocol Population, the Overall Completers (Observed Case) Population, the Overall Safety Population, and the Overall PK-Evaluable Population. Within each of the six populations, the TNFα antagonist failure subpopulation was defined as all patients who met the TNFα antagonist failure criterion collected in the IVRS at the time of randomization.

In accordance with an ITT approach, the Overall ITT Population consisted of all randomized patients who received any amount of blinded study drug. The TNFα Antagonist Failure ITT Subpopulation was a subset of the Overall ITT Population in which all patients met the TNFα antagonist failure criterion.

This population was used for proportion-based endpoints, such as clinical remission or enhanced clinical response. Patients in this population were to be analyzed according to the treatment to which they were randomized, regardless of any errors in dosing.

The Overall Modified ITT Population consisted of all randomized patients who received any amount of blinded study drug and had a baseline and at least one post-randomization measurement for the endpoint under consideration. The TNFα Antagonist Failure Modified ITT Subpopulation was a subset of the Overall Modified ITT Population in which all patients met the TNFα antagonist failure criterion.

This population was used for change from baseline analyses of CDAI scores. Patients in this population were to be analyzed according to the treatment to which they were randomized, regardless of any errors of dosing

The Overall Per-Protocol Population was a subset of the Overall ITT Population. All criteria for excluding patients from the Overall Per-Protocol Population data set were decided prior to the unblinding of the study. The TNFα Antagonist Failure Per-Protocol Subpopulation was a subset of the Overall Per-Protocol Population in which all patients met the TNFα antagonist failure criterion.

Efficacy assessments throughout the study were based on CDAI scores. CDAI scores were also obtained at screening, Week 2, and at any unscheduled visit(s) due to disease exacerbation. On all dosing days, CDAI scores were determined based on components obtained prior to dosing. The applicant calculated all CDAI scores utilizing the sum of the most recently available eDiary CDAI score components.

Analyses using the Overall Per-Protocol Population were provided as sensitivity analyses for the primary and key secondary endpoints.

Patients were included in the Overall Per-Protocol Population, if they met the following criteria according to the specified hierarchy:

- Confirmed diagnosis of CD of at least three months' duration and an enrolling CDAI score between 210 and 410 (inclusive)
- Received the correct study medication as assigned
- Completed Week 10 assessments per protocol or met one or more of the following criteria for failure prior to the Week 10 assessments:
  - Received rescue medication for treatment of CD prior to Week 10
  - Had major surgery for CD
  - Had a drug-related AE that led to discontinuation
- Received all 3 doses of study drug as assigned or met 1 or more of the criteria for failure
- Did not receive concomitant corticosteroids or other potentially effective medications (except as permitted per protocol) for an unrelated comorbid condition (e.g., prednisone for idiopathic thrombocytopenic purpura)
- Completed the Week 10 visit and had a valid Week 10 CDAI assessment
- Did not have the treatment assignment unblinded by the investigator

The Overall Completers (Observed Case) Population was defined as all randomized patients who received any amount of blinded study drug who had assessments for the endpoint under consideration (e.g., CDAI score) at baseline and Weeks 6 and 10. The TNFα Antagonist Failure Completers (Observed Case) Subpopulation was a subset of the Overall Completers (Observed Case) Population in which all patients met the TNFα antagonist failure criterion.

The Overall Safety Population was defined as all patients who received any amount of study drug. The TNFα Antagonist Failure Safety Subpopulation was a subset of the Overall Safety Population in which all patients met the TNFα antagonist failure criterion.

The Overall Safety Population was used for all safety analyses; patients in this population were analyzed according to the treatment they received.

All patients who prematurely discontinued for any reason were to be considered as not achieving clinical remission for the primary efficacy analysis.

Baseline CDAI scores (obtained at Week 0) were used for the comparison with Week 6 and Week 10 scores to determine response and remission as defined below:

- Clinical Remission: CDAI score ≤ 150 points
- Sustained Clinical Remission: Clinical remission at both Week 6 and Week 10
- Enhanced Clinical Response: A ≥ 100-point decrease in CDAI score from baseline (Week 0)

The primary endpoint was:

- Proportion of patients in clinical remission at Week 6 in the TNFα antagonist failure Subpopulation.

The secondary endpoints were:

- Proportion of patients in clinical remission at Week 6 in the entire study population;
- Proportions of patients in clinical remission at Week 10 in the TNFα antagonist failure subpopulation and in the entire study population;
- Proportions of patients with sustained clinical remission (i.e., clinical remission at both Week 6 and Week 10) in the TNFα antagonist failure subpopulation and in the entire study population;
- Proportion of patients with enhanced clinical response at Week 6 in the TNFα antagonist failure subpopulation.

3.1.2.2 Pre-specified Analyses

Demographic and other baseline characteristics were summarized by treatment group and overall, using the Overall ITT Population and the TNFα Antagonist Failure ITT Subpopulation. Age was summarized as a continuous variable, and by grouping patient age categories (< 65, ≥ 65 and < 35, ≥ 35 years). Body mass index (BMI) was summarized as a continuous variable.

Selected demographic and CD-related baseline characteristics were compared between the treatment groups using unadjusted p-values based on the chi-square test or Fisher's exact test for categorical variables and the Kruskal-Wallis or Wilcoxon rank sum test for continuous variables.

The proportion-based endpoints, such as clinical remission, sustained clinical remission, and enhanced clinical response, were tested using the Cochran-Mantel-Haenszel (CMH) chi-square test at a 5% significance level with stratification according to concomitant use of oral corticosteroids and concomitant use of immunomodulators (6-MP, azathioprine, or methotrexate) for the TNFα antagonist failure subpopulation, or with stratification according to previous failure of TNFα antagonist therapy, concomitant use of oral corticosteroids, and concomitant use of immunomodulators (6-MP, azathioprine, or methotrexate) for the overall population. The CMH

chi-square p-value and the risk difference, along with its 95% two-sided confidence interval (CI), were provided.

To maintain the overall Type I error rate at 5%, the secondary endpoint analyses were performed sequentially. Specifically, clinical remission at Week 6 for the overall population was to be tested only if the primary endpoint comparison was significant; the set of analyses for clinical remission at Week 10 for the TNFα antagonist failure subpopulation and the overall population was to be tested only if the endpoint of clinical remission at Week 6 for the overall population was significant. The remaining secondary endpoints were to be tested only if the comparison for the previous secondary endpoint was significant.

In addition, the Hochberg method was applied to each secondary endpoint pair in order to control the overall Type 1 error rate at a 5% significance level. Specifically, for the two comparisons (TNFα antagonist failure subpopulation and the overall population) of clinical remission at Week 10 and for the two comparisons (TNFα antagonist failure subpopulation and the overall population) of sustained clinical remission, the Hochberg method was applied to each endpoint pair to control the overall Type I error rate across the two different analysis populations. If both p-values for each of the analysis populations within each set were ≤ 0.05, both the TNFα antagonist failure subpopulation and the overall population were declared significant. If one of the p-values within the set was > 0.05, the other p-value was to be tested at the 0.025 level and declared significant only if the p-value was ≤ 0.025. The sequential testing procedure for testing the next secondary endpoint was then to be used as described above. If neither population within the endpoint set was declared significant for the related secondary endpoint, no further formal statistical testing of subsequent endpoints was to be conducted.

For the assessment of the primary efficacy endpoint, the CMH chi-square test was used. To assess the robustness of the primary efficacy analysis, the following additional analyses were performed for the primary endpoint.

- CMH chi-square test using the TNFα Antagonist Failure Completers (Observed Case) Subpopulation (i.e., TNFα Antagonist Failure ITT Subpopulation patients who had baseline measurements, as well as Week 6 and Week 10 post-baseline CDAI score assessments);
- CMH chi-square test using the TNFα Antagonist Failure Per-Protocol Subpopulation.

Approximately 396 patients were planned to be enrolled in this study from approximately 150 sites worldwide. Of those, approximately 296 patients were to have previously failed (i.e., had an inadequate response to, loss of response to, or intolerance of) TNFα antagonist therapy and up to approximately 100 patients were to have no previous exposure to TNFα antagonists. Enrollment was defined as the point in time at which the patient began the first dose of study drug. Final enrollment was 416 patients, 315 (76%) of whom had previously failed TNFα antagonist therapy.

The study was adequately powered for the primary endpoint, as well as for the key secondary endpoints. Power estimates (provided in the table below) for the primary and secondary efficacy

endpoints were based on a total sample size of 396 for the overall study population and 296 for the TNFα antagonist failure subpopulation.

Table 38 Power Estimates for the Primary and Secondary Efficacy Analysis Study C13011

Objective	Endpoint Name	Assumed Response Rates	Sample Size per Group	Power
Primary	Clinical Remission at Week 6 in the TNFα antagonist failure subpopulation	placebo = 5% vedolizumab = 17%	placebo = 148 vedolizumab = 148	91%
Secondary	Clinical Remission at Week 6 in the overall study population	placebo = 10% vedolizumab = 23%	placebo = 198 vedolizumab = 198	93%
	Clinical Remission at Week 10 in the TNFα antagonist failure subpopulation	placebo = 7% vedolizumab = 19%	placebo = 148 vedolizumab = 148	87%
	Clinical Remission at Week 10 in the overall study population	placebo = 13% vedolizumab = 26%	placebo = 198 vedolizumab = 198	90%
	Sustained Clinical Remission at both Week 6 and Week 10 in the TNFα antagonist failure subpopulation	placebo = 4% vedolizumab = 14%	placebo = 148 vedolizumab = 148	85%
	Sustained Clinical Remission at both Week 6 and Week 10 in the overall study population	placebo = 8% vedolizumab = 19%	placebo = 198 vedolizumab = 198	89%
	Enhanced Clinical Response at Week 6 in the TNFα antagonist failure subpopulation	placebo = 21% vedolizumab = 36%	placebo = 148 vedolizumab = 148	81%

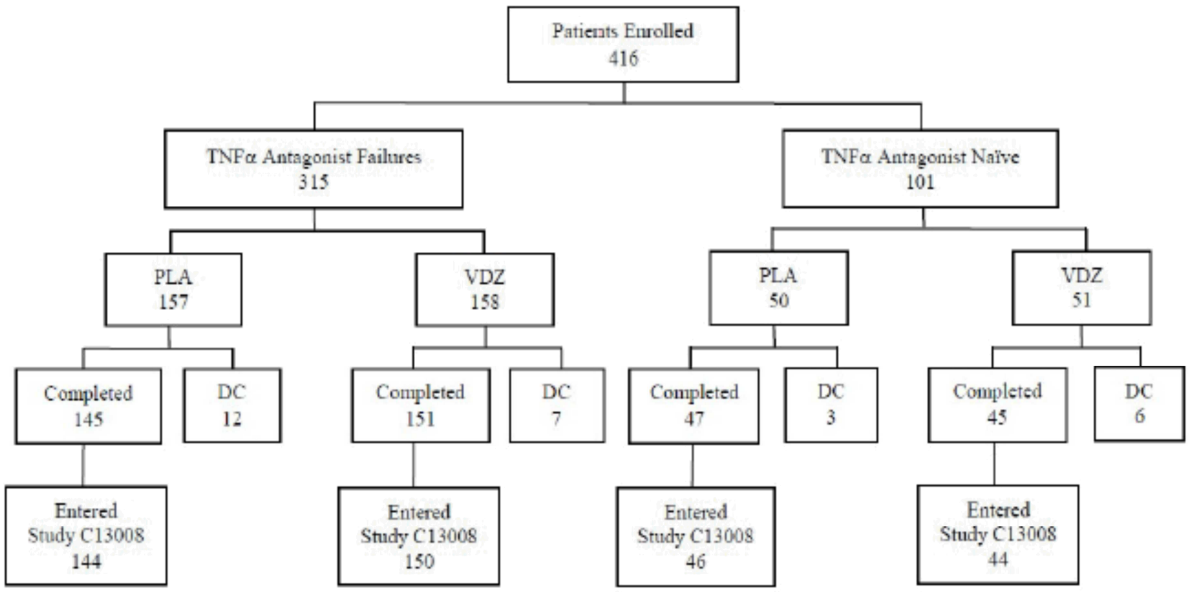
Copied from Table 9-2, page 60 CSR.

3.1.2.3 Applicant's Analyses

A total of 660 patients were screened for enrollment in the study (data obtained from IVRS). Of these, 244 patients failed screening due to the following reasons: did not meet enrollment criteria (209 patients); withdrew consent (11 patients); SAE (5 patients); protocol violation (1 patient); and other or unknown reason (18 patients). Thus, 416 patients were enrolled in the study and randomized to treatment.

Among the 416 randomized patients, 315 (76%) had previously failed TNFα antagonist therapy and 101 (24%) were naïve to TNFα antagonist therapy per data collected in the IVRS, 311 (75%) patients had previously failed TNFα antagonist therapy and 105 (25%) patients were naïve to TNFα antagonist therapy. A schematic of the study drug assignment and disposition of all study participants is presented in the figure below.

Figure 7 Study Drug Assignment and Disposition of All Patients
Study C13011



Copied from Figure 10-1, page 78 CSR.

3.1.2.3.1 Patient Disposition

Patient disposition is summarized by treatment group for the overall patient population and the TNFα antagonist failure patient subpopulation in the table below.

Table 39 Patient Disposition – Study C13011

	TNFα Antagonist Failure ITT Subpopulation			Overall ITT Population		
	PLA N = 157	VDZ N = 158	Total N = 315	PLA N = 207	VDZ N = 209	Total N = 416
Randomized	157	158	315	207	209	416
Safety Population[a]	157 (100)	158 (100)	315 (100)	207 (100)	209 (100)	416 (100)
ITT Population[b]	157 (100)	158 (100)	315 (100)	207 (100)	209 (100)	416 (100)
Per-Protocol Population[c]	145 (92)	147 (93)	292 (93)	194 (94)	192 (92)	386 (93)
Completed study[d]	145 (92)	151 (96)	296 (94)	192 (93)	196 (94)	388 (93)
Enrolled into C13008	144 (99)	150 (99)	294 (99)	190 (99)	194 (99)	384 (99)
Discontinued (Reason)	12 (8)	7 (4)	19 (6)	15 (7)	13 (6)	28 (7)
Adverse event	6 (4)	2 (1)	8 (3)	8 (4)	4 (2)	12 (3)
Protocol violation(s)	0	1 (< 1)	1 (< 1)	0	1 (< 1)	1 (< 1)
Lack of efficacy	4 (3)	0	4 (1)	5 (2)	1 (< 1)	6 (1)
Withdrawal of consent	2 (1)	3 (2)	5 (2)	2 (< 1)	4 (2)	6 (1)
Lost to follow-up	0	1 (< 1)	1 (< 1)	0	3 (1)	3 (< 1)

Copied from Table 10-1, page 80 CSR.

As seen from the table above, in the overall patient population, 207 patients were randomized to receive placebo and 209 patients were randomized to receive vedolizumab. In both treatment groups, all randomized patients received at least one dose of blinded study drug and were included in the Overall Safety and Overall ITT Populations. The Overall Per-Protocol Population included 94% of placebo-treated patients and 92% of vedolizumab-treated patients. In the Overall ITT Population, similar proportions of patients in both treatment groups completed Week 10 assessments (93% placebo; 94% vedolizumab); with 99% of the completed patients continuing into the long-term safety study (Study C13008). A greater proportion of placebo-treated patients prematurely discontinued from the study due to AEs (4%) than patients who received vedolizumab (2%); no other notable differences were observed between the treatment groups for reasons leading to premature discontinuation.

Among the 416 randomized patients, 76% had previously failed TNFα antagonist therapy. As the primary study objective was to addresses the therapeutic benefit of vedolizumab in patients who had experienced inadequate response, loss of response, or intolerance to other TNFα antagonists, the TNFα antagonist failure subpopulation was defined in which all patients met the TNFα antagonist failure criterion collected in the IVRS at the time of randomization. Each of the overall analysis populations has a corresponding TNFα antagonist failure subpopulation. Throughout the results, data are presented for both the overall population and the TNFα antagonist failure subpopulation.

Among the 315 TNFα antagonists failure patients, 157 received placebo and 158 received Vedolizumab; each of these patients was included in the TNFα Antagonist Failure Safety and ITT Subpopulations. The TNFα Antagonist Failure Per-Protocol Subpopulation included 92% of

placebo-treated patients and 93% of vedolizumab-treated patients. In the TNFα Antagonist Failure ITT Subpopulation, the majority of the patients in both treatment groups completed study (92% placebo; 96% vedolizumab); with 99% of the completed patients continuing into the long-term safety study (Study C13008). As was observed in the Overall ITT Population, a greater proportion of placebo-treated patients prematurely discontinued from the study due to AEs (4%) than patients who received vedolizumab (1%). In addition, 3% of the patients in the placebo group prematurely discontinued due to lack of efficacy, whereas none of the patients in the vedolizumab group discontinued for this reason. No other notable differences were observed between the treatment groups for reasons leading to premature discontinuation.

Inclusion/exclusion criteria deviations are summarized for the Overall ITT Population and the TNFα Antagonist Failure ITT Subpopulation in the table below.

Table 40 Inclusion and Exclusion Criteria Not Met
TNFα Antagonist Failure ITT Subpopulation and Overall ITT Population
Study C13011

	TNFα Antagonist Failure ITT Subpopulation			Overall ITT Population		
Type of Unmet Criteria[a], n (%)	PLA N = 157	VDZ N = 158	Total N = 315	PLA N = 207	VDZ N = 209	Total N = 416
Patients with at least 1 unmet entry criterion	5	10	15	8	14	22
Inclusion criteria						
CDAI of 220 to 400 and a) CRP > 2.87 mg/L, b) at least 3 non-anastomotic ulcerations, or c) fecal calprotectin > 250 µg/g	2 (1)	3 (2)	5 (2)	2 (< 1)	5 (2)	7 (2)
Inadequate or lost response/intolerance of steroids, immunomodulators, and/or TNFα antagonists	1 (< 1)	2 (1)	3 (< 1)	2 (< 1)	4 (2)	6 (1)
Initial steroid dose stable × 4 weeks or 2 weeks for patients tapering steroids	1 (< 1)	2 (1)	3 (< 1)	1 (< 1)	2 (< 1)	3 (< 1)
CD > 3 months duration (histo report) or > 6 months if report not available	0	1 (< 1)	1 (< 1)	0	1 (< 1)	1 (< 1)
Colonoscopy within 12 months for long-standing disease	1 (< 1)	0	1 (< 1)	1 (< 1)	0	1 (< 1)
Gastrointestinal exclusion criteria						
Extensive colonic resection, subtotal or total colectomy	0	1 (< 1)	1 (< 1)	0	1 (< 1)	1 (< 1)
C. difficile infection or other intestinal pathogen within 28 days	0	0	0	1 (< 1)	0	1 (< 1)
Infectious disease exclusion criteria						
Chronic HBV or HCV	0	0	0	1 (< 1)	0	1 (< 1)
General exclusion criteria						
ALT or AST > 3 × ULN	0	1 (< 1)	1 (< 1)	0	1 (< 1)	1 (< 1)

Copied from Table 10-2, page 81 CSR.

As seen from the table above, in the Overall ITT Population, a total of 22 patients (8 patients in placebo; 14 patients in vedolizumab) had at least one unmet entry criterion. In both treatment groups, the most common deviations were failure to meet the inclusion criterion for baseline CDAI score of 220 to 400 with either a CRP level > 2.87 mg/L, a minimum of three nonanastomotic ulcerations or 10 aphthous ulcerations consistent with CD, or a fecal calprotectin > 250 µg/g (2 placebo; 5 vedolizumab), and inadequate or lost response/intolerance of steroids, immunomodulators, and/or TNFα antagonists (2 in placebo; 4 in vedolizumab). Similar results were observed in the TNFα Antagonist Failure ITT Subpopulation.

Criteria that led to exclusion from the Per-Protocol Population are summarized for the Overall ITT Population and the TNFα Antagonist Failure ITT Subpopulation in the table below.

Table 41 Criteria Leading to Exclusion from the Per-Protocol Population TNFα Antagonist Failure ITT Subpopulation and Overall ITT Population Study C13011

	TNFα Antagonist Failure ITT Subpopulation			Overall ITT Population		
Criterion[a], n (%)	PLA N = 157	VDZ N = 158	Total N = 315	PLA N = 207	VDZ N = 209	Total N = 416
Number of patients excluded from the Per-Protocol Population	12	11	23	13	17	30
Baseline and screening CDAI scores < 210 OR > 410 or CD duration < 3 months	2 (1)	3 (2)	5 (2)	2 (< 1)	4 (2)	6 (1)
Received incorrect study medication as assigned at any study visit	0	0	0	0	0	0
Received < 3 doses of study medication, unless patient met 1 of the criteria for failure	4 (3)	5 (3)	9 (3)	5 (2)	7 (3)	12 (3)
Received concomitant corticosteroids or other potentially effective medications for unrelated comorbid condition	2 (1)	0	2 (< 1)	2 (< 1)	0	2 (< 1)
Invalid Week 10 assessment[b]	8 (5)	8 (5)	16 (5)	9 (4)	13 (6)	22 (5)
Patient who had blind broken	0	0	0	0	0	0

Copied from Table 10-3, page 83 CSR.

As seen from the table above, In the Overall ITT Population, a total of 30 patients (13 in placebo; 17 in vedolizumab) met at least one criterion that led to exclusion from the Overall Per-Protocol Population. The most common reasons for exclusion in both treatment groups were invalid Week 10 assessment (9 in placebo; 13 in vedolizumab) and receipt of < 3 doses of study medication, unless the patient met the criteria for failure (5 in placebo; 7 in vedolizumab). Six patients (2 in placebo; 4 in vedolizumab) had screening/baseline CDAI scores out of range and were excluded from the Overall Per-Protocol Population; two placebo patients and three vedolizumab patients had baseline scores ranging from 418 to 564 and one vedolizumab patient had a baseline CDAI score of 203.

In the TNFα Antagonist Failure ITT Subpopulation, the number of patients who met at least one criterion that led to exclusion from the TNFα Antagonist Failure Per-Protocol Subpopulation was similar between the treatment groups (12 in placebo; 11in vedolizumab). Similar to the Overall ITT Population, the most common reasons for exclusion in both treatment groups were invalid Week 10 assessment (8 in placebo; 8 in vedolizumab) and receipt of < 3 doses of study medication unless the patient met the criteria for failure (4 in placebo; 5 in vedolizumab). The two placebo patients and three vedolizumab patients with baseline CDAI scores > 410 were also excluded from the TNFα Antagonist Failure Per-Protocol Subpopulation.

3.1.2.3.2 Treatment Group Comparability

Baseline demographic characteristics of the Overall ITT Population and the TNFα Antagonist Failure ITT Subpopulation are summarized by treatment group in Appendix Table 24,

As seen from Appendix Table 24, baseline demographic characteristics were generally similar between the treatment groups in the Overall ITT Population. Among all patients, there were a higher proportion of female patients than male patients (57% vs. 43%). Most patients were White and non-Hispanic. The mean age was 37.9 years; most patients were ≥ 35 years of age (54%) and few patients were ≥ 65 years (2%). More placebo-treated patients (51%) than vedolizumab-treated patients (42%) were < 35 years. The mean body weight was 70.4 kg and the mean body mass index (BMI) was 24.3 kg/m_2. With respect to geographic distribution, 28% were enrolled at sites in the US and 72% were enrolled at sites outside of the US, including 21% at Central European sites, 19% at Canadian sites, 18% at Western/Northern European sites, 8% at sites located in Asia, Australia, and Africa, and 6% at Eastern European sites.

The demographic characteristics of the TNFα Antagonist Failure ITT Subpopulation were similar to those observed for the Overall ITT Population, except that the difference between the treatment groups in patients < 35 years of age was less pronounced (placebo 46%; vedolizumab 41%). In addition, the TNFα Antagonist Failure ITT Subpopulation had greater proportions of patients enrolled at sites in North America and smaller proportions of patients enrolled at sites in Central Europe than the Overall ITT Population.

Appendix Table 25 presents a comparison between the treatment groups for selected demographic characteristics.

As seen from Appendix Table 25, there were no significant differences between the treatment groups for these selected parameters in either the Overall ITT Population or the TNFα Antagonist Failure ITT Subpopulation.

Baseline (Week 0) CD disease characteristics of the Overall ITT Population and the TNFα Antagonist Failure ITT Subpopulation are summarized by treatment group in Appendix Table 26.

As seen from Appendix Table 26, consistent with the study's inclusion criteria, patients with moderately to severely active CD were enrolled, as demonstrated by the baseline disease characteristics of the treatment groups. In the Overall ITT Population, the mean duration of disease was 10.3 years, with the majority of the patients having been diagnosed for ≥ 7 years (57%). Thirty seven percent (37%) of vedolizumab-treated patients had a baseline CDAI score > 330 compared with 29% of the placebo-treated patients. The majority of the patients had a baseline CRP > 10 mg/L (50%), a baseline fecal calprotectin > 500 μg/g (58%), and disease involvement of both the ileum and colon (61%). A history of prior surgery for CD was reported for 44% of the patients. The majority of the patients in both treatment groups had no history of fistulizing disease, and only 12% of the patients had a draining fistula at baseline. Extraintestinal manifestations of the disease were present at baseline in 59% of the patient s. Most patients in both treatment groups had never smoked or were former smokers (70%).

The baseline CD characteristics of the TNFα Antagonist Failure ITT Subpopulation were similar to those observed for the Overall ITT Population; except for disease duration and baseline CDAI score. The mean duration of disease was somewhat longer in the TNFα Antagonist Failure ITT Subpopulation, with 64% of the patients having been diagnosed for ≥ 7 years.

Treatment failure to CD therapies is summarized by treatment group for the Overall ITT Population in Appendix Table.27.

Information regarding prior use of CD medications, previous treatment failure, and concomitant medications was captured at both screening and Week 0, and during the study. Therefore, the numbers of patients in this table and subsequent summaries of baseline and concomitant medication use may vary based on how the data were collected (IVRS versus eCRF).

As seen from Appendix Table 27, of the 416 patients in the Overall ITT Population, 75% had previously failed TNFα antagonist therapy per the eCRF. The remaining 25% of the patients in the Overall ITT Population were naïve to TNFα antagonist therapy. The proportions of patients who had previously failed TNFα antagonist therapy or were naïve to TNFα antagonist therapy were similar between the treatment groups. In addition, the treatment groups were similar with respect to the number of TNFα antagonist therapies patients had previously failed.

A hierarchical approach was used to categorize treatment failure to TNFα antagonists immunomodulators, and corticosteroids ("worst treatment failure"). TNFα antagonist failure was prioritized over failure of immunomodulators, which was prioritized over failure of corticosteroids. Within each treatment category, patients were categorized by type of failure to a particular agent. For TNFα antagonists, patients were categorized as having had an inadequate response (persistently active disease despite induction treatment), loss of response (recurrence of symptoms during maintenance treatment following prior clinical benefit), or intolerance (treatment-related toxicity). For immunomodulators and corticosteroids, treatment failure was categorized as either inadequate response (persistently active disease despite a 4-week regimen of corticosteroids or an 8-week regimen of immunomodulators) or intolerance. As patients may have had more than 1 definition of treatment failure, only 1 category was assigned to each patient.

A worst treatment failure was assigned using a hierarchical approach, with inadequate response considered worse than loss of response, and loss of response worse than intolerance. According to this categorization in the Overall ITT Population, 75% of patients had history of failure to TNFα antagonists and 21% had failed immunomodulators (without TNFα antagonist failure). Few patients failed corticosteroids alone (3%). In patients who had experienced TNFα antagonist failure, 43% had an inadequate response (i.e., primary `treatment` failures) and 45% had loss of response (i.e., secondary treatment failures). Medication failure categories were comparable between the treatment groups.

Baseline CD therapy is summarized for the Overall ITT Population and for the TNFα Antagonist Failure ITT Subpopulation in Appendix Table 28,

As seen from Appendix Table 28, in the Overall ITT Population, the majority (52%) of the patients reported corticosteroid use at baseline; 35% were treated with corticosteroids alone. Approximately one-third (34%) of the patients reported immunomodulator use at baseline; 16% were treated with immunomodulators alone. Baseline CD therapy use in the TNF α Antagonist Failure ITT Subpopulation was similar to that observed for the Overall ITT Population.

Appendix Table 29 presents a comparison between the treatment groups of selected baseline Crohn's disease characteristics.

As seen from Appendix Table 29, in the Overall ITT Population, the mean baseline disease activity, as assessed by the baseline CDAI score, was statistically significantly higher in the vedolizumab group (313.9) than the placebo group (301.3); this difference was marginally significant in the TNFα Antagonist Failure ITT Subpopulation (306.1 placebo; 316.1 vedolizumab). In both patient populations, no statistically significant differences were observed between the treatment groups for the proportions of patients who were receiving corticosteroids or who were receiving immunomodulators at baseline.

3.1.2.3.3 Analysis Populations

Table below summarizes the analysis populations in this study by treatment group for the overall patient population and the TNFα antagonist failure patient subpopulation. All randomized patients received at least one dose of blinded study drug and were included in the Overall Safety and ITT Populations. The TNFα Antagonist Failure ITT Subpopulation and the Overall ITT Population were the primary analysis populations for the evaluation of efficacy.

Table 42 Summary of Analysis Populations – Study C13011

	TNFα Antagonist Failure Patient Subpopulation		Overall Patient Population	
Data Set, n (%)	PLA N = 157	VDZ N = 158	PLA N = 207	VDZ N = 209
Randomized patients	157	158	207	209
Safety Population[a]	157 (100)	158 (100)	207 (100)	209 (100)
ITT Population[b]	157 (100)	158 (100)	207 (100)	209 (100)
Modified ITT Population[c]	155 (99)	155 (98)	205 (99)	206 (99)
Per-Protocol Population[d]	145 (92)	147 (93)	194 (94)	192 (92)
Completers (Observed Case) Population[e]	137 (87)	147 (93)	184 (89)	191 (91)

Copied from Table 11-1, page 97 CSR.

3.1.2.3.4 Applicant's Analyses of the Primary Efficacy Endpoint

The primary efficacy endpoint for this study was the proportion of patients in clinical remission at Week 6 in the TNFα Antagonist Failure ITT Subpopulation.

Clinical remission is defined as CDAI score ≤ 150 points. Enhanced clinical response is defined as a ≥ 100 point reduction from baseline in CDAI score.

The CMH chi-square test was performed with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) concomitant immunomodulator use (yes/no).

Results from these analyses are summarized by treatment groups is given in the table below

Table 43 Clinical Remission at Week 6
TNFα Antagonist Failure ITT Subpopulation
Study C13011

Clinical Remission[a]	PLA N = 157	VDZ N = 158
Number (%) achieving clinical remission	19 (12.1)	24 (15.2)
95% CI	(7.0, 17.2)	(9.6, 20.8)
Difference from placebo		3.0
95% CI for difference from placebo[b]		(-4.5, 10.5)
P-value for difference from placebo[c]		0.4332
Relative risk[d]		1.2
95% CI for relative risk		(0.7, 2.2)

Copied from Table 11-2, page 99 CSR.

As seen from the table above, in the TNFα Antagonist Failure ITT Subpopulation, no statistically significant difference was observed between the vedolizumab and placebo groups for the proportions of patients in clinical remission at Week 6. Of the 158 patients who received vedolizumab, 24 (15.2%) achieved clinical remission at Week 6 compared with 19 of 157 (12.1%) patients who received placebo. The treatment difference from placebo was 3.0% (95% CI: -4.5, 10.5; p = 0.4332).

The proportions of patients who achieved clinical remission at Week 6 are presented in Appendix Table 30 for the TNFα Antagonist Failure Per-Protocol Subpopulation and in Appendix Table 31 for the TNFα Antagonist Failure Completers (Observed Case) Subpopulation.

As seen from Appendix Tables 30 and 31, the proportions of patients who achieved clinical remission at Week 6 in the TNFα Antagonist Failure Per-Protocol Subpopulation and in the TNFα Antagonist Failure Completers (Observed Case) Subpopulation were similar to those observed in the TNFα Antagonist Failure ITT Subpopulation, with treatment differences from placebo of 3.7% (95% CI: -4.2, 11.6; p = 0.3626) and 3.1% (95% CI: -5.1, 11.2; p = 0.4603), respectively.

3.1.2.3.5 Applicant's Analyses of the Secondary Efficacy Endpoints

The four sets of secondary efficacy endpoints for this study are presented by the treatment groups as follows:

- Proportion of patients in clinical remission at Week 6 in the Overall ITT Population
- Proportion of patients in clinical remission at Week 10 in the TNFα Antagonist Failure ITT Subpopulation and in the Overall ITT Population
- Proportion of patients with sustained clinical remission (i.e., clinical remission at both Week 6 and Week 10) in the TNFα Antagonist Failure ITT Subpopulation and in the Overall ITT Population
- Proportion of patients with enhanced clinical response at Week 6 in the TNFα Antagonist Failure ITT Subpopulation

Analyses of the secondary efficacy endpoints were performed using a sequential testing procedure, and only if the primary endpoint comparison was significant. Since the primary efficacy endpoint did not reach statistical significance, formal hypothesis testing could not be performed for the ranked secondary endpoints. However, observed p-values, relative risks, and 95% confidence intervals are presented for descriptive purposes.

3.1.2.3.5.1 Clinical Remission at Week 6 – Overall ITT Population

The CMH chi-square test was performed with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) previous failure of, or naïve to TNFα antagonists; 3) concomitant immunomodulator use (yes/no).

The proportions of patients who achieved clinical remission at Week 6 in the Overall ITT Population are summarized by treatment group in the table below.

Table 44 Clinical Remission at Week 6
Overall ITT Population
Study C13011

Clinical Remission[a]	PLA N = 207	VDZ N = 209
Number (%) achieving clinical remission	25 (12.1)	40 (19.1)
95% CI	(7.6, 16.5)	(13.8, 24.5)
Difference from placebo		6.9
95% CI for difference from placebo[b]		(0.1, 13.8)
P-value for difference from placebo[c]		0.0478
Relative risk[d]		1.6
95% CI for relative risk		(1.0, 2.5)

Copied from Table 11-3, page 101 CSR.

As seen from the table above, in the Overall ITT Population, 19.1% of vedolizumab-treated patients and 12.1% of placebo-treated patients achieved clinical remission at Week 6; the treatment difference from placebo was 6.9%.

3.1.2.3.5.2 Clinical Remission at Week 10 – TNFα Antagonist Failure ITT Subpopulation and Overall ITT Population

For TNFα antagonist failure ITT subpopulation, the CMH chi-square test was performed with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) concomitant immunomodulator use (yes/no).

For overall ITT population, the CMH chi-square test was performed with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) previous failure of, or naïve to TNFα antagonists; 3) concomitant immunomodulator use (yes/no).

The proportions of patients who achieved clinical remission at Week 10 in the TNFα Antagonist Failure ITT Subpopulation and the Overall ITT Population are summarized by treatment groups in the table below.

Table 45 Clinical Remission at Week 10
TNFα Antagonist Failure ITT Subpopulation and Overall ITT Population
Study C13011

	TNFα Antagonist Failure ITT Subpopulation		Overall ITT Population	
Clinical Remission[a]	PLA N = 157	VDZ N = 158	PLA N = 207	VDZ N = 209
Number (%) achieving clinical remission	19 (12.1)	42 (26.6)	27 (13.0)	60 (28.7)
95% CI	(7.0, 17.2)	(19.7, 33.5)	(8.5, 17.6)	(22.6, 34.8)
Difference from placebo		14.4		15.5
95% CI for difference from placebo[b]		(5.7, 23.1)		(7.8, 23.3)
P-value for difference from placebo[c]		0.0012		< 0.0001
Relative risk[d]		2.2		2.2
95% CI for relative risk		(1.3, 3.6)		(1.4, 3.3)

Copied from Table 11-4, page 105 CSR.

As seen from the table above, In the TNFα Antagonist Failure ITT Subpopulation, 26.6% of vedolizumab-treated patients and 12.1% of placebo-treated patients achieved clinical remission at Week 10; the treatment difference from placebo was 14.4%. The proportion of patients who achieved clinical remission at Week 10 increased from 15.2% at Week 6 in the vedolizumab group and was essentially unchanged from Week 6 (12.1%) in the placebo group.

In the Overall ITT Population, 28.7% of vedolizumab-treated patients and 13.0% of placebo-treated patients achieved clinical remission at Week 10; the treatment difference from placebo was 15.5%. The proportion of patients who achieved clinical remission increased from 19.1% at

Week 6 in the vedolizumab group and showed little change from Week 6 (12.1%) in the placebo group.

3.1.2.3.5.3 Sustained Clinical Remission - TNFα Antagonist Failure ITT Subpopulation and Overall ITT Population

Sustained clinical remission was defined as clinical remission at both Week 6 and Week 10.

For TNFα antagonist failure ITT subpopulation, the CMH chi-square test was performed with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) concomitant immunomodulator use (yes/no).

For overall ITT population, the CMH chi-square test was performed with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) previous failure of, or naïve to TNFα antagonists; 3) concomitant immunomodulator use (yes/no).

The proportions of patients who achieved sustained clinical remission at both Week 6 and Week 10in the TNFα antagonist failure ITT subpopulation and the overall ITT population are summarized by the treatment groups in the table below.

Table 46 Sustained Clinical Remission
TNFα Antagonist Failure ITT Subpopulation and Overall ITT Population
Study C13011

	TNFα Antagonist Failure ITT Subpopulation		Overall ITT Population	
Sustained Clinical Remission[a]	PLA N = 157	VDZ N = 158	PLA N = 207	VDZ N = 209
Number (%) achieving sustained clinical remission	13 (8.3)	19 (12.0)	17 (8.2)	32 (15.3)
95% CI	(4.0, 12.6)	(7.0, 17.1)	(4.5, 12.0)	(10.4, 20.2)
Difference from placebo		3.7		7.0
95% CI for difference from placebo[b]		(-2.9, 10.3)		(0.9, 13.1)
P-value for difference from placebo[c]		0.2755		0.0249
Relative risk[d]		1.4		1.9
95% CI for relative risk		(0.7, 2.8)		(1.1, 3.2)

Copied from Table 11-5, page 112 CSR.

As seen from the table above, in the TNFα Antagonist Failure ITT Subpopulation, 12.0% of vedolizumab-treated patients and 8.3% of placebo-treated patients achieved sustained clinical remission; the treatment difference from placebo was 3.7%. In the Overall ITT Population, 15.3% of vedolizumab treated patients and 8.2% of placebo-treated patients achieved sustained clinical remission; the treatment difference from placebo was 7.0%.

3.1.2.3.5.4 Enhanced Clinical Response at Week 6 TNFα Antagonist Failure ITT Subpopulation

Enhanced clinical response was defined as a ≥ 100 point reduction in CDAI from baseline.

The CMH chi-square test was performed with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) concomitant immunomodulator use (yes/no).

The proportion of patients who achieved enhanced clinical response at Week 6 in the TNFα Antagonist Failure ITT Subpopulation is summarized by treatment group in the table below

Table 47 Enhanced Clinical Response at Week 6
TNFα Antagonist Failure ITT Subpopulation
Study C13011

Enhanced Clinical Response[a]	PLA N = 157	VDZ N = 158
Number (%) achieving enhanced clinical response	35 (22.3)	62 (39.2)
95% CI	(15.8, 28.8)	(31.6, 46.9)
Difference from placebo		16.9
95% CI for difference from placebo[b]		(6.7, 27.1)
P-value for difference from placebo[c]		0.0011
Relative risk[d]		1.8
95% CI for relative risk		(1.2, 2.5)

Copied from Table 11-6, page 114 CSR.

As seen from the table above, in the TNFα Antagonist Failure ITT Subpopulation, 39.2% of vedolizumab-treated patients and 22.3% of placebo-treated patients achieved enhanced clinical response at Week 6; the treatment difference from placebo was 16.9%.

3.1.2.4 Reviewer's Comments and Evaluation

Per the FDA's requested, the summary tables below for the ITT Population for Studies C13011 are provided by the applicant for analyses using the alternative definition of clinical remission regardless of the CDAI score.

Table 48 Clinical Remission at Week 6 – Based on Alternative Definition from FDA TNFα Antagonist Failure ITT Subpopulation Study C13011

Clinical remission	Placebo (n=157)	VDZ (n=158)
N (%) achieve clinical remission at week 6	7 (4.5)	10 (6.3)
95% CI	(1.2, 7.7)	(2.5,10.1)
Difference from placebo		1.8
95% CI for difference from placebo		(-3.1,6.8)
p-value for difference from placebo		0.4642

Complied from Table 39.2.1.1 by this reviewer.
P-values were based on the CMH chi-square test was performed with 2 stratifications: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to TNFα antagonists and/or concomitant immunomodulator use (yes/no)

Table 49 Clinical Remission at Week 6 – Based on Alternative Definition from FDA Intent-to-Treat Population Study C13011

Clinical remission	Placebo (n=207)	VDZ (n=209)
N (%) achieve clinical remission at week 6	10 (4.8)	20 (9.6)
95% CI	(1.9, 7.8)	(5.6,13.6)
Difference from placebo		4.7
95% CI for difference from placebo		(-0.2,9.5)
p-value for difference from placebo		0.060

Complied from Table 48.2.1.1Bby this reviewer.
P-values were based on the CMH chi-square test was performed with 2 stratifications: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to TNFα antagonists and/or concomitant immunomodulator use (yes/no)

As seen from tables above, based on the alternative definition from FDA, vedolizumab-treated patients failed to achieve statistical significance for the clinical remission at Week 6 for TNFα antagonist failure ITT subpopulation.

3.2 Evaluation of Safety

3.2.1 Study C13007

3.2.1.1 Induction Phase

An overall summary of AEs during the Induction Phase is presented for the Induction Phase Safety Population in the table below.

Table 50 Overall Summary of Adverse Event
Induction Phase Safety Population
Study C13007

Adverse Event Category, n (%)	Induction Cohort 1 ITT Population[a] PLA N = 148	Induction Cohort 1 ITT Population[a] VDZ N = 220	Induction Cohort 2[b] Open-Label VDZ N = 747	Combined VDZ N = 967
Any adverse event	88 (59)	124 (56)	426 (57)	550 (57)
Drug-related adverse event	31 (21)	51 (23)	165 (22)	216 (22)
Adverse event resulting in study discontinuation	9 (6)	9 (4)	24 (3)	33 (3)
Serious adverse event	9 (6)	20 (9)	52 (7)	72 (7)
Serious infection adverse events	2 (1)	1 (< 1)	10 (1)	11 (1)
Drug-related serious adverse event	0	3 (1)	4 (< 1)	7 (< 1)
Serious adverse event resulting in discontinuation	5 (3)	5 (2)	15 (2)	20 (2)
Deaths	0	0	1 (< 1)	1 (< 1)

Copied from Table 33, page 161 CSR.

As seen from the table above, the overall incidence of AEs in the Induction Phase Safety Population was similar across the treatment groups, with 59% of placebo-treated patients and 56% of vedolizumab-treated patients in the ITT Population, and 57% of the patients who received open-label vedolizumab experiencing at least one AE during the study. Drug-related AEs, as considered by the investigator, were reported for 21% of the placebo patients and 23% of the vedolizumab patients in the ITT population, and for 22% of the open-label vedolizumab patients. Premature discontinuation from study due to AEs was highest among placebo patients (6%), followed by vedolizumab patients in the ITT Population (4%), and patients who received open-label vedolizumab (3%). Serious AEs were experienced by 6% of the placebo patients and 9% of the vedolizumab patients in the ITT Population, and by 7% of patients who received open-label vedolizumab. Serious infection AEs and drug-related SAEs occurred in ≤ 1% of patients in each of the treatment groups. Serious AEs that resulted in study discontinuation were experienced by 3% of the placebo patients and 2% of the vedolizumab patients in the ITT Population, and by 2% of the patients who received open-label vedolizumab. None of the placebo or vedolizumab patients in the ITT Population died during the Induction Phase of the study. One patient who received open-label vedolizumab died due to myocarditis, 75 days after his last dose of study drug.

3.2.1.2 Maintenance Phase

An overall summary of AEs during the Induction Phase is presented for the Maintenance Phase Safety Population in the table below.

Table 51 Overall Summary of Adverse Event
Maintenance Phase Safety Population
Study C13007

	Maintenance ITT[a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Maintenance Non-ITT		Combined	
Adverse Event Category n (%)	PLA N = 153	VDZ Q8W N = 154	VDZ Q4W N = 154	PLA[b] (from Week 0) N = 148	VDZ Q4W[c] (Week 6 Nonresponders) N = 506	PLA N = 301	VDZ N = 814
Any adverse event	128 (84)	135 (88)	130 (84)	118 (80)	441 (87)	246 (82)	706 (87)
Drug-related adverse event	51 (33)	63 (41)	63 (41)	45 (30)	191 (38)	96 (32)	317 (39)
Adverse event resulting in study discontinuation	15 (10)	12 (8)	9 (6)	14 (9)	70 (14)	29 (10)	91 (11)
Serious adverse event	23 (15)	28 (18)	25 (16)	23 (16)	146 (29)	46 (15)	199 (24)
Serious infection adverse events	5 (3)	6 (4)	9 (6)	4 (3)	30 (6)	9 (3)	45 (6)
Drug-related serious adverse event	4 (3)	5 (3)	6 (4)	2 (1)	24 (5)	6 (2)	35 (4)
Serious adverse event resulting in discontinuation	7 (5)	9 (6)	5 (3)	8 (5)	45 (9)	15 (5)	59 (7)
Deaths	0	1 (< 1)	0	1 (< 1)	3 (< 1)	1 (< 1)	4 (< 1)

Copied from Table 77, page 301 CSR.

As seen from the table above, the overall incidence of AEs in the Maintenance Study ITT Population was similar among the treatment groups, with 84% of the placebo patients, 88% of vedolizumab Q8W patients, and 84% of the vedolizumab Q4W patients experiencing at least one AE during the study. As AE rates are influenced by patients' duration on study, incidence density analyses were performed to adjust for differences in overall exposure. The number of AEs per 100 patient-years was similar among the treatment groups (placebo 688.6; vedolizumab Q8W 578.3; vedolizumab Q4W 685.7).

Drug-related AEs, as considered by the investigator, were reported for somewhat greater proportions of patients in the vedolizumab groups (41% each) compared with the placebo group (33%). Premature discontinuation from study due to AEs was highest among placebo patients (10%), followed by the vedolizumab Q8W (8%) and vedolizumab Q4W (6%) treatment groups.

Serious AEs were experienced by similar proportions of patients in each of the treatment groups in the Maintenance Study ITT Population (placebo 15%; vedolizumab Q8W 18%; vedolizumab Q4W 16%). The incidences of SAEs leading to discontinuation were generally similar among the treatment groups. The serious infection AE rates were 4% in the vedolizumab Q8W group, 6% in the vedolizumab Q4W group and 3% in placebo. When adjusted for patients' duration on study, the number of serious infection events per 100 patient-years was similar among the treatment groups (placebo 7.3; vedolizumab Q8W 6.4; vedolizumab Q4W 8.4).

In the all vedolizumab combined group, which includes ITT vedolizumab patients and non-ITT vedolizumab patients (non-responders to vedolizumab induction treatment), 87% of patients reported at least one AE. The SAE rate was 24% in the all vedolizumab combined group and 16□ in the non-ITT placebo group (patients treated with placebo for the entire duration of the study). The rates of serious infections for the all vedolizumab combined and non-ITT placebo groups were 6% and 3%, respectively. The number of SAEs per 100 patient-years for the all vedolizumab combined and non-ITT placebo groups was 51.5 and 36.5, respectively. The number of serious infections per 100 patient-years was 9.3 for the all vedolizumab combined

group and 4.9 for the non-ITT placebo group. The higher rates in the combined vedolizumab arm appear to be driven by higher rates in the non-ITT vedolizumab group, who were non-responders to vedolizumab induction therapy and had greater severity of disease than the vedolizumab responder group.

Five deaths were reported in this study. One death occurred in a vedolizumab patient during the Induction Phase (myocarditis) that was considered not related to study drug. Three deaths occurred in vedolizumab patients during the Maintenance Phase; of these, two were considered related to study drug (CD and sepsis in one patient and septic shock in one patient) and 1 was considered not related (intentional overdose). One death occurred in a non-ITT placebo patient (bronchopneumonia) and was considered not related. In addition, one death (cardio-respiratory arrest) occurred post study, 660 days (nearly two years) after the patient's last dose of vedolizumab.

3.2.2 Study C13011

An overall summary of AEs is presented for the Overall Safety Population and the TNFα Antagonist Failure Safety Subpopulation in the table below.

Table 52 Overall Summary of Adverse Event
Induction Phase Safety Population
Study C13011

	TNFα Antagonist Failure Safety Subpopulation		Overall Safety Population	
Adverse Event Category, n (%)	PLA N = 157	VDZ N = 158	PLA N = 207	VDZ N = 209
Any adverse event	102 (65)	94 (59)	124 (60)	117 (56)
Drug-related adverse event	30 (19)	31 (20)	34 (16)	34 (16)
Adverse event resulting in study discontinuation	6 (4)	2 (1)	8 (4)	4 (2)
Serious adverse event	14 (9)	8 (5)	16 (8)	13 (6)
Serious infection adverse event	0	2 (1)	0	2 (< 1)
Drug-related serious adverse event	1 (< 1)	0	1 (< 1)	1 (< 1)
Serious adverse event resulting in discontinuation	5 (3)	2 (1)	5 (2)	4 (2)
Deaths	0	0	0	0

Copied from Table 12-2, page 148 CSR.

As seen from the table above, in the Overall Safety Population, 60% of placebo-treated patients and 56% of vedolizumab treated patients experienced at least one AE during the study. Drug-related AEs, as considered by the investigator, were reported in 16% of the patients in both treatment groups. A greater proportion of placebo-treated patients (4%) than vedolizumab-treated patients (2%) experienced an AE that resulted in study discontinuation.

No patient deaths were reported. A total of 16 (8%) placebo patients and 13 (6%) vedolizumab patients experienced an SAE. Serious infection AEs and drug-related SAEs occurred in < 1% of patients in both treatment groups; SAEs that resulted in study discontinuation were experienced by 2% of the patients in both treatment groups. Results observed in the TNFα Antagonist Failure Safety Subpopulation were similar to those observed in the Overall Safety Population. In the TNFα Antagonist Failure Safety Subpopulation, greater proportions of placebo-treated patients experienced SAEs (9%) and AEs that resulted in study discontinuation (4%) compared with vedolizumab-treated patients (5% and 1%, respectively).

4. FINDINGS IN SPECIAL/SUBGROUP POPULATION

4.1 Gender, Race, Age, Other Special/Subgroup Population

Subgroup analyses for clinical response at Week 6 in the Induction Study ITT population and clinical remission at Week 52 in the Maintenance Study ITT population are provided based on: age (age < 35, age ≥ 35 years. age < 65, age ≤ 65 years, gender, race, duration from UC diagnosis to first dose, geographic region, baseline (Week 0) disease activity, baseline (Week 0) fecal calprotectin (≤ 250 µg/g, > 250 µg/g; ≤ 500 µg/g, > 500 µg/g), and disease localization.

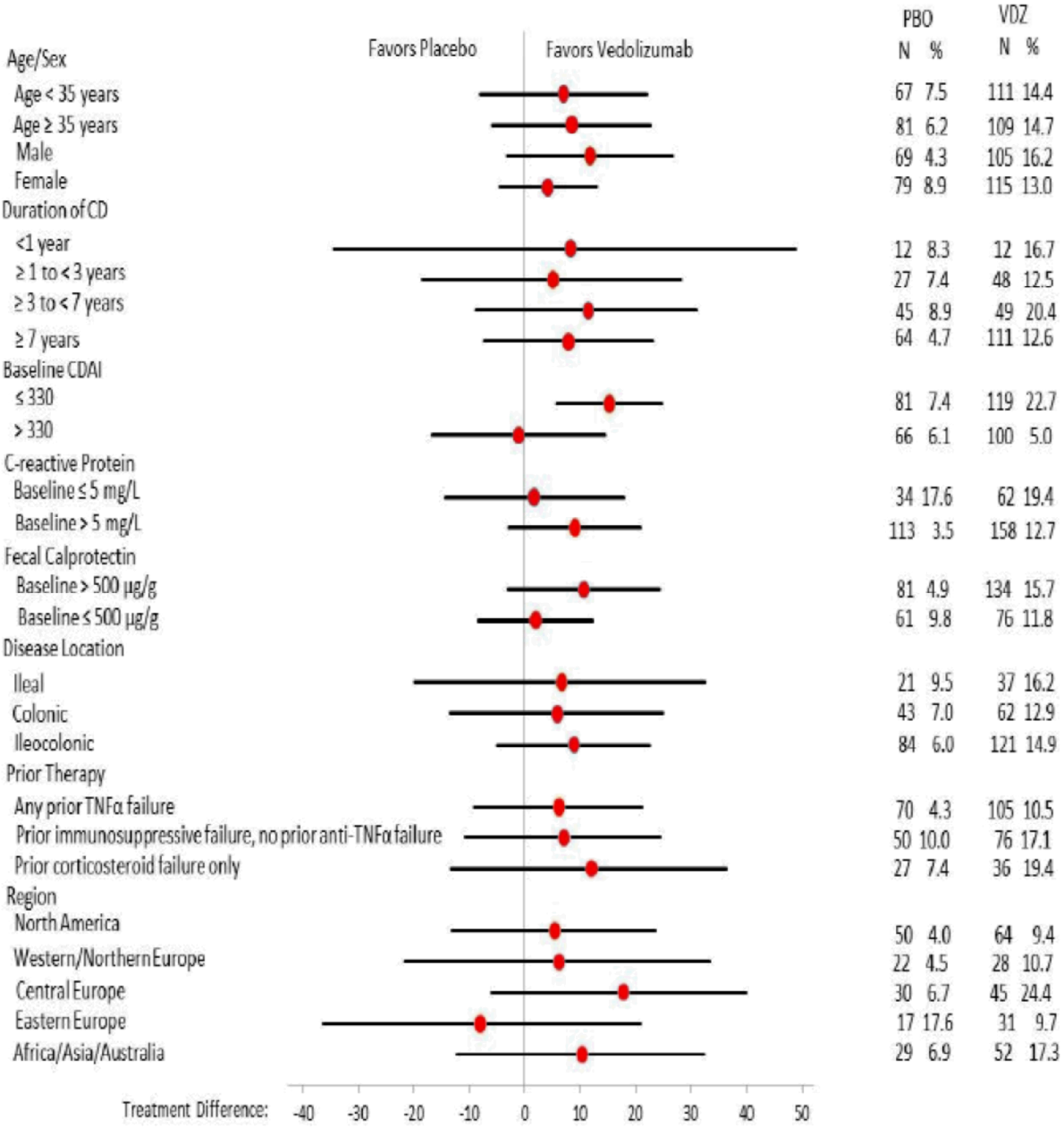
4.1.1 Study C13007

4.1.1.1 Induction Phase

4.1.1.1.1 Clinical Remission at Week 6

Figure below summarizes the risk differences (percentages) from placebo for the primary endpoint of clinical remission at Week 6 in patient subgroups according to demographic characteristics and measure of disease activity in the Induction Study ITT Population.

Figure 8 Treatment Difference in Percentage Points for Clinical Remission at Week 6 with the 95% Confidence Interval by Baseline Subgroups- Induction ITT Population
Study C13007



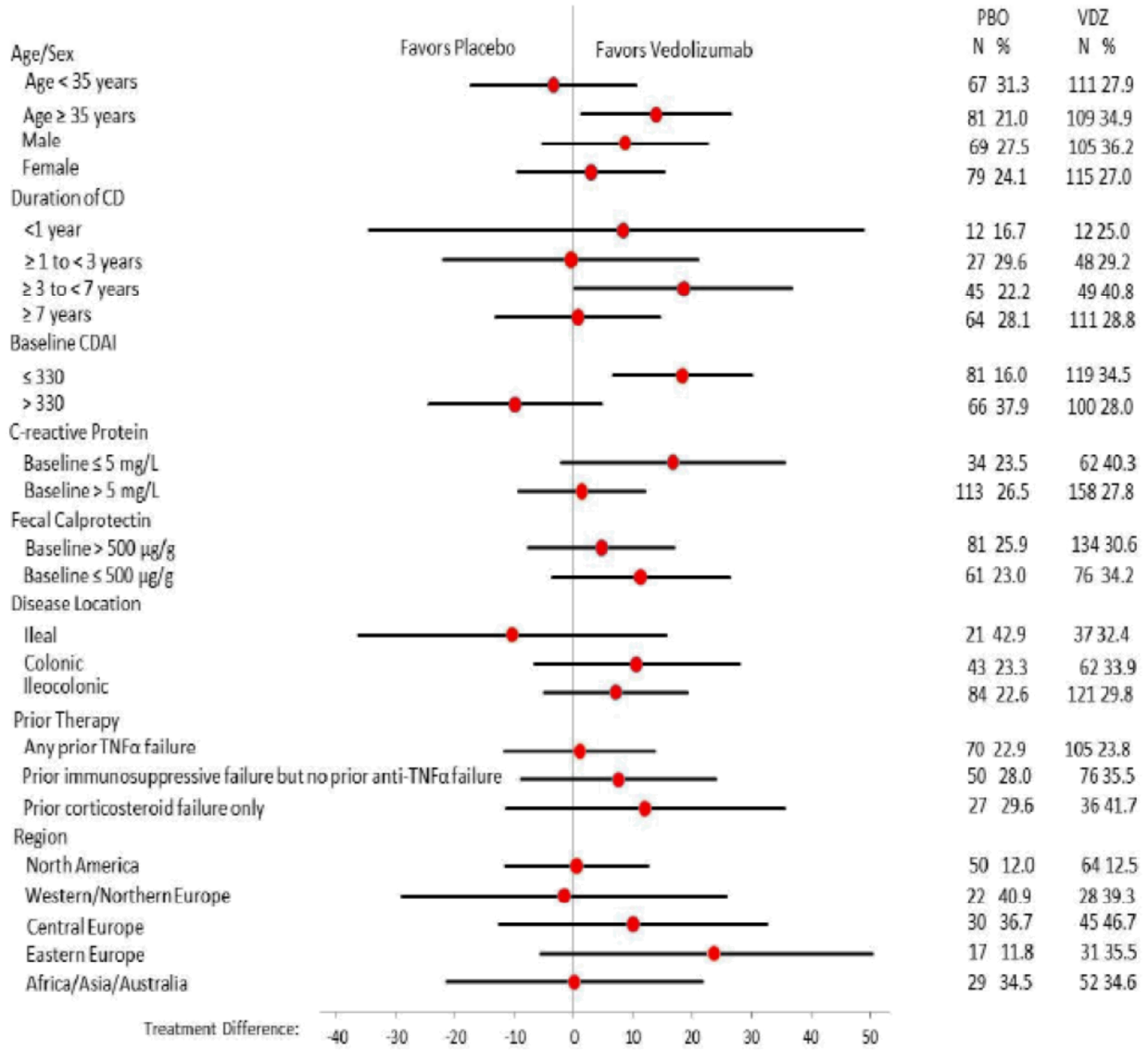
Copied from Figure 7-17, Applicant's Advisory Committee Briefing Document.

As seen from the figure above, the risk difference from placebo favored vedolizumab in the majority of the subgroup analyses, although there was greater variability and the 95% CIs for the differences from placebo often included zero in these analyses.

4.1.1.1.2 Enhanced Clinical Response at Week 6

Figure below summarizes the risk differences (percentages) from placebo for the primary endpoint of enhanced clinical response at Week 6 in patient subgroups according to demographic characteristics and measure of disease activity in the Induction Study ITT Population.

Figure 9 Treatment Difference in Percentage Points for Enhanced Clinical Response (CDAI-100 Response) at Week 6 with 95% Confidence Interval by Baseline Subgroups-Induction ITT Population
Study C13007



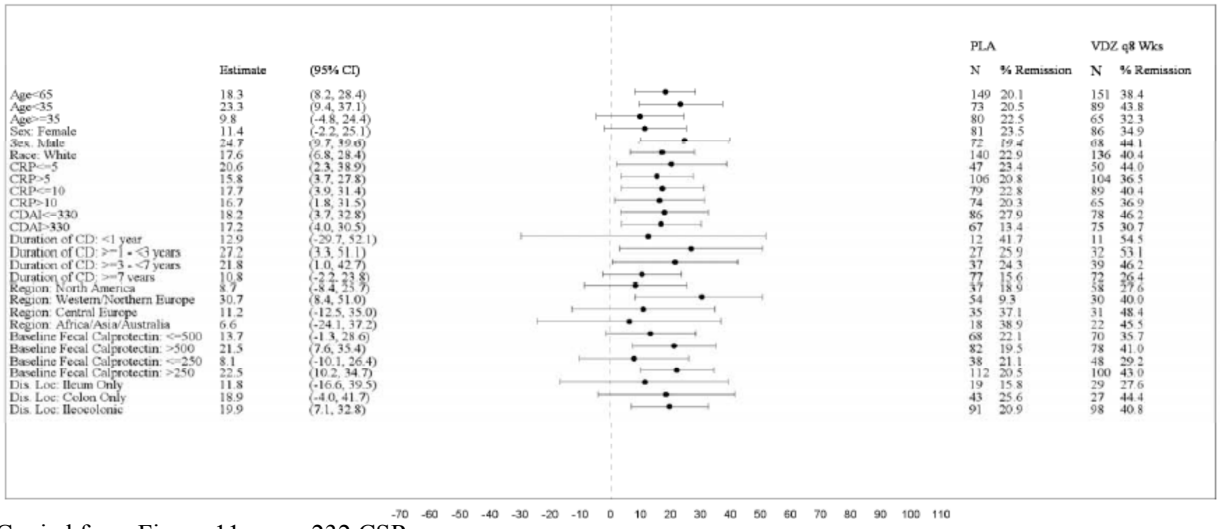
Copied from Applicant's Advisory Committee Briefing Document.

As seen from the figure above, for clinical response at Week 6, the risk difference from placebo favored vedolizumab in the subgroups of patients aged ≥ 35 and of patients who had baseline CDAI ≤ 330 point. There was greater variability and the 95% CIs for the differences from placebo often included zero for most of subgroups.

4.1.1.2 Maintenance Phase

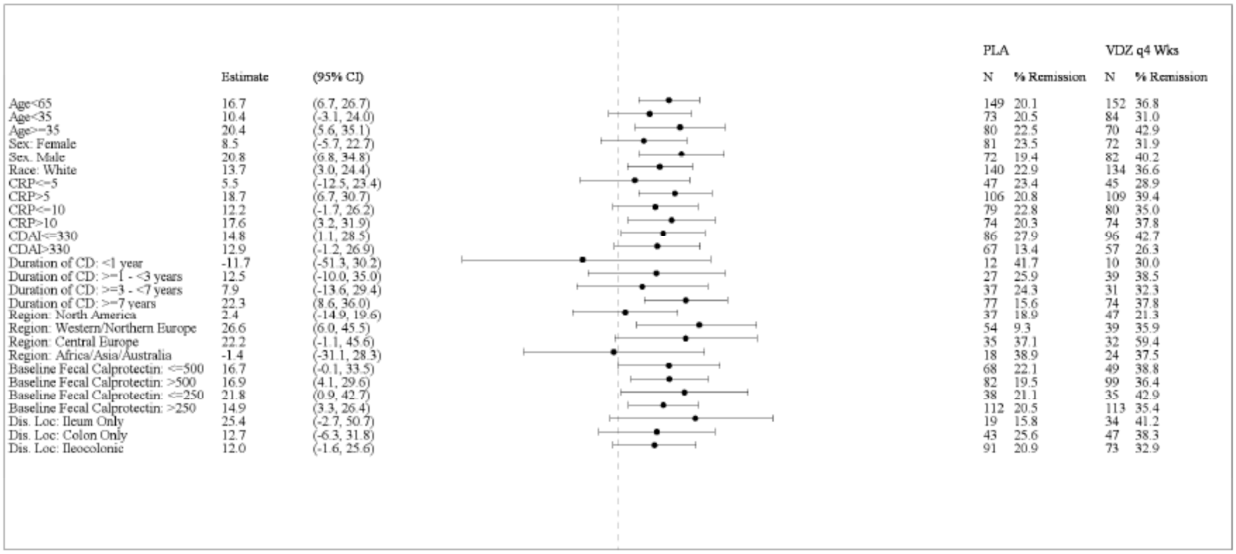
Figures below summarize the risk differences (percentages) from placebo for the vedolizumab Q8W and Q4W treatment groups, respectively, for the primary endpoint of clinical remission at Week 52 in patient subgroups according to demographic characteristics and measures of disease severity in the Maintenance Study ITT Population.

Figure 10 Risk Difference (Percentage) and the 95% Confidence Intervals for Subgroup Analyses of Clinical Remission at Week 52 for Vedolizumab Q8W vs. Placebo Maintenance Study ITT Population Study C13007



Copied from Figure 11, page 232 CSR.

Figure 11 Risk Difference (Percentage) and the 95% Confidence Intervals for Subgroup Analyses of Clinical Remission at Week 52 for Vedolizumab Q4W vs. Placebo Maintenance Study ITT Population
Study C13007



Copied from Figure 13, page 234

As seen from Figures 10 and 11, the treatment benefit of vedolizumab for the maintenance of clinical remission at Week 52 in the Maintenance Study ITT Population was preserved in patient subgroups according to demographic variables and disease characteristics. In both vedolizumab groups, the treatment effect was observed in the majority of the patient subgroups by age, gender, race, and geographic region, although not all of the treatment difference 95% CIs excluded zero. Both males and females had a positive response to treatment, but the treatment differences from placebo were greater in males compared with females in both the vedolizumab Q8W and Q4W treatment groups. With respect to age, the treatment difference in the vedolizumab Q8W treatment group was greater for patients < 35 years of age than for patients 35 years of age or older. Conversely, in the vedolizumab Q4W treatment group, patients 35 years of age or older had a greater treatment difference from placebo compared to patients < 35 years of age.

Similar results were also observed for subgroups according to disease activity and severity, including CDAI, baseline CRP, baseline fecal calprotectin, and disease location. Consistent with the results observed for age, treatment differences from placebo were greater among patients in the vedolizumab Q8W treatment group with a disease duration ≥ 1 to < 3 years and ≤ 3 to < 7 years, compared to those with a disease duration ≥ 7 years, whereas the converse was observed in the vedolizumab Q4W treatment group.

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

5.1.1 Induction Studies

Two studies (C13007 and C13011) were conducted to evaluate vedolizumab as induction therapy for moderate to severe CD. Study C13007 evaluated vedolizumab 300 mg in CD patients, of which 50% subjects who were naïve to TNFα antagonists and 50% patients with previous TNFα antagonists. Study C13011 included approximately 75% patients who had previously failed TNFα antagonist therapy and approximate 25% patients who were naïve to TNFα antagonist therapy.

During the Study C13007 Induction Phase, the applicant elevated the first key secondary endpoint, enhanced clinical response (decrease in CDAI of ≥ 100 points), to a "co-primary" endpoint. The applicant further specified that the primary objective of the study would be met by achieving statistical significance for either of the co-primary endpoints, and the Hochberg method would be used to adjust for multiple comparisons.

There was disagreement between the applicant and FDA regarding the definition of co-primary endpoints. The following statements were conveyed to the applicant in October 1, 2009.

- The term of co-primary endpoint that you have defined for Study C13007is not commonly used for regulatory purposes.

- Two or more primary endpoints are called co-primary if each must be show statistically significant treatment benefit at a pre-specified significance level α (e.g., α=0.025, by one-sided tests).

The applicant performed analyses of clinical remission at Week 6 and enhance clinical response at Week 6 using the Cochran-Mantel-Haenszel (CMH) test, with stratification according to:

4) concomitant use of oral corticosteroids (yes/no);
5) previous exposure to TNFα antagonists (yes.no);
6) concomitant immunomodulator use (yes/no).

Based on the Induction Study ITT Population, a statistically significant greater proportion of vedolizumab-treated patients achieved clinical remission at Week 6 compared with patients who received placebo. The treatment difference was 7.8% (95% CI: 1.2, 14.3; p = 0.0206).

The difference between the vedolizumab and placebo groups was not statistically significant for the other primary endpoint of enhanced clinical response at Week 6. The difference was 5.7% (95% CI: -3.6, 15.0; p = 0.2322).

This reviewer found that 20 ITT patients (10 patients in each group) who had a baseline CDAI score of less than 220 and 2 patients (1 patient in each group) with baseline CDAI missing were enrolled in this study.

Among these 20 patients with a baseline CDAI < 220, a greater proportion of vedolizumab-treated patients achieved clinical remission at Week 6 as compared with patients who received placebo [70% (7/10) vs. 20% (2/10)].

This reviewer also found 18 patients (9 in each treatment group) who had a baseline CDAI score of greater than 450. Among these 18 patients with a baseline CDAI > 450, proportions of vedolizumab-treated and placebo patients achieved clinical remission at Week 6 were zeros.

According to the inclusion criteria (a CDAI score of 220 to 450), these 40 patients should not be enrolled in the study, this reviewer performed a post-hoc sensitivity analysis by excluding these 40 patients. The resulting treatment different would be 6.25% with nominal p-value of 0.0893 (Fisher's Exact test). If these 40 patients were considered as "non-responders", the resulting treatment difference would be 5.95% with nominal p-value of 0.0622 (Fisher's Exact test).

This reviewer also found that 16 Per-Protocol (PP) patients (8 patients in each group) who had a baseline CDAI score of less than 220 were included in the PP analysis in this study.

Among these 16 patients with baseline CDAI < 220, a greater proportion of vedolizumab-treated patients achieved clinical remission at Week 6 as compared with patients who received placebo [62.5% (5/8) vs. 12.5% (1/8)].

According to the final inclusion criteria (a CDAI score of 220 to 450), these 16 patients should not be enrolled in the study. This reviewer performed a post-hoc sensitivity analysis by excluding these 16 patients. The treatment different would be 6.67% with a nominal p-value of 0.0606 based on the Fisher's Exact test. If these 16 patients were considered as "non-responders", the treatment difference would be 6.53% with a nominal p-value of 0.0611 based on the Fisher's Exact test.

This reviewer performed a post-hoc unadjusted analysis of clinical remission at Week 6 using the Fisher's exact test to see whether the applicant's result was robust and method independent.

The resulting p-value from the Fisher's exact test yielded 0.0287 which is greater than 0.025, level of significance. So, the applicant's analysis of clinical remission at Week 6 might not be robust and was method dependent.

Furthermore, it was observed that the treatment group differed with respect to the proportion of patients enrolled by geographic site with p-value of 0.0610.

This reviewer also performed a post-hoc analysis of clinical remission at Week 6 adjusted for geographic site using the CMH chi-square test to see whether the applicant's result was robust and method independent. The resulting p-value from the CMH chi-square test yielded 0.0279

which is greater than 0.025, level of significance. So, the applicant's analysis of clinical remission at Week 6 might not be robust and was method dependent.

This reviewer performed a post-hoc Breslow-Day test to evaluate the homogeneity of subgroup by the baseline CDAI (≤330 vs. >330). The p-value from the Breslow-Day test yielded 0.0636 which is smaller than 0.10, the usual level of significance used for testing interaction. It was suggested that vedolizumab might be more effective for patients with baseline CDAI ≤ 330. However, it needs to be reconfirmed by the other study.

Per FDA's requested, for the Induction Study ITT Population for Studies C13007, the applicant performed a post hoc analysis using the following alternative definition of clinical remission:

- Total number of liquid/very soft stools of ≤ 10 per day in the relevant week; and
- Abdominal pain rated as 0 or 1 for each day in the relevant week.

Based on alternative definition from FDA, vedolizumab-treated patients failed to achieve statistical significance for clinical remission at Week 6 for vedolizumab group with the difference of 4.8% (95 CI: -0.7, 10.3; p = 0.0848).

For subjects who failed TNFα antagonist therapy, Study C13011 failed to demonstrate that there was no statistically significant difference between the vedolizumab and placebo groups for the proportion of patients in clinical remission at Week 6 with treatment difference of 3.0% (95 CI: -4.5, 10.5; p=0.4332). Study C13007 revealed a trend with treatment difference of 6.2% (95% CI: -9.1, 21.3).

5.1.2 Maintenance Study

One study (C13007) was conducted to evaluate vedolizumab as maintenance therapy for moderate to severe Crohn's disease.

The Maintenance Phase included three groups of patients who were to be assigned to treatment groups based on their Induction Phase treatment assignments and responses to the study therapy. Vedolizumab-treated patients from both Induction Cohort 1 (double-blind) and Cohort 2 (open label) who demonstrated a clinical response according to protocol-specified criteria, as assessed by the investigator, were to be randomized in a 1:1:1 ratio to double-blind treatment with vedolizumab administered every 4 weeks (Q4W), vedolizumab administered every 8 weeks (Q8W), or placebo. Randomization was to be stratified by three factors:

- Enrollment in Cohort 1 or Cohort 2 in the Induction Phase
- Concomitant use of oral corticosteroids
- Previous exposure to TNFα antagonists or concomitant immunomodulator use

These patients who were randomized into the Maintenance Phase comprised the Maintenance Study ITT Population, the primary efficacy population.

Vedolizumab-treated patients who did not demonstrate response at Week 6 of the Induction Phase were to continue treatment with open-label vedolizumab, administered Q4W. Patients who had been treated with double-blind placebo in the Induction Study were to continue on double-blind placebo during the Maintenance Phase, regardless of treatment response during induction. The Maintenance Phase began at Week 6, included study drug dosing at Week 6 and Q4W or Q8W thereafter, and concluded with Week 52 assessments.

In the Induction Phase, a total of 220 vedolizumab patients were enrolled into Cohort 1; a total of 747 additional patients were enrolled into Cohort 2. Among 220 vedolizumab patients in Cohort 1, 96 patients (43.6%) were Week 6 responders. Among 747 additional patients in Cohort 2, 365 patients (48.8%) were Week 6 responders.

With substantial number of patients (79%) from Cohort 2 enrolling this study, results from this study were difficult to interpret from statistical perspective. Furthermore, results for vedolizumab against placebo of overall analysis might be driven by that of Cohort 2.

A majority of patients (79%) who were randomized into the Maintenance Phase were from Cohort 2. Compared to Cohort 1, Cohort 2 had a greater proportion of patients who had prior TNFα antagonist use (68%) and failure (63%) vs. 50% and 48%, respectively. Cohort 2 also had more patients at sites in Western/Northern Europe and fewer patients entering at sites in Asia/Australia/Africa and Eastern Europe than was observed for the Cohort 1.

The applicant also noted that an imbalance across the treatment groups in the proportion of patients who had achieved clinical remission at Week 6 was observed due to randomization at Week 6 was not stratified by the remission status. Clinical remission at Week 6 was achieved by 27.9% of patients in the vedolizumab Q4W group and 33.8% of patients in the vedolizumab Q8W group compared with 36.6% of patients in the placebo group.

This imbalance may have had an impact on the analyses of the clinical remission-based endpoints in favor of placebo, especially for the vedolizumab Q4W group versus the placebo group because the clinical remission rate at Week 6 for the vedolizumab Q4W group was about 9% lower than that of the placebo .

Results by cohort were inconsistent between dosing regimens (Q4W vs. Q8W) between cohorts. The results for Q8W against placebo of overall analysis were driven by that of Cohort 2**.**

For the Maintenance Study's primary and secondary endpoints, the applicant used a Hochberg and sequential testing procedure in order to maintain the overall Type I error rate of 0.05. This multiplicity adjustment mehod may not be able to properly control the study-wise Type I error.

This reviewer provided the following comments in the Statistical Review and Evaluation for applicant's IND 9-125 submission S/N 0411 dated January 9, 2012.

However, the Hochberg procedure is generally not recommended for sequencing testing. It is not assumption free. Furthermore, it is known to provide overall α-control for independent and for

certain types of positively correlated endpoints. But its properties for other types of dependent endpoints are not fully known.

> We recommend you use a Bonferroni based gatekeeping procedures to test all endpoints in the primary endpoint family and proceed to the secondary family of endpoints only if there has been statistical success in the primary family. When used as a gatekeeping strategy to test the primary family endpoints, the Bonferroni method has an important property of preserving some alpha for testing the secondary endpoint family when at least one of the endpoints in the primary family is statistical significant. The endpoint-specific alpha from each test that successfully rejects the null hypothesis is summed and becomes the alpha available to the secondary endpoint family.

So, results from secondary efficacy endpoints were difficult to interpret from statistical perspective.

For clinical remission at Week 52 and enhanced clinical response at Week 52, the results by cohort were inconsistent between dosing regimens (Q4W vs. Q8W) between cohorts. The results for Q8W against placebo of overall analysis were driven by that of Induction Cohort 2 for clinical remission at Week 52 and enhanced clinical response at Week 52.

There were more than 58% of the data missing for placebo, more than 53% of the data missing for vedolizumab Q8W and more than 47% of the data missing for vedolizumab Q4W, the observed treatment effect might not be reliable.

5.2 Conclusion and Recommendation

The applicant has submitted two phase 3 studies (C13007 and C13011) to support the indication of Crohn's disease (CD).

Both studies were conducted to evaluate vedolizumab as an induction therapy for moderate to severe CD. Study C13007 evaluated vedolizumab 300 mg in CD patients, of which 50% subjects naïve to TNFα antagonists and 50% patients with previous TNFα antagonists. Study C13011 included approximately 75% patients who had previously failed TNFα antagonist therapy and approximate 25% patients who were naïve to TNFα antagonist therapy.

Two primary efficacy endpoints, clinical remission and enhanced clinical response at Week 6, were pre-specified for Study C13007. Based on the Induction Study ITT Population, study C13007 showed that a statistically significant greater proportion of vedolizumab-treated patients achieved clinical remission at Week 6 compared with patients who received placebo. The treatment difference from placebo was 7.8% (95% CI: 1.2, 14.3; p = 0.0206). However, the treatment difference was found to be marginal and might not be robust.

However, the treatment comparisons on both the other primary endpoint of enhanced clinical response at Week 6 and the secondary efficacy endpoint of changes from baseline in CRP at Week 6, failed to achieve statistical significance.

For subjects who failed TNFα antagonist therapy, Study C13011 failed to demonstrate a statistically significant difference between the vedolizumab and placebo groups for the proportion of patients in clinical remission at Week 6 with treatment difference of 3.0% (95% CI: -4.5, 10.5; p=0.4332). Study C13007 revealed a trend with treatment difference of 6.2% (95% CI: -9.1, 21.3) in this subpopulation favoring vedolizumab.

Only Study C13007 was performed to evaluate vedolizumab as a maintenance therapy for moderate to severe CD. The results from Study C13007 Maintenance Phase showed statistically significant difference on the primary efficacy endpoint of clinical remission at Week 52 for the every eight week (Q8W) regimen. Statistically significant treatment differences were also observed for two of three key secondary efficacy endpoints.

However, for the maintenance phase, results from Study C13007 by the Induction Phase Cohort were inconsistent between dosing regimens (Q4W vs. Q8W) and between Cohorts. The results for vedolizumab Q8W against placebo of overall analysis were driven by Cohort 2**.**

With more than 58% of the data missing for placebo, more than 53% of the data missing for vedolizumab Q8W, and more than 47% of the data missing for vedolizumab Q4W, the observed treatment effect might not be reliable.

Evidence of efficacy given in Study C13007 might not be statistically persuasive.

6. Appendix

Table 1 Baseline Demographic – Induction Phase Safety Population – Study C13007

Parameter	Induction Cohort 1 ITT Population[a] PLA N = 148	Induction Cohort 1 ITT Population[a] VDZ N = 220	Induction Cohort 2[b] Open-label VDZ N = 747	Combined VDZ N = 967	Total N = 1115
Gender, n (%)					
Male	69 (47)	105 (48)	346 (46)	451 (47)	520 (47)
Female	79 (53)	115 (52)	401 (54)	516 (53)	595 (53)
Race, n (%)					
White	124 (84)	182 (83)	689 (92)	871 (90)	995 (89)
Black	3 (2)	3 (1)	17 (2)	20 (2)	23 (2)
Asian	19 (13)	35 (16)	35 (5)	70 (7)	89 (8)
Other	2 (1)	0	6 (< 1)	6 (< 1)	8 (< 1)
Ethnicity, n (%)					
Hispanic or Latino	5 (3)	2 (< 1)	19 (3)	21 (2)	26 (2)
Not Hispanic or Latino	139 (94)	214 (97)	712 (95)	926 (96)	1065 (96)
Not reported	4 (3)	4 (2)	16 (2)	20 (2)	24 (2)
Age (yrs)[c]					
Mean (Std Dev)	38.6 (13.16)	36.3 (11.57)	35.6 (12.01)	35.7 (11.91)	36.1 (12.12)
Median	36.7	34.8	33.0	33.6	34.0
Minimum, maximum	19, 75	18, 77	18, 76	18, 77	18, 77
Age (yrs), n (%)					
< 35	67 (45)	111 (50)	404 (54)	515 (53)	582 (52)
≥ 35	81 (55)	109 (50)	343 (46)	452 (47)	533 (48)
Age (yrs), n (%)					
< 65	142 (96)	218 (99)	732 (98)	950 (98)	1092 (98)
≥ 65	6 (4)	2 (< 1)	15 (2)	17 (2)	23 (2)
Body weight (kg)					
Mean (Std Dev)	68.7 (18.90)	67.1 (19.07)	70.8 (19.56)	69.9 (19.50)	69.8 (19.42)
Median	66.0	65.2	67.0	66.6	66.2
Minimum, maximum	32, 130	30, 167	30, 161	30, 167	30, 167
BMI (kg/m^2)					
Mean (Std Dev)	23.7 (5.77)	23.1 (5.62)	24.2 (6.02)	24.0 (5.95)	23.9 (5.93)
Median	22.3	22.2	22.9	22.9	22.9
Minimum, maximum	12, 45	13, 56	14, 50	13, 56	12, 56
Geographic region[d], n (%)					
North America	50 (34)	64 (29)	291 (39)	355 (37)	405 (36)
Western/Northern Europe	22 (15)	28 (13)	210 (28)	238 (25)	260 (23)
Central Europe	30 (20)	45 (20)	133 (18)	178 (18)	208 (19)
Eastern Europe	17 (11)	31 (14)	42 (6)	73 (8)	90 (8)

Table 1 Baseline Demographic – Induction Phase Safety Population (continued) – Study C13007

Parameter	Induction Cohort 1 ITT Population[a]		Induction Cohort 2[b] Open-label	Combined	Total
	PLA N = 148	VDZ N = 220	VDZ N = 747	VDZ N = 967	N = 1115
Asia/Australia/Africa	29 (20)	52 (24)	71 (10)	123 (13)	152 (14)

Source: Table 14.1.1.5CP.

Abbreviations: BMI = body mass index; ITT = intent-to-treat; PLA = placebo; Std Dev = standard deviation; VDZ = vedolizumab.

a All patients enrolled in Cohort 1 who were randomized to blinded induction treatment with vedolizumab or placebo.

b All patients enrolled in Cohort 2 who received open-label vedolizumab induction treatment.

c Age is defined as (1 + first dose date - birth date)/365.25.

d The countries of each geographic region are specified in Table 14.1.1.3BP.

Table 10 presents a comparison of selected baseline demographic characteristics of patients randomized to placebo versus patients randomized to vedolizumab in the Induction Study ITT Population. No statistically significant differences were noted between the treatment groups for selected baseline demographic characteristics including gender, race, age, body weight, and geographic region.

Table 2 Comparison by Treatment Group of Selected Baseline Demographic Characteristics – Induction Study Populations – Study C13007

Parameter	PLA N = 148	VDZ N = 220	P-value[a]
Gender, n (%)			0.8350
Male	69 (47)	105 (48)	
Female	79 (53)	115 (52)	
Race, n (%)			0.7906
White	124 (84)	182 (83)	
Other	24 (16)	38 (17)	
Age (yrs)[b]			0.1803
Mean (Std Dev)	38.6 (13.16)	36.3 (11.57)	
Body weight (kg)			0.4130
Mean (Std Dev)	68.7 (18.90)	67.1 (19.07)	
Geographic region[c], n (%)			0.5237
North America	50 (34)	64 (29)	
Europe (Western, Central and Eastern)	69 (47)	104 (47)	
Asia/Australia/Africa	29 (20)	52 (24)	

Source: Table 14.1.1.5B.

Abbreviations: PLA = placebo; Std Dev = standard deviation; VDZ = vedolizumab.

a P-values are from chi-square test for categorical variables and from Kruskal Wallis Test for continuous variables.

b Age is defined as (1 + first dose date – birth date)/365.25

c The countries of each geographic region are specified in Table 14.1.1.3A.

Table 3 Baseline Crohn's Disease Characteristics – Induction Phase Safety Population – Study 13007

Crohn's Disease Characteristic	Induction Cohort 1 ITT Population[a] PLA N = 148	Induction Cohort 1 ITT Population[a] VDZ N = 220	Induction Cohort 2[b] Open-label VDZ N = 747	Combined VDZ N = 967	Total N = 1115
Duration of CD (yrs)[c]					
Mean (Std Dev)	8.2 (7.80)	9.2 (8.18)	9.2 (7.63)	9.2 (7.76)	9.0 (7.77)
Median	6.1	7.1	7.2	7.2	7.0
Minimum, maximum	0.3, 42.0	0.5, 43.6	0.2, 42.5	0.2, 43.6	0.2, 43.6
Duration of CD - categorical, n (%)					
< 1 year	12 (8)	12 (5)	45 (6)	57 (6)	69 (6)
≥ 1 - < 3 years	27 (18)	48 (22)	126 (17)	174 (18)	201 (18)
≥ 3 - < 7 years	45 (30)	49 (22)	191 (26)	240 (25)	285 (26)
≥ 7 years	64 (43)	111 (50)	385 (52)	496 (51)	560 (50)
Baseline disease activity – CDAI[d]					
n	147	219	743	962	1109
Mean (Std Dev)	324.6 (78.08)	327.3 (70.67)	322.2 (67.17)	323.4 (67.98)	323.6 (69.37)
Median	319.0	324.0	320.0	321.0	321.0
Minimum, maximum	155, 584	132, 500	93, 548	93, 548	93, 584
Baseline disease activity – categorical, n (%)					
CDAI ≤ 330	81 (55)	119 (54)	418 (56)	537 (56)	618 (55)
CDAI > 330	66 (45)	100 (45)	325 (44)	425 (44)	491 (44)
Missing	1	1	4	5	6
Baseline CRP (mg/L)					
n	147	220	747	967	1114
Mean (Std Dev)	23.6 (27.85)	24.1 (27.23)	20.4 (27.40)	21.2 (27.39)	21.5 (27.45)

Table 3 Baseline Crohn's Disease Characteristics – Induction Phase Safety Population – Study 13007 (continued)

Crohn's Disease Characteristic	Induction Cohort 1 ITT Population[a] PLA N = 148	Induction Cohort 1 ITT Population[a] VDZ N = 220	Induction Cohort 2[b] Open-label VDZ N = 747	Combined VDZ N = 967	Total N = 1115
Median	13.7	15.3	10.2	10.6	11.5
Minimum, maximum	0.2, 159.0	0.2, 164.0	0.2, 295.0	0.2, 295.0	0.2, 295
Baseline CRP - categorical, n (%)					
≤ 2.87 mg/L	20 (14)	37 (17)	130 (17)	167 (17)	187 (17)
> 2.87 to ≤ 5 mg/L	14 (9)	25 (11)	75 (10)	100 (10)	114 (10)
> 5 to ≤ 10 mg/L	28 (19)	38 (17)	160 (21)	198 (20)	226 (20)
> 10 mg/L	85 (57)	120 (55)	382 (51)	502 (52)	587 (53)
Missing	1	0	0	0	1
Baseline fecal calprotectin					
n	142	210	719	929	1071
Mean (Std Dev)	1421.2 (2076.11)	1839.9 (2624.92)	1050.1 (1558.93)	1228.7 (1881.84)	1254.2 (1908.82)
Median	652.6	852.2	656.8	688.3	685.8
Minimum, maximum	23.8, 12429.0	23.8, 13672.5	23.8, 18607.5	23.8, 18607.5	23.8, 18607.5
Baseline fecal calprotectin - categorical, n (%)					
≤ 250 µg/g	34 (23)	51 (23)	201 (27)	252 (26)	286 (26)
> 250 to ≤ 500 µg/g	27 (18)	25 (11)	112 (15)	137 (14)	164 (15)
> 500 µg/g	81 (55)	134 (61)	406 (54)	540 (56)	621 (56)
Missing	6	10	28	38	44
Disease localization, n (%)					
Ileum only	21 (14)	37 (17)	123 (16)	160 (17)	181 (16)
Colon only	43 (29)	62 (28)	211 (28)	273 (28)	316 (28)
Ileocolonic (both ileum and colon)	84 (57)	121 (55)	413 (55)	534 (55)	618 (55)
Other (extra ileum, extra colon)	0	0	0	0	0
History of prior surgery for CD, n (%)	54 (36)	98 (45)	314 (42)	412 (43)	466 (42)
History of fistulizing disease, n (%)	56 (38)	90 (41)	264 (35)	354 (37)	410 (37)
Draining fistula at baseline, n (%)					
Yes	23 (16)	38 (17)	104 (14)	142 (15)	165 (15)
All closed	2 (1)	1 (< 1)	8 (1)	9 (< 1)	11 (< 1)
No	123 (83)	181 (82)	635 (85)	816 (84)	939 (84)
Smoking status, n (%)					
Current smoker	34 (23)	54 (25)	210 (28)	264 (27)	298 (27)
Nonsmoker (never smoked)	85 (57)	120 (55)	351 (47)	471 (49)	556 (50)

Table 3 Baseline Crohn's Disease Characteristics – Induction Phase Safety Population – Study 13007 (continued)

Crohn's Disease Characteristic	Induction Cohort 1 ITT Population[a] PLA N = 148	Induction Cohort 1 ITT Population[a] VDZ N = 220	Induction Cohort 2[b] Open-label VDZ N = 747	Combined VDZ N = 967	Total N = 1115
Former smoker	29 (20)	46 (21)	185 (25)	231 (24)	260 (23)
Missing	0	0	1	1	1
Baseline extraintestinal manifestations, n (%)	107 (72)	133 (60)	456 (61)	589 (61)	696 (62)
History of extraintestinal manifestations, n (%)	123 (83)	177 (80)	619 (83)	796 (82)	919 (82)

Source: Table 14.1.1.6CP.

Abbreviations: CD = Crohn's disease; CDAI = Crohn's Disease Activity Index; CRP = C-reactive protein; ITT = intent-to-treat; PLA = placebo; Std Dev = standard deviation; VDZ = vedolizumab.

a All patients enrolled in Cohort 1 who were randomized to blinded induction treatment with vedolizumab or placebo.

b All patients enrolled in Cohort 2 who received open-label vedolizumab induction treatment.

c Duration of CD is defined as (1 + first dose date – diagnosis date)/365.25.

d Baseline disease activity represents the baseline CDAI score.

Table 4 Categorization of Patients by Prior TNTα Antagonist Use and Worst Prior Treatment Failure, Induction Phase Safety Population – Study C13007

Medication Use/Failure Failure Category	Induction Cohort 1 ITT Population[a]		Induction Cohort 2[b] Open-label	Combined	
	PLA N = 148	VDZ N = 220	VDZ N = 747	VDZ N = 967	Total N = 1115
Prior TNFα antagonist use[c], n (%)	72 (49)	111 (50)	506 (68)	617 (64)	689 (62)
No prior TNFα antagonist use[c], n (%)	76 (51)	109 (50)	241 (32)	350 (36)	426 (38)
Any prior TNFα antagonist failure[d], n (%)	70 (47)	105 (48)	470 (63)	575 (59)	645 (58)
Inadequate response[e]	41 (59)	56 (53)	223 (47)	279 (49)	320 (50)
Loss of response[f]	22 (31)	40 (38)	189 (40)	229 (40)	251 (39)
Intolerance[g]	7 (10)	9 (9)	58 (12)	67 (12)	74 (11)
Prior immunomodulators failure but no TNFα antagonist failure, n (%)	50 (34)	76 (35)	199 (27)	275 (29)	325 (29)
Inadequate response[e]	35 (70)	53 (70)	146 (73)	199 (72)	234 (72)
Intolerance[g]	15 (30)	23 (30)	53 (27)	76 (28)	91 (28)
Prior corticosteroid failure only, n (%)	27 (18)	36 (17)	72 (10)	108 (11)	135 (12)
Inadequate response[e]	23 (85)	31 (86)	66 (92)	97 (90)	120 (89)
Intolerance[g]	4 (15)	5 (14)	6 (8)	11 (10)	15 (11)

Source: Table 14.1.1.6CP, Table 14.1.1.12BP.

Abbreviations: CDRX = Crohn's Disease therapy; eCRF = electronic case report form; ITT = intent-to-treat; IVRS = interactive voice response system; PLA = placebo; TNFα = tumor necrosis factor alpha; VDZ = vedolizumab.

Each patient is counted in only 1 medication class with the worst outcome counted according to the following hierarchy: inadequate response considered worse than loss of response; loss of response considered worse than intolerance.

a All patients enrolled in Cohort 1 who were randomized to blinded induction treatment with vedolizumab or placebo.

b All patients enrolled in Cohort 2 who received open-label vedolizumab induction treatment.

c Data for prior TNFα antagonist use at randomization are obtained from the IVRS.

d n represents patients with information on prior treatment failure from the CDRX eCRF; 10 patients were missing prior treatment failure category and are excluded from the denominator for calculating percentages.

e Inadequate response to TNFα antagonists is defined as persistently active disease despite induction treatment (as listed in Section 4.2) with specified medications. For immunomodulators and corticosteroids, inadequate response includes patients who had an inadequate response, lost response (immunomodulators only), or who were being treated with these agents at the time of study entry and had active disease.

f Loss of response to TNFα antagonists is defined as recurrence of symptoms during maintenance dosing following prior clinical benefit.

g Intolerance is defined as occurrence of treatment-related toxicities (as listed in Section 4.2).

Table 5 Prior Therapies for Crohn's Disease – Induction Study ITT Population – Study C13007

Therapies[a], n (%)	PLA N = 148	VDZ N = 220	Total N = 368
Any prior therapies[b]	148 (100)	220 (100)	368 (100)
Any systemic corticosteroids	140 (95)	200 (91)	340 (92)
Only systemic corticosteroids	26 (18)	28 (13)	54 (15)
Any immunomodulators	113 (76)	174 (79)	287 (78)
Only immunomodulators	3 (2)	11 (5)	14 (4)
Any TNFα antagonists	75 (51)	117 (53)	192 (52)
Only TNFα antagonists	1 (< 1)	3 (1)	4 (1)
Immunomodulators and TNFα antagonists	66 (45)	101 (46)	167 (45)

Source: Table 14.1.1.8A.

Abbreviations: CD = Crohn's disease; eCRF = electronic case report form; IBD = inflammatory bowel disease; ITT = intent-to-treat; PLA = placebo; TNFα = tumor necrosis factor alpha; VDZ = vedolizumab.

"Any" refers to any exposure to the medication. "Only" refers to exclusive exposure to the medication (eg, no exposure to other medication for IBD).

a Patients may have been exposed to more than 1 prior therapy; therapies include prior or ongoing therapies for CD.

b As captured by the prior medications eCRF.

Table 6 Crohn's Disease Therapy at Baseline – Induction Phase Safety Population – Study C13007

Therapy at Baseline	Induction Cohort 1 ITT Population[a]		Induction Cohort 2[b] Open-label	Combined	
	PLA N = 148	VDZ N = 220	VDZ N = 747	VDZ N = 967	Total N = 1115
Corticosteroids[c], n (%)	71 (48)	105 (48)	394 (53)	499 (52)	570 (51)
Immunomodulators[d], n (%)	51 (34)	75 (34)	244 (33)	319 (33)	370 (33)
Corticosteroids only, n (%)	45 (30)	67 (30)	269 (36)	336 (35)	381 (34)
Immunomodulators only, n (%)	25 (17)	37 (17)	119 (16)	156 (16)	181 (16)
Corticosteroids and immunomodulators, n (%)	26 (18)	38 (17)	125 (17)	163 (17)	189 (17)
No corticosteroids or immunomodulators, n (%)	52 (35)	78 (35)	234 (31)	312 (32)	364 (33)

Source: Table 14.1.1.6B, Table 14.1.1.6CP. (Corticosteroid use for the open-label vedolizumab group, the combined vedolizumab group, and the overall total were derived by adding the numbers of patients who received corticosteroids only to those who received corticosteroids and immunomodulators. Similarly, immunomodulator use was derived by adding the numbers of patients who received immunomodulators only to those who received corticosteroids and immunomodulators.)

Abbreviations: ITT = intent-to-treat; IVRS = interactive voice response system; PLA = placebo; VDZ = vedolizumab.

Baseline CD medication use data were obtained from the IVRS for purposes of randomization stratification.

a All patients enrolled in Cohort 1 who were randomized to blinded induction treatment with vedolizumab or placebo.

b All patients enrolled in Cohort 2 who received open-label vedolizumab induction treatment.

c Corticosteroid use with and without immunomodulator use.

d Immunomodulator use with and without corticosteroid use.

Table 7 Comparisons by Treatment Group of Selected Baseline Crohn's Disease Characteristics – Induction Study ITT Population – Study C13007

Parameter	PLA N = 148	VDZ N = 220	P-value[a]
Duration of Crohn's disease (yrs)[b]			
Mean (Std Dev)	8.22 (7.803)	9.18 (8.184)	0.2052
Corticosteroid use at randomization, n (%)	71 (48)	105 (48)	0.9631
Immunomodulator use at randomization, n (%)	51 (34)	75 (34)	0.9418
Prior TNFα antagonist use[c], n (%)	72 (49)	111 (50)	0.7341
Baseline disease activity[d]			
Mean (Std Dev)	324.6 (78.08)	327.3 (70.67)	0.5209
Prior TNFα antagonist failure[e], n (%)	70 (47)	105 (48)	0.9355

Source: Table 14.1.1.6B.

Abbreviations: CD = Crohn's disease; CDAI = Crohn's Disease Activity Index; CDRX = Crohn's Disease therapy; eCRF = electronic case report form; ITT = intent-to-treat; IVRS = interactive voice response system; PLA = placebo; Std Dev = standard deviation; TNFα = tumor necrosis factor alpha; VDZ = vedolizumab.

a P-values are from chi-square test for categorical variables and from Kruskal Wallis Test for continuous variables.

b Duration of CD is defined as (1 + first dose date – diagnosis date)/365.25.

c Data for prior TNFα antagonist use were obtained from the IVRS.

d Baseline disease activity represents the baseline CDAI.

e Data for prior TNFα antagonist failure status were obtained from the CDRX eCRF.

Table 8 Clinical Remission at Week 6 – Per protocol Population – Study C13007

Table 14.3.1.2C
Clinical Remission at Week 6
Per Protocol Population

Clinical Remission [a]	PLA N=141	VDZ N=205
Number (%) Achieving Clinical Remission	9 (6.4)	30 (14.6)
95% CI	(2.3, 10.4)	(9.8, 19.5)
Difference from Placebo [b]		8.3
95% CI for Difference from Placebo		(1.6, 15.0)
P-value for Difference from Placebo [c]		0.0153
Relative Risk [d]		2.3
95% CI for Relative Risk		(1.1, 4.7)

(a) Clinical remission is defined as CDAI score <= 150 points.
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI
(c) P-value is based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no)
(d) Adjusted Relative Risk and its 95% CI.

Table 9 Table 8 Clinical Remission at Week 6 – Based on Patients Who Had Baseline and Week 6 Visit (Observed Case) – Study C13007

Table 14.3.1.2B
Clinical Remission at Week 6 - Based on Patients Who Had Baseline and Week 6 Visit (Observed Case)

Clinical Remission [a]	PLA N=136	VDZ N=200
Number (%) Achieving Clinical Remission	10 (7.4)	32 (16.0)
95% CI	(3.0, 11.7)	(10.9, 21.1)
Difference from Placebo [b]		8.7
95% CI for Difference from Placebo		(1.5, 15.8)
P-value for Difference from Placebo [c]		0.0174
Relative Risk [d]		2.2
95% CI for Relative Risk		(1.1, 4.3)

(a) Clinical remission is defined as CDAI score <= 150 points.
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI
(c) P-value is based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no)
(d) Adjusted Relative Risk and its 95% CI.

Table 10 Enhanced Clinical Response at Week 6 – Per protocol Population – Study C13007

Table 14.3.1.4C
Enhanced Clinical Response at Week 6
Per Protocol Population

Enhanced Clinical Response [a]	PLA N=141	VDZ N=205
Number (%) Achieving Enhanced Clinical Response	38 (27.0)	68 (33.2)
95% CI	(19.6, 34.3)	(26.7, 39.6)
Difference from Placebo [b]		6.4
95% CI for Difference from Placebo		(-3.3, 16.2)
P-value for Difference from Placebo [c]		0.1972
Relative Risk [d]		1.2
95% CI for Relative Risk		(0.9, 1.7)

(a) Enhanced clinical response is defined as a >= 100 point reduction in CDAI score from baseline
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI.
(c) P-value is based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no)
(d) Adjusted Relative Risk and its 95% CI.

Table 11 Enhanced Clinical Response at Week 6 – Based on Patients Who Had Baseline and Week 6 Visit (Observed Case) – Study C13007

Table 14.3.1.4B
Enhanced Clinical Response at Week 6 - Based on Patients Who Had Baseline and Week 6 Visit (Observed Case)

Enhanced Clinical Response [a]	PLA N=136	VDZ N=200
Number (%) Achieving Enhanced Clinical Response	38 (27.9)	69 (34.5)
95% CI	(20.4, 35.5)	(27.9, 41.1)
Difference from Placebo [b]		6.8
95% CI for Difference from Placebo		(-3.3, 16.8)
P-value for Difference from Placebo [c]		0.1871
Relative Risk [d]		1.2
95% CI for Relative Risk		(0.9, 1.7)

(a) Enhanced clinical response is defined as a >= 100 point reduction in CDAI score from baseline
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI.
(c) P-value is based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no)
(d) Adjusted Relative Risk and its 95% CI.

Table 12 Baseline Demographic – Maintenance Phase Safety Population – Study C13007

Parameter	Maintenance ITT[a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Maintenance Non-ITT		Combined	
	PLA N = 153	VDZ Q8W N = 154	VDZ Q4W N = 154	PLA[b] (from Week 0) N = 148	VDZ Q4W[c] (Week 6 Nonresponders) N = 506	PLA N = 301	VDZ N = 814
Gender, n (%)							
Male	72 (47)	68 (44)	82 (53)	69 (47)	229 (45)	141 (47)	379 (47)
Female	81 (53)	86 (56)	72 (47)	79 (53)	277 (55)	160 (53)	435 (53)
Race, n (%)							
White	140 (92)	136 (88)	134 (87)	124 (84)	461 (91)	264 (88)	731 (90)
Black	4 (3)	4 (3)	2 (1)	3 (2)	10 (2)	7 (2)	16 (2)
Asian	9 (6)	14 (9)	15 (10)	19 (13)	32 (6)	28 (9)	61 (7)
Other	0	0	3 (2)	2 (1)	3 (< 1)	2 (< 1)	6 (< 1)
Ethnicity, n (%)							
Hispanic or Latino	2 (1)	3 (2)	2 (1)	5 (3)	14 (3)	7 (2)	19 (2)
Not Hispanic or Latino	148 (97)	148 (96)	149 (97)	139 (94)	481 (95)	287 (95)	778 (96)
Not reported	3 (2)	3 (2)	3 (2)	4 (3)	11 (2)	7 (2)	17 (2)
Age[d] (yrs)							
Mean (Std Dev)	37.2 (11.95)	35.1 (12.23)	34.9 (12.20)	38.6 (13.16)	35.8 (11.70)	37.9 (12.56)	35.5 (11.89)
Median	36.0	32.5	32.7	36.7	33.5	36.2	33.1
Minimum, maximum	18, 68	18, 72	19, 77	19, 75	18, 76	18, 75	18, 77
Age (yrs), n (%)							
< 35	73 (48)	89 (58)	84 (55)	67 (45)	269 (53)	140 (47)	442 (54)
≥ 35	80 (52)	65 (42)	70 (45)	81 (55)	237 (47)	161 (53)	372 (46)
Age (yrs) n (%)							
< 65	149 (97)	151 (98)	152 (99)	142 (96)	498 (98)	291 (97)	801 (98)

Table 12 Baseline Demographic – Maintenance Phase Safety Population (continued) – Study C13007

Parameter	Maintenance ITT[a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Maintenance Non-ITT		Combined	
	PLA N = 153	VDZ Q8W N = 154	VDZ Q4W N = 154	PLA[b] (from Week 0) N = 148	VDZ Q4W[c] (Week 6 Nonresponders) N = 506	PLA N = 301	VDZ N = 814
≥ 65	4 (3)	3 (2)	2 (1)	6 (4)	8 (2)	10 (3)	13 (2)
Body weight (kg)							
Mean (Std Dev)	69.0 (18.15)	68.5 (18.56)	71.5 (18.38)	68.7 (18.90)	70.2 (20.49)	68.9 (18.50)	70.1 (19.75)
Median	67.0	64.0	68.5	66.0	66.8	66.7	66.2
Minimum, maximum	30, 124	33, 123	40, 148	32, 130	30, 167	30, 130	30, 167
BMI (kg/m^2)							
Mean (Std Dev)	24.0 (5.93)	23.6 (5.67)	24.2 (5.28)	23.7 (5.77)	24.0 (6.24)	23.9 (5.84)	24.0 (5.96)
Median	22.5	22.6	23.3	22.3	22.8	22.4	22.9
Minimum, maximum	14, 50	14, 46	15, 49	12, 45	13, 56	12, 50	13, 56
Geographic region[e], n (%)							
North America	37 (24)	58 (38)	47 (31)	50 (34)	213 (42)	87 (29)	318 (39)
Western/Northern Europe	54 (35)	30 (19)	39 (25)	22 (15)	115 (23)	76 (25)	184 (23)
Central Europe	35 (23)	31 (20)	32 (21)	30 (20)	80 (16)	65 (22)	143 (18)
Eastern Europe	9 (6)	13 (8)	12 (8)	17 (11)	39 (8)	26 (9)	64 (8)
Asia/Australia/Africa	18 (12)	22 (14)	24 (16)	29 (20)	59 (12)	47 (16)	105 (13)

Source: Table 14.1.1.5AM.

Baseline refers to Week 0.

Abbreviations: ITT = intent-to-treat; PLA = placebo; Q4W = dosing every 4 weeks; Q8W = dosing every 8 weeks; Std Dev = standard deviation; VDZ = vedolizumab.

a Maintenance ITT includes patients who received vedolizumab during the Induction Phase, determined to be responders to induction therapy, and were randomized to the Maintenance ITT Population at Week 6.

b Maintenance Non-ITT placebo includes patients who received placebo during the Induction Phase and were assigned to continue placebo during the Maintenance Phase.

c Maintenance Non-ITT vedolizumab Q4W includes patients who received vedolizumab in the Induction Phase, did not achieve clinical response at Week 6, and were assigned to receive vedolizumab Q4W during the Maintenance Phase.

d Age is defined as (1+first dose date-birth date)/365.25.

e The countries of each geographic region are specified in Table 14.1.1.3AM.

Table 13 Comparison by Treatment Group of Selected Baseline Demographic Characteristics – Maintenance Study Populations – Study C13007

Parameter	PLA N = 153	VDZ Q8W N = 154	VDZ Q4W N = 154	P-value[a]
Gender, n (%)				
Male	72 (47)	68 (44)	82 (53)	0.2646
Female	81 (53)	86 (56)	72 (47)	
Race, n (%)				
White	140 (92)	136 (88)	134 (87)	0.4350
Other	13 (8)	18 (12)	20 (13)	
Age (yrs)[b]				
Mean (Std Dev)	37.2 (11.95)	35.1 (12.23)	34.9 (12.20)	0.0926
Body weight (kg)				
Mean (Std Dev)	69.0 (18.15)	68.5 (18.56)	71.5 (18.38)	0.2242
Geographic region[c], n (%)				
North America	37 (24)	58 (38)	47 (31)	0.0610
Europe (Western/Northern, Central and Eastern)	98 (64)	74 (48)	83 (54)	
Asia/Australia/Africa	18 (12)	22 (14)	24 (16)	

Source: Table 14.1.1.5BM.

Abbreviations: ITT = intent-to-treat; PLA = placebo; Q4W = dosing every 4 weeks; Q8W = dosing every 8 weeks; Std Dev = standard deviation; VDZ = vedolizumab.

Baseline refers to Week 0.

a P-values for categorical variables are from Chi-Square Test and for continuous variables are from Kruskal Wallis Test.

b Age is defined as (1+first dose date – birth date)/ 365.25.

c The countries of each geographic region are specified in Table 14.1.1.3AM.

Table 14 Baseline Crohn's Disease Characteristics – Maintenance Phase Safety Population – Study 13007

	Maintenance ITT[a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Maintenance Non-ITT		Combined	
Disease Characteristic	PLA N = 153	VDZ Q8W N = 154	VDZ Q4W N = 154	PLA[b] (from Week 0) N = 148	VDZ Q4W[c] (Week 6 Nonresponders) N = 506	PLA N = 301	VDZ N = 814
Duration of Crohn's disease (yrs)[d]							
Mean (Std Dev)	9.6 (8.85)	8.4 (7.28)	7.7 (6.78)	8.2 (7.80)	9.7 (7.77)	8.9 (8.37)	9.1 (7.54)
Median	7.0	6.5	6.4	6.1	8.0	6.3	7.2
Minimum, maximum	0.3, 43.6	0.3, 34.7	0.2, 42.5	0.3, 42.0	0.3, 42.8	0.3, 43.6	0.2, 42.8
Duration of Crohn's disease - categorical, n (%)							
< 1 year	12 (8)	11 (7)	10 (6)	12 (8)	24 (5)	24 (8)	45 (6)
≥ 1 - < 3 years	27 (18)	32 (21)	39 (25)	27 (18)	76 (15)	54 (18)	147 (18)
≥ 3 - < 7 years	37 (24)	39 (25)	31 (20)	45 (30)	133 (26)	82 (27)	203 (25)
≥ 7 years	77 (50)	72 (47)	74 (48)	64 (43)	273 (54)	141 (47)	419 (51)
Baseline disease activity - CDAI[e]							
n	153	153	153	147	503	300	809
Mean (Std Dev)	325.2 (65.58)	325.5 (68.76)	317.0 (65.99)	324.6 (78.08)	324.2 (69.13)	324.9 (71.86)	323.1 (68.46)
Median	315.0	322.0	316.0	319.0	322.0	317.5	322.0
Minimum, maximum	166, 500	149, 486	132, 548	155, 584	93, 517	155, 584	93, 548
Baseline disease activity – categorical, n (%)							
CDAI ≤ 330	86 (56)	78 (51)	96 (62)	81 (55)	277 (55)	167 (55)	451 (55)
CDAI > 330	67 (44)	75 (49)	57 (37)	66 (45)	226 (45)	133 (44)	358 (44)
Missing	0	1	1	1	3	1	5
Baseline CRP (mg/L)							
n	153	154	154	147	506	300	814
Mean (Std Dev)	17.2 (21.86)	17.9 (29.47)	16.9 (18.68)	23.6 (27.85)	24.8 (29.93)	20.3 (25.14)	22.0 (28.26)
Median	9.8	8.6	9.8	13.7	14.0	12.7	10.6
Minimum, maximum	0.2, 165.0	0.2, 295.0	0.2, 118.0	0.2, 159.0	0.2, 234.0	0.2, 165.0	0.2, 295.0
Baseline CRP - categorical, n (%)							
≤ 2.87 mg/L	24 (16)	35 (23)	25 (16)	20 (14)	83 (16)	44 (15)	143 (18)
> 2.87 - ≤ 5 mg/L	23 (15)	15 (10)	20 (13)	14 (9)	42 (8)	37 (12)	77 (9)
> 5 - ≤ 10 mg/L	32 (21)	39 (25)	35 (23)	28 (19)	92 (18)	60 (20)	166 (20)
> 10 mg/L	74 (48)	65 (42)	74 (48)	85 (57)	289 (57)	159 (53)	428 (53)
Missing	0	0	0	1	0	1	0
Baseline fecal calprotectin							
n	150	148	148	142	483	292	779
Mean (Std Dev)	1142.5 (1429.34)	1044.6 (1502.03)	1219.3 (1784.00)	1421.2 (2076.11)	1314.7 (2123.14)	1278.0 (1775.96)	1245.3 (1957.32)
Median	683.7	583.5	776.3	652.6	702.0	662.1	689.4
Minimum, maximum	23.8, 7581.3	23.8, 9479.0	23.8, 11978.8	23.8, 12429.0	23.8, 18607.5	23.8, 12429.0	23.8, 18607.5
Baseline fecal calprotectin – categorical, n (%)							
≤ 250 µg/g	38 (25)	48 (31)	35 (23)	34 (23)	131 (26)	72 (24)	214 (26)
> 250 - ≤ 500 µg/g	30 (20)	22 (14)	14 (9)	27 (18)	71 (14)	57 (19)	107 (13)
> 500 µg/g	82 (54)	78 (51)	99 (64)	81 (55)	281 (56)	163 (54)	458 (56)

Table 14 Baseline Crohn's Disease Characteristics – Maintenance Phase Safety Population – Study 13007 (continued)

	Maintenance ITT[a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Maintenance Non-ITT		Combined	
Disease Characteristic	PLA N = 153	VDZ Q8W N = 154	VDZ Q4W N = 154	PLA[b] (from Week 0) N = 148	VDZ Q4W[c] (Week 6 Nonresponders) N = 506	PLA N = 301	VDZ N = 814
Missing	3	6	6	6	23	9	35
Disease localization, n (%)							
Ileum only	19 (12)	29 (19)	34 (22)	21 (14)	78 (15)	40 (13)	141 (17)
Colon only	43 (28)	27 (18)	47 (31)	43 (29)	156 (31)	86 (29)	230 (28)
Ileocolonic (both ileum and colon)	91 (59)	98 (64)	73 (47)	84 (57)	272 (54)	175 (58)	443 (54)
Other (extra ileum, extra colon)	0	0	0	0	0	0	0
History of prior surgery for Crohn's disease, n (%)	57 (37)	57 (37)	61 (40)	54 (36)	237 (47)	111 (37)	355 (44)
History of fistulizing disease, n (%)	57 (37)	47 (31)	49 (32)	56 (38)	201 (40)	113 (38)	297 (36)
Draining fistula at baseline, n (%)							
Yes	18 (12)	17 (11)	22 (14)	23 (16)	85 (17)	41 (14)	124 (15)
All closed	2 (1)	1 (< 1)	0	2 (1)	6 (1)	4 (1)	7 (< 1)
No fistula at baseline	133 (87)	136 (88)	132 (86)	123 (83)	415 (82)	256 (85)	683 (84)
Smoking status, n (%)							
Current smoker	48 (31)	48 (31)	39 (25)	34 (23)	129 (25)	82 (27)	216 (27)
Nonsmoker (never smoked)	64 (42)	74 (48)	77 (50)	85 (57)	256 (51)	149 (50)	407 (50)
Former smoker	41 (27)	31 (20)	38 (25)	29 (20)	121 (24)	70 (23)	190 (23)
Missing	0	1	0	0	0	0	1
Baseline extraintestinal manifestations, n (%)	95 (62)	87 (56)	91 (59)	107 (72)	316 (62)	202 (67)	494 (61)
History of extraintestinal manifestations, n (%)	125 (82)	124 (81)	124 (81)	123 (83)	423 (84)	248 (82)	671 (82)

Source: Table 14.1.1.6AM.

Abbreviations: CDAI = Crohn's Disease Activity Index; CRP= C-reactive protein; ITT = intent-to-treat; PLA = placebo; Q4W = dosing every 4 weeks; Q8W = dosing every 8 weeks; Std Dev = standard deviation; VDZ = vedolizumab.

Baseline refers to Week 0.

a Maintenance ITT includes patients who received vedolizumab during the Induction Phase, determined to be responders to induction therapy, and were randomized to the Maintenance ITT Population at Week 6.

b Maintenance Non-ITT placebo includes patients who received placebo during the Induction Phase and were assigned to continue placebo during the Maintenance Phase.

c Maintenance Non-ITT vedolizumab Q4W includes patients who received vedolizumab in the Induction Phase, did not achieve clinical response at Week 6, and were assigned to receive vedolizumab Q4W during the Maintenance Phase.

d Duration of Crohn's Disease is defined as (1 + first dose date - diagnosis date)/ 365.25.

e Baseline disease activity represents the baseline CDAI score.

Table 15 Categorization of Patients by Prior TNTα Antagonist Use and Worst Prior Treatment Failure, Maintenance Phase Safety Population – Study C13007

Medication Use/Failure Failure Category	Maintenance ITT[a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6) PLA N = 153	VDZ Q8W N = 154	VDZ Q4W N = 154	Maintenance Non-ITT PLA[b] (from Week 0) N = 148	VDZ Q4W[c] (Week 6 Nonresponders) N = 506	Combined PLA N = 301	VDZ N = 814
Prior TNFα antagonist use[d], n (%)	82 (54)	88 (57)	83 (54)	72 (49)	364 (72)	154 (51)	535 (66)
No prior TNFα antagonist use[d], n (%)	71 (46)	66 (43)	71 (46)	76 (51)	142 (28)	147 (49)	279 (34)
Any prior TNFα antagonist failure[e], n (%)	78 (51)	82 (55)	77 (50)	70 (48)	338 (67)	148 (49)	497 (62)
Inadequate response[f]	35 (45)	37 (45)	31 (40)	41 (59)	176 (52)	76 (51)	244 (49)
Loss of response[g]	29 (37)	35 (43)	33 (43)	22 (31)	132 (39)	51 (34)	200 (40)
Intolerance[h]	14 (18)	10 (12)	13 (17)	7 (10)	30 (9)	21 (14)	53 (11)
Prior immunomodulators failure but not TNFα antagonist failure, n (%)	49 (32)	48 (32)	54 (35)	50 (34)	124 (25)	99 (33)	226 (28)
Inadequate response[f]	34 (69)	29 (60)	38 (70)	35 (70)	98 (79)	69 (70)	165 (73)
Intolerance[h]	15 (31)	19 (40)	16 (30)	15 (30)	26 (21)	30 (30)	61 (27)
Prior corticosteroids failure only, n (%)	25 (16)	20 (13)	22 (14)	27 (18)	41 (8)	52 (17)	83 (10)
Inadequate response[f]	22 (88)	19 (95)	21 (95)	23 (85)	35 (85)	45 (87)	75 (90)
Intolerance[h]	3 (12)	1 (5)	1 (5)	4 (15)	6 (15)	7 (13)	8 (10)

Source: Table 14.1.1.6AM, Table 14.1.1.12AM.

Abbreviations: CDRX = Crohn's Disease therapy; eCRF = electronic case report form; ITT = intent-to-treat; IVRS = interactive voice response system; PLA = placebo; Q4W = dosing every 4 weeks; Q8W = dosing every 8 weeks; TNFα = tumor necrosis factor alpha; VDZ = vedolizumab.

a Maintenance ITT includes patients who received vedolizumab during the Induction Phase, determined to be responders to induction therapy, and were randomized to the Maintenance ITT Population at Week 6.

b Maintenance Non-ITT placebo includes patients who received placebo during the Induction Phase and were assigned to continue placebo during the Maintenance Phase.

c Maintenance Non-ITT vedolizumab Q4W includes patients who received vedolizumab in the Induction Phase, did not achieve clinical response at Week 6, and were assigned to receive vedolizumab Q4W during the Maintenance Phase.

d Data for prior TNFα antagonist use were obtained from the IVRS.

e n represents patients with information on prior treatment failure from the CDRX eCRF; 10 patients were missing prior treatment failure category because they did not meet per protocol failure criteria.

f Inadequate response to TNFα antagonists is defined as persistently active disease despite induction treatment (as defined by the protocol) with specified medications. For immunomodulators and corticosteroids, inadequate response includes patients who had an inadequate response, lost response (immunomodulators only), or who were being treated with these agents at the time of study entry and had active disease.

g Loss of response to TNFα antagonists is defined as recurrence of symptoms during maintenance dosing following prior clinical benefit.

h Intolerance is defined as occurrence of treatment-related toxicities (as defined by the protocol).

Table 16 Prior Therapies for Crohn's Disease – Maintenance Study ITT Population – Study C13007

	Maintenance ITT[a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Maintenance Non-ITT		Combined	
Therapies[b], n (%)	PLA N = 153	VDZ Q8W N = 154	VDZ Q4W N = 154	PLA[c] (from Week 0) N = 148	VDZ Q4W[d] (Week 6 Nonresponders) N = 506	PLA N = 301	VDZ N = 814
Any Prior Therapies[e]	153 (100)	154 (100)	154 (100)	148 (100)	506 (100)	301 (100)	814 (100)
Any Systemic Corticosteroids	144 (94)	140 (91)	142 (92)	140 (95)	462 (91)	284 (94)	744 (91)
Only Systemic Corticosteroids	16 (10)	18 (12)	17 (11)	26 (18)	32 (6)	42 (14)	67 (8)
Any Immunomodulators	128 (84)	124 (81)	128 (83)	113 (76)	435 (86)	241 (80)	687 (84)
Only Immunomodulators	4 (3)	4 (3)	6 (4)	3 (2)	17 (3)	7 (2)	27 (3)
Any TNFα antagonists	82 (54)	87 (56)	86 (56)	75 (51)	365 (72)	157 (52)	538 (66)
Only TNFα antagonists	1 (< 1)	3 (2)	1 (< 1)	1 (< 1)	5 (< 1)	2 (< 1)	9 (1)
Immunomodulators and TNFα antagonists	74 (48)	76 (49)	78 (51)	66 (45)	327 (65)	140 (47)	481 (59)

Source: Table 14.1.1.8AM.

Abbreviations: ITT = intent-to-treat; PLA = placebo; Q4W = dosing every 4 weeks; Q8W = dosing every 8 weeks; TNFα = tumor necrosis factor alpha; VDZ = vedolizumab.

"Any" refers to any exposure to the medication. "Only" refers to exclusive exposure to the medication (eg, no exposure to other medication for IBD).

a Maintenance ITT includes patients who received vedolizumab during the Induction Phase, determined to be responders to induction therapy, and were randomized to the Maintenance ITT Population at Week 6.

b Patients may have been exposed to more than 1 prior therapy; therapies include prior or ongoing therapies for Crohn's disease.

c Maintenance Non-ITT placebo includes patients who received placebo during the Induction Phase and were assigned to continue placebo during the Maintenance Phase.

d Maintenance Non-ITT vedolizumab Q4W includes patients who received vedolizumab in the Induction Phase, did not achieve clinical response at Week 6, and were assigned to receive vedolizumab Q4W during the Maintenance Phase.

e As captured by the prior medications eCRF.

Table 17 Crohn's Disease Therapy at Baseline – Maintenance Phase Safety Population – Study C13007

	Maintenance ITT[a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Maintenance Non-ITT		Combined	
Therapy at Baseline	PLA N = 153	VDZ Q8W N = 154	VDZ Q4W N = 154	PLA[b] (from Week 0) N = 148	VDZ Q4W[c] (Week 6 Nonresponders) N = 506	PLA N = 301	VDZ N = 814
Corticosteroids[d], n (%)	82 (54)	82 (53)	80 (52)	71 (48)	255 (50)	153 (51)	417 (51)
Immunomodulators[e], n (%)	49 (32)	50 (32)	53 (34)	51 (34)	167 (33)	100 (33)	270 (33)
Corticosteroids only, n (%)	56 (37)	59 (38)	58 (38)	45 (30)	163 (32)	101 (34)	280 (34)
Immunomodulators only, n (%)	23 (15)	27 (18)	31 (20)	25 (17)	75 (15)	48 (16)	133 (16)
Corticosteroids and immunomodulators, n (%)	26 (17)	23 (15)	22 (14)	26 (18)	92 (18)	52 (17)	137 (17)
No corticosteroids or immunomodulators, n (%)	48 (31)	45 (29)	43 (28)	52 (35)	176 (35)	100 (33)	264 (32)

Source: Table 14.1.1.6AM, Table 14.1.1.6BM. (Corticosteroid use for the Maintenance Non-ITT placebo and vedolizumab groups and the Combined placebo and vedolizumab groups were derived by adding the numbers of patients who received corticosteroids only to those who received corticosteroids and immunomodulators. Similarly, immunomodulator use was derived by adding the numbers of patients who received immunomodulators only to those who received corticosteroids and immunomodulators.)

Abbreviations: ITT = intent-to-treat; PLA = placebo; Q4W = dosing every 4 weeks; Q8W = dosing every 8 weeks; VDZ = vedolizumab.

Baseline refers to Week 0.

a Maintenance ITT includes patients who received vedolizumab during the Induction Phase, determined to be responders to induction therapy, and were randomized to the Maintenance ITT Population at Week 6.

b Maintenance Non-ITT placebo includes patients who received placebo during the Induction Phase and were assigned to continue placebo during the Maintenance Phase.

c Maintenance Non-ITT vedolizumab Q4W includes patients who received vedolizumab in the Induction Phase, did not achieve clinical response at Week 6, and were assigned to receive vedolizumab Q4W during the Maintenance Phase.

d Corticosteroid use with and without immunomodulator use.

e Immunomodulator use with and without corticosteroid use.

Table 18 Comparisons by Treatment Group of Selected Baseline Crohn's Disease Characteristics – Maintenance Study ITT Population – Study C13007

Parameter	PLA N = 153	VDZ Q8W N = 154	VDZ Q4W N = 154	P-value[a]
Duration of Crohn's disease (yrs)[b]				
Mean (Std Dev)	9.6 (8.85)	8.4 (7.28)	7.7 (6.78)	0.2797
Corticosteroid use at randomization, n (%)	82 (54)	82 (53)	80 (52)	0.9546
Immunomodulator use at randomization, n (%)	49 (32)	50 (32)	53 (34)	0.8937
Prior TNFα antagonist use[c], n (%)	82 (54)	88 (57)	83 (54)	0.7863
Baseline disease activity[d]				
Mean (Std Dev)	325.2 (65.58)	325.5 (68.76)	317.0 (65.99)	0.4413
Prior TNFα antagonist failure[e], n (%)	78 (51)	82 (53)	77 (50)	0.8429

Source: Table 14.1.1.6BM.

Abbreviations: CDAI = Crohn's Disease Activity Index; CDRX = Crohn's Disease therapy; eCRF = electronic case report form; ITT = intent-to-treat; IVRS = interactive voice response system; PLA = placebo; Q4W = dosing every 4 weeks; Q8W = dosing every 8 weeks; Std Dev = standard deviation; TNFα = tumor necrosis factor alpha; VDZ = vedolizumab.

a P-values for categorical variables are from Chi-Square Test and for continuous variables are from Kruskal Wallis Test.

b Duration of Crohn's Disease is defined as (1+first dose date – diagnosis date)/365.25.

c Data for prior TNFα antagonist use were obtained from the IVRS.

d Baseline (Week 0) disease activity represents the baseline (Week 0) CDAI score.

e Data for prior TNFα antagonist failure were obtained from the CDRX eCRF.

Table 19 Clinical Remission at Week 52 – Per protocol Population – Study C13007

Table 14.3.1.2CM
Clinical Remission at Week 52
Per Protocol Population

Clinical Remission [a]	PLA N=147	VDZ q8 wks N=149	VDZ q4 wks N=144
Number (%) Achieving Clinical Remission	33 (22.4)	56 (37.6)	55 (38.2)
95% CI	(15.7, 29.2)	(29.8, 45.4)	(30.3, 46.1)
Difference from Placebo [b]		15.0	15.9
95% CI for Difference from Placebo		(4.7, 25.3)	(5.4, 26.3)
P-value for Difference from Placebo [c]		0.0042	0.0029
Relative Risk [d]		1.7	1.7
95% CI for Relative Risk		(1.2, 2.4)	(1.2, 2.4)

(a) Clinical remission is defined as CDAI score <= 150 points.
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI.
(c) P-values are based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with 3 stratification factors: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no); 3) enrollment in cohort 1 or cohort 2 in the induction phase
(d) Adjusted Relative Risk and its 95% CI.

Table 20 Clinical Remission at Week 52 – Based on Patients Who Had Baseline and Week 52 Post-Baseline CDAI Assessment (Observed Case) – Study C13007

Table 14.3.1.2BM
Clinical Remission at Week 52 - Based on Patients who had Baseline and Week 52 Post-Baseline CDAI assessment (Observed Case)

Clinical Remission [a]	PLA N=63	VDZ q8 wks N=72	VDZ q4 wks N=81
Number (%) Achieving Clinical Remission	33 (52.4)	59 (81.9)	56 (69.1)
95% CI	(40.0, 64.7)	(73.1, 90.8)	(59.1, 79.2)
Difference from Placebo [b]		30.5	17.2
95% CI for Difference from Placebo		(15.2, 45.9)	(1.2, 33.3)
P-value for Difference from Placebo [c]		<0.0001	0.0350
Relative Risk [d]		1.6	1.3
95% CI for Relative Risk		(1.2, 2.1)	(1.0, 1.8)

(a) Clinical remission is defined as CDAI score <= 150 points.
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI.
(c) P-values are based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with 3 stratification factors: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no); 3) enrollment in cohort 1 or cohort 2 in the induction phase
(d) Adjusted Relative Risk and its 95% CI.

Table 21 Summaries of Differences in Clinical Response at Week 6 Based on IVRS vs. Reported by Treatment of Maintenance Phase Safety Population – Study C13007.

Table 14.3.1.37AM
Summary of Differences in Clinical Response at Week 6 Based on IVRS vs Reported by Treatment of Maintenance Phase Safety Population

	Clinical Response		ITT			Non-ITT		
Mis-categorized Patients	IVRS	Reported [a]	PLA N=153	VDZ q8 wks N=154	VDZ q4 wks N=154	PLA N=148	VDZ q4 wks N=506	Total N=1115
Overall			16 (10)	19 (12)	16 (10)	19 (13)	37 (7)	107 (10)
	Yes	No	16 (10)	19 (12)	16 (10)	14 (9)	1 (<1)	66 (6)
	No	Yes	0	0	0	5 (3)	36 (7)	41 (4)
By Cohort in Induction Phase								
Cohort 1	Yes	No	5 (3)	3 (2)	0	14 (9)	0	22 (2)
	No	Yes	0	0	0	5 (3)	8 (2)	13 (1)
Cohort 2	Yes	No	11 (7)	16 (10)	16 (10)	0	1 (<1)	44 (4)
	No	Yes	0	0	0	0	28 (6)	28 (3)

Table 22 Clinical Remission at Week 52 for Patients Who Achieved Clinical Response at Week 6 – ITT Population – Study C13007

Table 14.3.1.27BM
Clinical Remission at Week 52 For Patients Who Achieved Clinical Response at Week 6
Intent-to-Treat Population

Clinical Remission [a]	PLA N=137	VDZ q8 wks N=135	VDZ q4 wks N=138
Number (%) Achieving Clinical Remission	31 (22.6)	54 (40.0)	54 (39.1)
95% CI	(15.6, 29.6)	(31.7, 48.3)	(31.0, 47.3)
Difference from Placebo [b]		17.7	16.6
95% CI for Difference from Placebo		(6.8, 28.6)	(5.8, 27.4)
P-value for Difference from Placebo [c]		0.0014	0.0027
Relative Risk [d]		1.8	1.7
95% CI for Relative Risk		(1.2, 2.6)	(1.2, 2.5)

(a) Clinical remission is defined as CDAI score <= 150 points.
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI.
(c) P-values are based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with 3 stratification factors: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no); 3) enrollment in cohort 1 or cohort 2 in the induction phase
(d) Adjusted Relative Risk and its 95% CI.

Table 23 Enhanced Clinical Response at Week 52 for Patients Who Achieved Clinical Response at Week 6 – ITT Population – Study C13007

Table 14.3.1.27CM
Enhanced Clinical Response at Week 52 For Patients Who Achieved Clinical Response at Week 6
Intent-to-Treat Population

Clinical Remission [a]	PLA N=137	VDZ q8 wks N=135	VDZ q4 wks N=138
Number (%) Achieving Enhanced Clinical Response	42 (30.7)	62 (45.9)	67 (48.6)
95% CI	(22.9, 38.4)	(37.5, 54.3)	(40.2, 56.9)
Difference from Placebo [b]		15.4	18.0
95% CI for Difference from Placebo		(4.0, 26.9)	(6.6, 29.5)
P-value for Difference from Placebo [c]		0.0082	0.0021
Relative Risk [d]		1.5	1.6
95% CI for Relative Risk		(1.1, 2.0)	(1.2, 2.2)

(a) Enhanced clinical response is defined as a >=100 point decrease in CDAI score.
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI.
(c) P-values are based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with 3 stratification factors: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no); 3) enrollment in cohort 1 or cohort 2 in the induction phase
(d) Adjusted Relative Risk and its 95% CI.

Table 24 Baseline Demographic – Induction Phase Safety Population – Study C13011

	TNFα Antagonist Failure ITT Subpopulation			Overall ITT Population		
Parameter	PLA N = 157	VDZ N = 158	Total N = 315	PLA N = 207	VDZ N = 209	Total N = 416
Gender, n (%)						
Male	62 (39)	68 (43)	130 (41)	89 (43)	91 (44)	180 (43)
Female	95 (61)	90 (57)	185 (59)	118 (57)	118 (56)	236 (57)
Race, n (%)						
White	142 (90)	143 (91)	285 (90)	186 (90)	188 (90)	374 (90)
Black	5 (3)	4 (3)	9 (3)	5 (2)	4 (2)	9 (2)
Asian	3 (2)	5 (3)	8 (3)	9 (4)	9 (4)	18 (4)
Other	7 (4)	5 (3)	12 (4)	7 (3)	6 (3)	13 (3)
Not reported	0	1 (< 1)	1 (< 1)	0	2 (< 1)	2 (< 1)
Ethnicity, n (%)						
Hispanic or Latino	2 (1)	3 (2)	5 (2)	4 (2)	4 (2)	8 (2)
Not Hispanic or Latino	152 (97)	154 (97)	306 (97)	199 (96)	204 (98)	403 (97)
Not reported	3 (2)	1 (< 1)	4 (1)	4 (2)	1 (< 1)	5 (1)

Table 24 Baseline Demographic – Induction Phase Safety Population (continued) – Study C13011

	TNFα Antagonist Failure ITT Subpopulation			Overall ITT Population		
Parameter	PLA N = 157	VDZ N = 158	Total N = 315	PLA N = 207	VDZ N = 209	Total N = 416
Age (yrs)[a]						
Mean (Std Dev)	38.4 (13.81)	38.7 (12.15)	38.6 (12.98)	37.1 (13.15)	38.6 (12.14)	37.9 (12.66)
Median	36.6	37.5	37.1	34.8	36.9	36.2
Min, Max	19, 77	20, 69	19, 77	19, 77	20, 69	19, 77
Age (yrs), n (%)						
< 35	72 (46)	64 (41)	136 (43)	105 (51)	88 (42)	193 (46)
≥ 35	85 (54)	94 (59)	179 (57)	102 (49)	121 (58)	223 (54)
Age (yrs), n (%)						
< 65	152 (97)	155 (98)	307 (97)	202 (98)	206 (99)	408 (98)
≥ 65	5 (3)	3 (2)	8 (3)	5 (2)	3 (1)	8 (2)
Body weight (kg)						
Mean (Std Dev)	71.2 (19.14)	70.3 (18.97)	70.7 (19.03)	71.3 (19.22)	69.5 (17.76)	70.4 (18.50)
Median	65.3	66.5	66.0	66.7	66.0	66.2
Min, Max	41, 125	40, 144	40, 144	41, 147	40, 144	40, 147
BMI (kg/m^2)						
Mean (Std Dev)	24.8 (6.32)	24.1 (5.38)	24.5 (5.87)	24.6 (6.13)	24.0 (5.13)	24.3 (5.65)
Median	23.3	23.3	23.3	23.3	23.3	23.3
Min, Max	15, 48	15, 43	15, 48	15, 48	15, 43	15, 48
Geographic region, n (%)						
North America	90 (57)	84 (53)	174 (55)	95 (46)	102 (49)	197 (47)
Western/Northern Europe	32 (20)	33 (21)	65 (21)	37 (18)	38 (18)	75 (18)
Central Europe	17 (11)	20 (13)	37 (12)	46 (22)	41 (20)	87 (21)
Eastern Europe	14 (9)	10 (6)	24 (8)	15 (7)	10 (5)	25 (6)
Asia/Australia/ Africa	4 (3)	11 (7)	15 (5)	14 (7)	18 (9)	32 (8)

Source: Table 14.1.1.5AT, Table 14.1.1.5A.

Abbreviations: BMI = body mass index; ITT = intent-to-treat; Max = maximum; Min = minimum; PLA = placebo; Std Dev = standard deviation; TNFα = tumor necrosis factor alpha; VDZ = vedolizumab.

a Age is defined as (1+first dose date – birth date)/365.25.

Table 25 Comparison by Treatment Group of Selected Baseline Demographic Characteristics – Induction Study Populations – Study C13011

	TNFα Antagonist Failure ITT Subpopulation			Overall ITT Population		
Parameter	PLA N = 157	VDZ N = 158	P-value[a]	PLA N = 207	VDZ N = 209	P-value[a]
Gender, n (%)			0.5225			0.9106
Male	62 (39)	68 (43)		89 (43)	91 (44)	
Female	95 (61)	90 (57)		118 (57)	118 (56)	
Race, n (%)			0.9854			0.9738
White	142 (90)	143 (91)		186 (90)	188 (90)	
Other	15 (10)	15 (9)		21 (10)	21 (10)	
Age (yrs)[b]			0.5794			0.1179
Mean (Std Dev)	38.4 (13.81)	38.7 (12.15)		37.1 (13.15)	38.6 (12.14)	
Body weight (kg)			0.7883			0.5091
Mean (Std Dev)	71.2 (19.14)	70.3 (18.97)		71.3 (19.22)	69.5 (17.76)	
Geographic region, n (%)			0.1764			0.5565
North America	90 (57)	84 (53)		95 (46)	102 (49)	
Europe (Western, Northern, Central, Eastern)	63 (40)	63 (40)		98 (47)	89 (43)	
Asia/Australia/Africa	4 (3)	11 (7)		14 (7)	18 (9)	

Source: Table 14.1.1.5BT, Table 14.1.1.5B.

Abbreviations: ITT = intent-to-treat; PLA = placebo; Std Dev = standard deviation; TNFα = tumor necrosis factor alpha; VDZ = vedolizumab.

a P-values are from chi-square test for categorical variables and from Kruskal-Wallis test for continuous variables.

b Age is defined as (1+first dose date – birth date)/365.25.

Table 26 Baseline Crohn's Disease Characteristics – Induction Phase Safety Population – Study 13011

Crohn's Disease (CD) Characteristic	TNFα Antagonist Failure ITT Subpopulation			Overall ITT Population		
	PLA N = 157	VDZ N = 158	Total N = 315	PLA N = 207	VDZ N = 209	Total N = 416
Duration of CD (yrs)a						
Mean (Std Dev)	11.5 (8.09)	11.6 (8.64)	11.6 (8.36)	10.0 (7.98)	10.6 (8.75)	10.3 (8.37)
Median	9.6	9.4	9.5	8.0	8.4	8.0
Min, Max	1.0, 42.9	0.5, 41.8	0.5, 42.9	0.3, 42.9	0.3, 41.8	0.3, 42.9
Duration of CD – categorical, n (%)						
< 1 year	1 (< 1)	2 (1)	3 (< 1)	12 (6)	11 (5)	23 (6)
≥ 1 - < 3 years	12 (8)	17 (11)	29 (9)	25 (12)	28 (13)	53 (13)
≥ 3 - < 7 years	39 (25)	42 (27)	81 (26)	52 (25)	52 (25)	104 (25)
≥ 7 years	105 (67)	97 (61)	202 (64)	118 (57)	118 (56)	236 (57)

Table 26 Baseline Crohn's Disease Characteristics – Induction Phase Safety Population – Study 13011 (Continued)

Crohn's Disease (CD) Characteristic	TNFα Antagonist Failure ITT Subpopulation			Overall ITT Population		
	PLA N = 157	VDZ N = 158	Total N = 315	PLA N = 207	VDZ N = 209	Total N = 416
Baseline disease activity – CDAIb						
Mean (Std Dev)	306.1 (55.43)	316.1 (52.63)	311.1 (54.19)	301.3 (54.97)	313.9 (53.17)	307.7 (54.38)
Median	301.0	317.0	311.0	298.0	313.0	304.0
Min, Max	166, 564	196, 524	166, 564	166, 564	196, 524	166, 564
Baseline disease activity – categorical, n (%)						
CDAI ≤ 330	107 (68)	99 (63)	206 (65)	148 (71)	132 (63)	280 (67)
CDAI > 330	50 (32)	59 (37)	109 (35)	59 (29)	77 (37)	136 (33)
Baseline CRP (mg/L)						
Mean (Std Dev)	18.8 (23.58)	20.7 (24.70)	19.8 (24.13)	18.5 (21.98)	19.0 (23.17)	18.8 (22.56)
Median	9.4	10.1	9.7	10.5	9.7	9.8
Min, Max	0.2, 118.0	0.2, 168.0	0.2, 168.0	0.2, 118.0	0.2, 168.0	0.2, 168.0
Baseline CRP – categorical, n (%)						
≤ 2.87 mg/L	34 (22)	31 (20)	65 (21)	41 (20)	46 (22)	87 (21)
> 2.87 to ≤ 5 mg/L	16 (10)	11 (7)	27 (9)	19 (9)	14 (7)	33 (8)
> 5 to ≤ 10 mg/L	31 (20)	37 (23)	68 (22)	42 (20)	48 (23)	90 (22)
> 10 mg/L	76 (48)	79 (50)	155 (49)	105 (51)	101 (48)	206 (50)
Baseline fecal calprotectin (µg/g)						
N	157	154	311	206	204	410
Mean (Std Dev)	1459.5 (2475.01)	1249.2 (2071.60)	1355.3 (2282.93)	1426.5 (2357.76)	1148.1 (1878.58)	1288.0 (2134.79)
Median	647.0	693.6	658.0	665.4	618.3	656.8
Min, Max	23.8, 20000.0	23.8, 20000.0	23.8, 20000.0	23.8, 20000.0	23.8, 20000.0	23.8, 20000.0
Baseline fecal calprotectin categorical, n (%)						
≤ 250 µg/g	42 (27)	37 (23)	79 (25)	47 (23)	52 (25)	99 (24)

Table 26 Baseline Crohn's Disease Characteristics – Induction Phase Safety Population – Study 13011 (continued)

Crohn's Disease (CD) Characteristic	TNFα Antagonist Failure ITT Subpopulation			Overall ITT Population		
	PLA N = 157	VDZ N = 158	Total N = 315	PLA N = 207	VDZ N = 209	Total N = 416
> 250 to ≤ 500 µg/g	23 (15)	26 (16)	49 (16)	35 (17)	35 (17)	70 (17)
> 500 µg/g	92 (59)	91 (58)	183 (58)	124 (60)	117 (57)	241 (58)
Missing	0	4	4	1	5	6
Disease localization, n (%)						
Ileum only	20 (13)	21 (13)	41 (13)	29 (14)	33 (16)	62 (15)
Colon only	40 (25)	40 (25)	80 (25)	52 (25)	48 (23)	100 (24)
Ileocolonic (both ileum and colon)	97 (62)	97 (61)	194 (62)	126 (61)	128 (61)	254 (61)
Other (extra ileum, extra colon)	0	0	0	0	0	0
History of prior surgery for CD, n (%)	80 (51)	73 (46)	153 (49)	89 (43)	92 (44)	181 (44)
Smoking status, n (%)						
Current smoker	47 (30)	45 (28)	92 (29)	58 (28)	65 (31)	123 (30)
Never smoked	77 (49)	75 (47)	152 (48)	102 (49)	93 (44)	195 (47)
Former smoker	33 (21)	38 (24)	71 (23)	47 (23)	51 (24)	98 (24)
History of fistulizing disease, n (%)	67 (43)	57 (36)	124 (39)	77 (37)	71 (34)	148 (36)
Draining fistula at baseline, n (%)						
Yes	18 (11)	19 (12)	37 (12)	25 (12)	25 (12)	50 (12)
All closed	0	(< 1)	1 (< 1)	0	1 (< 1)	1 (< 1)
No fistula	139 (89)	138 (87)	277 (88)	182 (88)	183 (88)	365 (88)
Extraintestinal manifestations at baseline, n (%)	103 (66)	85 (54)	188 (60)	130 (63)	116 (56)	246 (59)

Source: Table 14.1.1.6AT, Table 14.1.1.6A.

Abbreviations: CD = Crohn's disease; CDAI = Crohn's Disease Activity Index; CRP = C-reactive protein; ITT = intent-to-treat; Max = maximum; Min = minimum; PLA = placebo; Std Dev = standard deviation; TNFα = tumor necrosis factor alpha; VDZ = vedolizumab.

a Duration of CD is defined as (1+first dose date – diagnosis date)/365.25.

b Baseline disease activity represents the baseline CDAI score.

Table 27 Categorization of Patients by Prior TNTα Antagonist Use and Worst Prior Treatment Failure, Induction Phase Safety Population – Study C13011

Medication Use/Failure Failure Category, n (%)	PLA N = 207	VDZ N = 209	Total N = 416
Prior TNFα antagonist failure (IVRS)[a]	157 (76)	158 (76)	315 (76)
Any prior TNFα antagonist failure (eCRF)[b]	156 (76)	155 (75)	311 (75)
Inadequate response[c]	69 (44)	66 (43)	135 (43)
Loss of response[d]	69 (44)	71 (46)	140 (45)
Intolerance[e]	18 (12)	18 (12)	36 (12)
Prior immunomodulator failure but no TNFα antagonist failure	45 (22)	44 (21)	89 (21)
Inadequate response[c]	28 (62)	33 (75)	61 (69)
Intolerance[e]	17 (38)	11 (25)	28 (31)
Prior corticosteroid failure only	5 (2)	9 (4)	14 (3)
Inadequate response[c]	5 (100)	8 (89)	13 (93)
Intolerance[e]	0	1 (11)	1 (7)

Source: Table 14.1.1.6A, Table 14.1.1.12A.

Abbreviations: CDRX = Crohn's disease therapy; eCRF = electronic case report form; ITT = intent-to-treat; IVRS = interactive voice response system; PLA = placebo; TNFα = tumor necrosis factor alpha; VDZ = vedolizumab.

Each patient is counted in only 1 medication class with the worst outcome counted according to the following hierarchy: inadequate response considered worse than loss of response; loss of response considered worse than intolerance.

a Data for prior TNFα antagonist failure at randomization are obtained from the IVRS.

b n represents patients with information on prior treatment failure from the CDRX eCRF; 2 patients were missing prior treatment failure category and are excluded from the denominator for calculating percentages.

c Inadequate response to TNFα antagonists is defined as persistently active disease despite induction treatment (as listed in Section 4.2) with specified medications. For immunomodulators and corticosteroids, inadequate response includes patients who had an inadequate response, lost response (immunomodulators only), or who were being treated with these agents at the time of study entry and had active disease.

d Loss of response to TNFα antagonists is defined as recurrence of symptoms during maintenance dosing following prior clinical benefit.

e Intolerance is defined as occurrence of treatment-related toxicities (as listed in Section 4.2).

Table 28 Crohn's Disease Therapy at Baseline – Induction Phase Safety Population – Study C13011

Therapy at Baseline, n (%)	TNFα Antagonist Failure ITT Subpopulation			Overall ITT Population		
	PLA N = 157	VDZ N = 158	Total N = 315	PLA N = 207	VDZ N = 209	Total N = 416
Corticosteroids[a]	85 (54)	86 (54)	171 (54)	108 (52)	110 (53)	218 (52)
Immunomodulators[b]	42 (27)	43 (27)	85 (27)	69 (33)	71 (34)	140 (34)
Corticosteroids only	61 (39)	62 (39)	123 (39)	72 (35)	73 (35)	145 (35)
Immunomodulators only	18 (11)	19 (12)	37 (12)	33 (16)	34 (16)	67 (16)
Corticosteroids and immunomodulators	24 (15)	24 (15)	48 (15)	36 (17)	37 (18)	73 (18)
No corticosteroids or immunomodulators	54 (34)	53 (34)	107 (34)	66 (32)	65 (31)	131 (31)

Source: Table 14.1.1.6AT, Table 14.1.1.6A, Table 14.1.1.6BT, Table 14.1.1.6B.

Abbreviations: ITT = intent-to-treat; PLA = placebo; TNFα = tumor necrosis factor alpha; VDZ = vedolizumab.

Baseline Crohn's disease medication use data were obtained from the IVRS for purposes of randomization stratification.

a Corticosteroid use with and without immunomodulator use.

b Immunomodulator use with and without corticosteroid use.

Table 29 Comparisons by Treatment Group of Selected Baseline Crohn's Disease Characteristics – Induction Study ITT Population – Study C13011

Crohn's Disease (CD) Characteristic	TNFα Antagonist Failure ITT Subpopulation			Overall ITT Population		
	PLA N = 157	VDZ N = 158	P-value[a]	PLA N = 207	VDZ N = 209	P-value[a]
Duration of CD (yrs)[b]						
Mean (Std Dev)	11.5 (8.09)	11.6 (8.64)	0.8864	10.0 (7.98)	10.6 (8.75)	0.6045
Baseline disease activity – CDAI[c]						
Mean (Std Dev)	306.1 (55.43)	316.1 (52.63)	0.0945	301.3 (54.97)	313.9 (53.17)	0.0153
Corticosteroid use at baseline[d], n (%)	85 (54)	86 (54)	0.9588	108 (52)	110 (53)	0.9255
Immunomodulator use at baseline[e], n (%)	42 (27)	43 (27)	0.9262	69 (33)	71 (34)	0.8905
Prior TNFα antagonist failure at randomization (IVRS), n (%)	NA	NA	NA	157 (76)	158 (76)	0.9531

Source: Table 14.1.1.6BT, Table 14.1.1.6B.

Abbreviations: CD = Crohn's disease; CDAI = Crohn's Disease Activity Index; ITT = intent-to-treat; IVRS = interactive voice response system; NA = not applicable; PLA = placebo; Std Dev = standard deviation; TNFα = tumor necrosis factor alpha; VDZ = vedolizumab.

a P-values are from chi-square test for categorical variables and from Kruskal-Wallis test for continuous variables.

b Duration of CD is defined as (1+first dose date – diagnosis date)/365.25.

c Baseline disease activity represents the baseline CDAI score.

d Corticosteroid use with and without immunomodulator use.

e Immunomodulator use with and without corticosteroid use.

Table 30 Clinical Remission at Week 6 – Anti-TNF Failure Per Protocol Population – Study C13011

Table 14.3.1.1CT
Clinical Remission at Week 6
Anti-TNF Failure Per Protocol Population

Clinical Remission [a]	PLA N=145	VDZ N=147
Number (%) Achieving Clinical Remission	18 (12.4)	24 (16.3)
95% CI	(7.0, 17.8)	(10.4, 22.3)
Difference from Placebo [b]		3.7
95% CI for Difference from Placebo		(-4.2, 11.6)
P-value for Difference from Placebo [c]		0.3626
Relative Risk [d]		1.3
95% CI for Relative Risk		(0.7, 2.3)

(a) Clinical remission is defined as CDAI score <= 150 points.
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI
(c) P values are based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with stratification according to: (1) concomitant use of oral corticosteroids (yes/no) ; and (2) concomitant immunomodulator use (yes/no)
(d) Adjusted Relative Risk and its 95% CI.

Table 31 Clinical Remission at Week 6 – Based on Patients who had a Baseline Week 6 Visit – Anti-TNF Failure Completers (Observed Case) Population – Study C13011

Table 14.3.1.1BT
Clinical Remission at Week 6 - Based on Patients who had a Baseline, Week 6 and Week 10 Visit
Anti-TNF Failure Completers (Observed Case) Population

Clinical Remission [a]	PLA N=137	VDZ N=147
Number (%) Achieving Clinical Remission	18 (13.1)	24 (16.3)
95% CI	(7.5, 18.8)	(10.4, 22.3)
Difference from Placebo [b]		3.1
95% CI for Difference from Placebo		(-5.1, 11.2)
P-value for Difference from Placebo [c]		0.4603
Relative Risk [d]		1.2
95% CI for Relative Risk		(0.7, 2.2)

Completers(Observed Case) population consists of all patients who received any amount of blinded study drug and have a baseline, week 6 and week 10 CDAI Score.
(a) Clinical remission is defined as CDAI score <= 150 points.
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI
(c) P values are based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with stratification according to: (1) concomitant use of oral corticosteroids (yes/no) ; and (2) concomitant immunomodulator use (yes/no)
(d) Adjusted Relative Risk and its 95% CI.

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

MILTON C FAN
05/19/2014

FREDA COONER
05/19/2014
See Statistical Team Leader Memorandum

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

BLA #:	125-476
Drug Name:	Entyvio (vedolizumab) (MLN002)
Indication:	For injection for the treatment of patients with moderately to severely active Ulcerative Colitis
Applicant:	Takeda Pharmaceuticals U.S.A., Inc.
Date:	Received June 20, 2013 PDUFA: February 18, 2014 (extended to May 20, 2014)
Review Priority:	Priority
Biometrics Division:	Division of Biometrics III
Statistical Reviewer:	Milton C. Fan, Ph.D, DB III
Concurring Reviewers:	Freda Cooner, Ph.D., Team Leader, DB III
Medical Division:	Gastroenterology and Inborn Errors Products (DGIEP)
Clinical Team:	Laurie Muldowney, M.D, Anil Rajpal, M.D. (TL) (DGIEP)
Project Manager:	Kevin Bugin (DGIEP)

Keywords: clinical studies, BLA review, placebo controlled

Table of Contents

1 EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

Only one pivotal study, Study C13006, was conducted to support Ulcerative Colitis (UC) for induction for vedolizumab 300 mg intravenous (IV) infusions. This study showed that in the Induction Phase, vedolizumab was statistically significant better than placebo in clinical response at Week 6 and clinical remission at Week 6 with treatment difference of 22% and 12%, respectively.

In Maintenance Phase, both vedolizumab dose regimens demonstrated statistically significant benefit compared to placebo in clinical remission at Week 52 and durable clinical response (clinical responses at both Weeks 6 and 52).

1.2 Brief Overview of Clinical Studies

1.2.1 Study C13006

This study was a pivotal, phase 3, multicenter, multinational, randomized, double-blind, placebo-controlled trial conducted to evaluate the efficacy and safety of both induction and maintenance treatment with vedolizumab in patients with moderately to severely active Ulcerative Colitis (UC), which was defined as a Mayo score of 6 to 12 points with an endoscopic subscore of ≥ 2.

This trial was designed to support the registration of vedolizumab for induction and maintenance treatment of a broad population of patients who have failed one or more standard therapies for UC, including corticosteroids, immunomodulators (azathioprine and 6-mercaptopurine), and TNFα antagonists.

This study was conducted at 211 sites worldwide. For study centers outside of the US, patients could have also failed treatment with corticosteroids. The applicant proposed that to ensure that the efficacy of vedolizumab could be evaluated in patients who were naïve to TNFα antagonists; enrollment of patients with previous TNFα antagonist exposure was limited to no more than 50% of the overall study population.

This study was consisted of two phases:

- The Induction Phase, designed to establish the efficacy and safety of vedolizumab for the induction of clinical response and remission, and

- The Maintenance Phase, designed to establish the efficacy and safety of vedolizumab for the maintenance of clinical response and remission.

All patients who completed the Induction Phase entered the Maintenance Phase. Treatment assignments were based on the Induction Phase treatment and the Investigator-assessed treatment response.

1.3 Statistical Issues and Findings

Induction Phase

Additional analyses have been conducted based on an alternative definition of clinical remission proposed by the FDA. The new definition is defined as a complete Mayo score of ≤ 2 points and no individual subscore > 1 point where rectal bleeding subscore = 0 and endoscopy subscore = 0. Results revealed that, fewer patients in either treatment group achieved this alternative definition of clinical remission.

The applicant defined mucosal healing as Mayo endoscopic subscore of ≤ 1 point and provided no histologic data to support a mucosal healing claim. Based on the applicant's definition, 40.9% of patients in the vedolizumab treatment group achieved mucosal healing, compared with 24.8% of patients receiving placebo, a 16.1% treatment difference (95% CI: 1.2, 2.3; p = 0.0012) was observed. When focusing only on the subset of patients who had an endoscopy subscore of 0 at Week 6, which indicates normal or inactive disease, there was no notable treatment difference observed 0.9%; (95 CI: -3.4, 5.1; p=0.6956) for the mucosal healing endpoint.

Per the medical officer's request, this review provided summary of this subgroup analysis for combined Induction Phase Cohorts (Cohort 1 and Cohort 2). Results show that the 95% confidence intervals overlapped for patients who had prior anti-TNF failures. For patients who did not have not prior anti-TNF failures, the 95% confidence intervals did not overlap.

All the analyses noted above were post-hoc sensitivity or subgroup analysis analyses. In general, the results from the sensitivity and subgroup analyses are consistently in favor of vedolizumab.

Maintenance Phase

To assess if the primary and secondary efficacy endpoints at Week 52 were affected by the Induction Phase Cohorts, additional analyses were requested during a Type C meeting held on July 24, 2012, after the phase 3 studies were completed.

Results from these analyses show that for clinical remission at Week 52 and durable clinical response at Week 52, vedolizumab in the every eight weeks (Q8W) and every four weeks (Q4W) treatment groups showed a treatment effect compared to placebo, regardless of whether patients were enrolled in Cohort 1 or 2 during the Induction Phase.

Per the medical officer's request, the applicant performed a subgroup analysis of clinical remission at Week 52 in subgroups based on anti-TNF failure status (inadequate response/loss of response). Results revealed that the 95% confidence intervals for the treatment differences of each of vedolizumab dose regimen from placebo included zero for patients who were prior anti-TNF failures. For patients who were not prior anti-TNF failures, the 95% confidence intervals excluded zero.

It should be noted that with more than 60% of data missing at Week 6 for placebo and more than 30% data missing at Week 52 for vedolizumab, the observed treatment difference might be overestimated when imputing all missing as non-responders. However, most of the missing data were due to lack of efficacy or adverse event.

2 INTRODUCTION

2.1 Overview

Vedolizumab is a humanized monoclonal antibody that binds to the α_4 β_7 integrin, which is expressed on discrete populations of leukocytes involved in gut mucosal immunity. The mechanism of action of MLN0002 reduces pathological bowel inflammation, thus providing a potential therapeutic option for patients with inflammatory bowel disease (IBD).

The applicant seeks marketing approval for the vedolizumab as an injection for the treatment of patients with moderately to severe active UC and Crohn's disease (CD).

This review is for the UC indication only and there is a separate statistical review for the CD indication.

2.2 Data Sources

The applicant has submitted three phase 3 studies (C13006, C13007, and C13011) for the proposed indication of injection for the treatment of patients with moderately to severely active UC or CD.

These three studies were entitled as follows:

- Clinical Protocol C13006: A Phase 3, Randomized, Placebo-Controlled, Blinded, Multicenter Study of the Induction and Maintenance of Clinical Response and Remission by Vedolizumab (MLN0002) in Patients with Moderate to Severe Ulcerative Colitis

- Clinical Protocol C13007: A Phase 3, Randomized, Placebo-Controlled, Blinded, Multicenter Study of the Induction and Maintenance of Clinical Response and Remission by Vedolizumab (MLN0002) in Patients with Moderate to Severe Crohn's Disease
- Clinical Protocol C13011: A Phase 3, Randomized, Placebo-controlled, Blinded, Multicenter Study of the Induction of Clinical Response and Remission by Vedolizumab (MLN0002) in Patients with Moderate to Severe Crohn's Diseases.

This review will focus on the study (C13006) for UC indication.

This original submission was submitted in eCTD dated June 20, 2013.

The electronic submission is located at
\\cdsesub1\bla\eCTD_Submissions\STN125476\0002.
The applicant submitted a response, dated September 9, 2013, to this reviewer's Information Request dated August 19, 2013.

The applicant submitted a response, dated October 4, 2013, to this reviewer's Information Request dated September 20, 2013.

The applicant submitted a response, dated October 21, 2013, to this reviewer's Information Request dated October 7, 2013.

3 STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

3.1.1 Study C13006

3.1.1.1 Study Design

This study was a pivotal, phase 3, multicenter, multinational, randomized, double-blind, placebo-controlled trial conducted to evaluate the efficacy and safety of both induction and maintenance treatment with vedolizumab in patients with moderately to severely active UC. Moderate to severe UC, is defined as a Mayo score of 6 to 12 points with an endoscopic subscore of ≥ 2 in this study.

This study was conducted at 211 sites. This trial was designed to support the registration of vedolizumab for induction and maintenance treatment of a broad population of patients who have failed one or more standard therapies for UC, including corticosteroids, immunomodulators (azathioprine and 6-mercaptopurine), and TNFα antagonists.

For study centers outside of the US, patients could have also failed treatment with corticosteroids. To ensure that the efficacy of vedolizumab could be evaluated in patients who are naïve to TNFα antagonists, enrollment of patients with previous TNFα antagonist exposure was limited to no more than 50% of the overall study population.

This study consisted of 2 phases:

- The Induction Phase, designed to establish the efficacy and safety of vedolizumab for the induction of clinical response and remission, and

- The Maintenance Phase, designed to establish the efficacy and safety of vedolizumab for the maintenance of clinical response and remission.

All patients who completed the Induction Phase entered the Maintenance Phase. Treatment assignments were based on the Induction Phase treatment and the Investigator-assessed treatment response.

3.1.1.2 Applicant's Analyses

3.1.1.2.1 Induction Phase

The 6-week Induction Phase contained two cohorts of patients: Cohort 1 patients were randomized and treated with study drug in a double-blind fashion, and Cohort 2 patients were treated with vedolizumab in an open-label fashion. The cohorts in the Induction Phase were enrolled sequentially, i.e., patients were enrolled in Cohort 2 after enrollment in Cohort 1 was complete. The eligibility criteria for both cohorts were identical. In Cohort 1, eligible UC patients who met the entry criteria were randomized to treatment with double-blind vedolizumab or placebo in a 3:2 ratio. The randomization was stratified for two factors that were specified as markers of disease severity:

- Concomitant use of oral corticosteroids
- Previous exposure to TNFα antagonists or concomitant immunomodulator (6-mercaptopurine or azathioprine) use

Randomized patients were treated with infusions of the double-blind study drug at Weeks 0 and 2. These patients comprised the population evaluated for efficacy and are referred to as the Induction ITT population. Randomization occurred via a central randomization interactive voice response system (IVRS). Treatment assignment was obtained from the system by the (unblinded) site pharmacist, who prepared the study drug and provided it to the site personnel (who remained blinded) in masked infusion bags.

The second cohort of patients was enrolled into the Induction Phase to ensure that the sample size of Induction responders randomized into the Maintenance Study could provide sufficient power for the Maintenance Study primary efficacy analysis. These patients did not contribute to the efficacy analyses done in the Induction Study. All patients in Cohort 2 were treated with open-label vedolizumab 300 mg, administered at Weeks 0 and 2. Patients in both cohorts were assessed for treatment response at Week 6.

3.1.1.2.1.1 Pre-specified Analyses

The primary efficacy assessment of the Induction Study tested for differences in the proportions of patients in the vedolizumab treatment regimen versus placebo who had clinical response using the complete Mayo score at Week 6. Clinical response by complete Mayo score was defined as a reduction in complete Mayo score of ≥ 3 points and ≥ 30% from baseline with an accompanying decrease in rectal bleeding subscore of ≥ 1 point or absolute rectal bleeding subscore of ≤1 point.

The primary comparison of the Induction Phase was performed using the Cochran-Mantel-Haenszel (CMH) chi-square test at a 5% significance level, with stratification according to the stratification factors (concomitant use of oral corticosteroids and previous exposure to TNFα antagonists or concomitant immunomodulator [6-mercaptopurine or azathioprine] use). The CMH chi-square p-value and the risk difference, along with its 95% confidence interval (CI) are provided. In addition, the relative risk and the 95% two-sided confidence interval (CI) are provided.

In addition to the primary comparisons, there were two secondary assessments of clinical efficacy (clinical remission and mucosal healing), which compared treatment differences between vedolizumab and placebo through formal closed testing procedures. Clinical remission was defined as a complete Mayo score of ≤ 2 and no individual subscore > 1 point. Mucosal healing was defined as Mayo endoscopic subscore of ≤ 1.

To maintain the overall Type I error rate at 5%, the key secondary assessments were performed sequentially (closed sequential method). The first secondary endpoint was to be tested only if the primary comparison was significant and the second key secondary endpoint was to be tested only if the first secondary endpoint was significant for vedolizumab. The testing order of the key secondary endpoints was finalized in the statistical analysis plan (SAP) before any formal unblinded data analyses.

The proportion-based key secondary endpoints were analyzed in the same fashion as the primary endpoint. The CMH chi-square p values and the relative risk estimates along with the 95% confidence intervals are provided. In addition, the absolute treatment differences in proportions are provided along with the 95% confidence intervals (CIs).

Disease activity for entry into this study and for efficacy assessments throughout the study was measured by the Mayo Score, a standard assessment tool to measure UC disease activity in clinical trials. The index consists of four components: two that are patient reported (rectal bleeding and stool frequency), a global assessment by the physician, and endoscopic subscore. The patient-reported components of the score were requested to be obtained on a daily basis from patients using the IVRS.

The major inclusion criteria are:

- Diagnosis of UC established at least six months prior to enrollment by clinical and endoscopic evidence and corroborated by a histopathology report
- Moderately to severely active UC as determined by a Mayo score of 6 to 12 with an endoscopic subscore ≥ 2 within seven days prior to the first dose of study drug
- Evidence of UC extending proximal to the rectum (≥ 15 cm of involved colon)
- Patients with extensive colitis or pancolitis of > 8 years duration or left-sided colitis of > 12 years duration must have documented evidence that a surveillance colonoscopy was performed within 12 months of the initial screening visit (may be performed during screening)
- Patients with a family history of colorectal cancer, personal history of increased colorectal cancer risk, age > 50 years, or other known risk factor must be up-to-date on colorectal cancer surveillance (may be performed during screening)
- Demonstrated, over the previous 5-year period, an inadequate response to, loss of response to, or intolerance of at least 1 of the following agents as defined below:

- Immunomodulators
    - Signs and symptoms of persistently active disease despite a history of at least one 8-week regimen of oral azathioprine (≥ 1.5 mg/kg) or 6-mercaptopurine (≥ 0.75 mg/kg) *OR*
    - History of intolerance of at least one immunomodulator (including but not limited to nausea/vomiting, abdominal pain, pancreatitis, liver function test abnormalities, lymphopenia, *TPMT* genetic mutation, infection)

- TNFα antagonists
    - Signs and symptoms of persistently active disease despite a history of at least one 4-week induction regimen of infliximab 5 mg/kg IV, 2 doses at least 2 weeks apart *OR*
    - Recurrence of symptoms during maintenance dosing following prior clinical benefit (discontinuation despite clinical benefit does not qualify) *OR*
    - History of intolerance of infliximab (including but not limited to infusion-related reaction, demyelination, congestive heart failure, infection)

ONLY APPLICABLE TO PATIENTS OUTSIDE THE US *(who may have been enrolled on the basis of corticosteroid treatment history):*

Corticosteroids

- Signs and symptoms of persistently active disease despite a history of at least one 4-week induction regimen that included a dose equivalent to prednisone 30 mg daily orally for 2 weeks or IV for 1 week, ***OR***
- Two failed attempts to taper corticosteroids to below a dose equivalent to prednisone 10 mg daily orally on 2 separate occasions, ***OR***
- History of intolerance of corticosteroids (including, but not limited to, Cushing's syndrome, osteopenia/osteoporosis, hyperglycemia, insomnia, and infection).

The Primary and secondary efficacy assessments for the Induction and Maintenance Phases were based on the Mayo scores. A complete Mayo score was obtained during screening, using patient diary entries within the 10 days prior to enrollment and flexible sigmoidoscopy results within 7 days prior to enrollment; this assessment was the baseline complete Mayo score. Sigmoidoscopy was done at Weeks 6 (prior to dosing) and a complete Mayo score was calculated for these visits. A validated IVRS was used for collection of the patient-reported outcome components of the complete Mayo Score.

A baseline (Week 0) complete Mayo score, calculated by adding the screening endoscopy subscore to the partial Mayo score obtained on Day 1 (Week 0), was used for the comparison with the Week 6 complete Mayo score to determine response and remission at Week 6.

The Week 6 complete Mayo score was calculated by the investigator or designee and recorded in the patient's source documents; this assessment determined whether the patient had achieved clinical response at Week 6, and, therefore, determined treatment assignment in the Maintenance Phase.

A partial Mayo score was to be derived at the visits where sigmoidoscopy was not performed (i.e., Weeks 0, 2, 4), and at any unscheduled visit(s) due to disease exacerbation. These scores were used to determine clinical response or disease worsening during the study.

Approximately 826 patients were planned to be enrolled into this study from approximately 300 sites worldwide. Enrollment was defined as the point in time at which the patient was assigned a treatment in the Induction Phase. An initial cohort (Cohort 1) of 375 patients was to be randomized in the Induction Phase, based on the sample size requirements of the Induction Study. Approximately 451 patients were then to be enrolled in Cohort 2. The number of patients to be enrolled in Cohort 2 was determined by the sample size requirements of the Maintenance Study. The protocol allowed for up to 100 additional patients to be enrolled into Cohort 2 (increasing the total number of study participants to 926), depending on the observed overall response rate in the combined cohorts, to ensure that at least 372 patients with clinical response at Week 6 to vedolizumab treatment were randomized in the Maintenance Phase. The Data and Safety Monitoring Board (DSMB) monitored the overall response and attrition rate (i.e., patients who were not willing to participate in the Maintenance Phase) to determine if additional enrollment of patients was required to ensure that the sample size for the Maintenance Study was achieved.

A total sample size of 826 was planned for the Induction Study. An initial cohort of 375 patients (Cohort 1) was randomized in a 3:2 ratio to receive vedolizumab (n = 225) or placebo (n = 150) in a double-blind manner. Following the randomization of this first cohort of 375 patients, 451 patients were planned to be enrolled into Cohort 2, and were to receive open-label vedolizumab induction dosing. Cohort 2 was necessary to provide sufficient power for the Maintenance analyses and was not included in the formal efficacy analysis of the Induction Study.

Power estimates for the primary and key secondary efficacy endpoints for the Induction Study are based on a total sample size of 375 patients at a 5% significance level and are provided in the table below.

Table 1 Power Estimates for Primary and Key Secondary Efficacy Analyses in Analyses in the Induction Study (Cohort 1)
Study C13006

Objective	Endpoint at Week 6 Vedolizumab vs Placebo	Assumed Response Rates	Sample Size per Group [a]	Power
Primary	Response	Placebo = 35% Vedolizumab = 53%	Placebo = 150 Vedolizumab = 225	93%
Key secondary	Remission	Placebo = 15% Vedolizumab = 29%	Placebo = 150 Vedolizumab = 225	89%
	Mucosal healing	Placebo = 8% Vedolizumab = 20%	Placebo = 150 Vedolizumab = 225	90%

Copied from Table 4, page 76 CSR.

3.1.1.2.1.2 Patient Disposition

The detailed patient disposition is given below.

Table 2 Patient Disposition– Induction Phase
Study C13006

	Induction Study ITT[a]		Non-ITT		
	PLA N = 149	VDZ Cohort 1 N = 225	VDZ Cohort 2 [b] N = 521	VDZ Combined N = 746	Total N = 895
Randomized/assigned	149	225	521	746	895
Study populations, n (%)					
Safety [c]	149 (100)	225 (100)	521 (100)	746 (100)	895 (100)
Intent-to-treat [d]	149 (100)	225 (100)	–	225 (30)	374 (42)
Per-Protocol [e]	138 (93)	215 (96)	–	215 (29)	353 (39)
Completed Induction Phase, n (%)[f]	135 (91)	218 (97)	485 (93)	703 (94)	838 (94)
Discontinued (reason)	14 (9)	7 (3)	36 (7)	43 (6)	57 (6)
Adverse event	4 (3)	0	7 (1)	7 (< 1)	11 (1)
Protocol violation(s)	1 (< 1)	1 (< 1)	6 (1)	7 (< 1)	8 (< 1)
Lack of efficacy	5 (3)	2 (< 1)	14 (3)	16 (2)	21 (2)
Study terminated by sponsor	0	0	0	0	0
Withdrawal of consent	3 (2)	4 (2)	8 (2)	12 (2)	15 (2)
Lost to follow-up	1 (< 1)	0	1 (< 1)	1 (< 1)	2 (< 1)
Other	0	0	0	0	0

Copied from Table 6, page 107 CSR.

Inclusion and exclusion criteria not met at Induction Phase entry are summarized for the Induction Study ITT population in the table below.

Table 3 Induction and Exclusion Criteria Not Met –Induction Study ITT Population Study C13006

Type of Unmet Criteria,[a] n (%)	PLA N = 149	VDZ N = 225	Total N = 374
Patients with at Least One Unmet Entry Criterion	13	11	24
Inclusion Criteria			
Inadequate or lost response/intolerance of steroids, immunomodulators and/or TNFα antagonists	3 (2)	5 (2)	8 (2)
Criteria for stability of corticosteroid dosing prior to enrollment [b]	2 (1)	2 (< 1)	4 (1)
UC diagnosed ≥6 months prior to enrollment	2 (1)	0	2 (< 1)
Gastrointestinal Exclusion Criteria			
C.difficile infection within 60 days, or other intestinal pathogen within 30 days prior to enrollment	5 (3)	0	5 (1)
5-ASA or steroid enemas/suppositories within 2 weeks of first dose	0	1 (< 1)	1 (< 1)
Use of non-biologic therapies (eg, cyclosporine, thalidomide) for the treatment of UC within 30 days prior to enrollment	0	1 (< 1)	1 (< 1)
History or evidence of colonic mucosal dysplasia	0	1 (< 1)	1 (< 1)
Infectious Disease Exclusion Criteria			
Positive TB test within 1 month prior to enrollment	2 (1)	1 (< 1)	3 (< 1)
Chest x-ray evidence of active or latent TB within 3 months prior to enrollment	1 (< 1)	0	1 (< 1)
General Exclusion Criteria			
History of major neurological disorders	0	1 (< 1)	1 (< 1)

Copied from Table 7, page 108 CSR.

A total of 24 ITT patients had at least one unmet entry criterion – 13 patients (8.7%) from the placebo group and 11 (4.9%) patients from the vedolizumab group. The most frequent reason was failure to meet the criteria for inadequate or lost response/intolerance of steroids, immunomodulators, and/ or TNFα antagonists, which occurred in eight patients. Overall, there were no notable trends between treatment groups. An additional 34 patients (6.5%) in Cohort 2 had violations of the inclusion/exclusion criteria; again, no notable trends were observed.

Deviations leading to exclusion from the Induction Study Per-Protocol population are summarized for the Induction Study ITT population in the table below.

Table 4 Protocol Deviations Leading to Exclusion from the Per-Protocol Population Induction Study ITT Population Study C13006

Type of Deviation,[a] n (%)	PLA N = 149	VDZ N = 225	Total N = 374
Patients with at least 1 protocol deviation	11	10	21
Baseline Mayo score < 6 or endoscopic subscore < 2 or duration of disease < 0.5 years	2 (1)	1 (< 1)	3 (< 1)
Received incorrect study medication as assigned at any study visit	0	0	0
Received < 2 doses of study medication	4 (3)	2 (< 1)	6 (2)
Received concomitant corticosteroids or other potentially effective medications for unrelated comorbid condition	1 (< 1)	0	1 (< 1)
Invalid Day 43/ET assessment [b]	8 (5)	8 (4)	16 (4)
Patients who had blind broken	0	0	0

Copied from Table 8, page 109 CSR.

A total of 21 patients [11 patients (7.4%) from the placebo group and 10 patients (4.4%) from the ITT vedolizumab group] had at least one protocol deviation and are excluded from the Per-Protocol population. For both groups, an invalid Day 43/Early Termination (ET) assessment was the most common protocol deviation. An invalid Day 43/ET assessment may have been due to either a clinical assessment outside Days 36 to 56 (inclusive) or a sigmoidoscopy performed outside Days 29 to 56 (inclusive).

3.1.1.2.1.3 Treatment Group Comparability

Baseline demographic characteristics of the Induction Phase Safety population are summarized by treatment group in Tables 1 and 2 in Appendix. ITT population was consisted of all patients who were randomized in Cohort 1 (double-blind). Non-ITT population was consisted of all patients who were enrolled in Chort2 (open-label).

As seen from Tables 1 and 2 in Appendix, overall, baseline demographics were similar for vedolizumab and placebo patients in the ITT population. There were no significant differences between treatment groups in selected demographic characteristics of patients randomized to placebo versus patients randomized to vedolizumab in the Induction Study ITT population.

Baseline UC disease characteristics of the Induction Phase Safety population are summarized in Appendix Table 3.

As seen from Appendix Table 3, the Induction Study ITT population treatment groups were comparable with respect to disease characteristics. The baseline disease characteristics of the Cohort 2 patients were similar to those of patients in Cohort 1.

The prior use of TNFα antagonists and treatment failure to UC therapies are summarized for the Induction Phase Safety population in Appendix Table 4.

Information regarding prior use of UC medications, previous treatment failure, and concomitant medications was captured at different time points during the study (screening in the IVRS system; Week 0 on the UCRX eCRF; and during the study on the concomitant medications eCRF). Therefore, methods used to collect this information may have resulted in inconsistencies in the numbers of patients in this table and subsequent summaries of baseline and concomitant medication use. Enrollment of patients with prior TNFα antagonist use was limited to no more than 50%; in the entire study population, 52% of patients were TNFα antagonist naive.

A hierarchical approach was used to categorize treatment failure to TNFα antagonists, immunomodulators, and corticosteroids. TNFα antagonist failure was prioritized over failure to immunomodulators, which was prioritized over failure of corticosteroids. Within each treatment category, patients were categorized by type of failure to a particular agent, per protocol definition. For TNFα antagonists, patients were categorized as having had an inadequate response (persistently active disease despite induction treatment), loss of response (recurrence of symptoms during maintenance treatment following prior clinical benefit), or intolerance (treatment-related toxicity). For immunomodulators and corticosteroids, treatment failure was categorized as either inadequate response (persistently active disease despite a 4-week regimen of corticosteroids or an 8-week regimen of immunomodulators) or intolerance, using similar definitions. As patients may have had more than one definition of treatment failure, only one category was assigned to each patient. Worst treatment failure was assigned using a hierarchical approach, with inadequate response considered worse than loss of response, and loss of response worse than intolerance. Using this approach, approximately 40% of patients had a history of failure to TNFα antagonists and a similar proportion had failed immunomodulators (without TNFα antagonist failure). Fewer patients failed corticosteroids alone (17%). In patients who had failed a TNFα antagonist, approximately half had inadequate response and approximately 40% had loss of response. The treatment groups were comparable with respect to the extent and nature of treatment failure to UC therapies.

Baseline UC therapy is summarized for the Induction Phase Safety population in Appendix Table 5.

As seen from Appendix Table 5, UC therapy use at baseline was similar between treatment groups and cohorts.

Appendix Table 6 shows comparisons of selected baseline UC disease characteristics in the Induction Study ITT population treatment groups.

As seen from Appendix Table 6, there were no significant differences in baseline UC disease characteristics between treatment groups for the duration of UC, baseline disease activity, use of concomitant therapies at baseline (corticosteroids or immunomodulators), or prior TNFα antagonist use.

3.1.1.2.1.4 Applicant's Analysis of Primary Efficacy Endpoint

The primary endpoint for the Induction Phase is the proportion of patients with clinical response at Week 6.Clinical response is defined as reduction in complete Mayo score of ≥ 3 points and ≥ 30% from baseline (Week 0) with an accompanying decrease in rectal bleeding subscore of ≥ 1 point or absolute rectal bleeding subscore of ≤ 1 point.

The CMH chi-square test was performed with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to TNFα antagonists and/or concomitant immunomodulator use (yes/no).

The results from analysis of the proportion of patients with clinical response at Week 6 are given below.

Table 5 Clinical Response at Week 6 by Cohort

Clinical response	Placebo (n=149)	VDZ (n=225)	VDZ open-label (n=521)
N (%) achieve clinical response at week 6	38 (25.5)	106 (47.1)	231 (44.3)
95% CI	(18.5, 32.5)	(40.6, 53.6)	(40.1, 48.6)
Difference from placebo		21.7	
95% CI for difference from placebo		(11.6, 31.7)	
p-value for difference from placebo		<0.0001	

Compiled by this reviewer from Table 18, CSR and Table 14.3.1.14 H.

As seen from the table above, patients who received vedolizumab treatment were significantly more likely to achieve a clinical response at Week 6 compared to patients who received placebo.

Patients who received vedolizumab treatment in Cohort 2 (open-label) had similar clinical response at Week 6 in Cohort 1.

Results of the primary endpoint for the per-protocol population are given in Appendix Table 7, and those for the completers (observed case) population are given in Appendix Table 8.

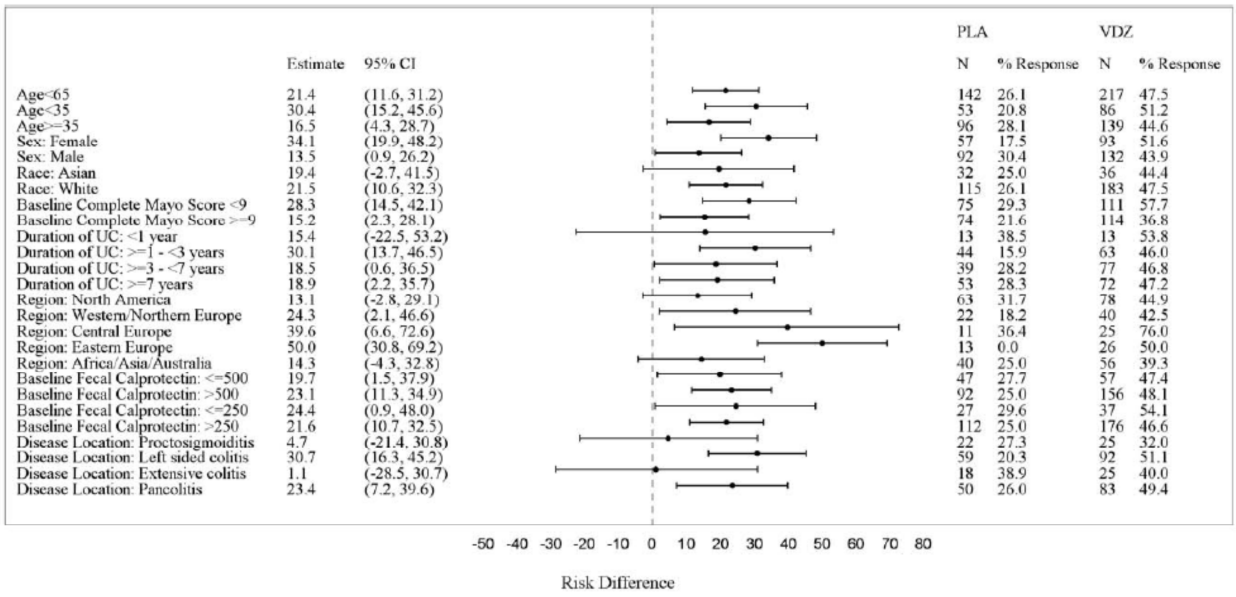
As seen from Appendix Tables 7 and 8, results for both populations were similar to those for the Induction Study ITT population.

3.1.1.2.1.4.1 Subgroup Analyses

Subgroup analyses for clinical response at Week 6 in the Induction Study ITT population are conducted for: age (< 35 vs. ≥ 35 years;.< 65 vs. ≤ 65 years), gender, race, duration from UC diagnosis to first dose, geographic region, baseline (Week 0) disease activity, baseline (Week 0) fecal calprotectin (≤ 250 µg/g vs. > 250 µg/g; ≤ 500 µg/g vs. > 500 µg/g, and disease localization.

The risk differences compared with placebo for the primary endpoint, clinical response, in subgroups according to demographic characteristics and measures of disease severity are summarized in the Figure below.

Figure 1 Risk Difference and 95% Confidence Interval for Subgroup Analyses of Clinical Response at Week 6 – Induction Study ITT Population



Copied from Figure 6, page 124 CSR.

As seen from the Figure above, the treatment benefit of vedolizumab as measured by the primary endpoint was statistically significant across age categories (< 35 and ≥ 35 years) and gender. For each of the other demographic subgroups (according to race and region), the risk differences consistently favored vedolizumab over placebo, although due to small sample sizes in some subgroups, not all of the treatment differences were significant. There were no apparent differences in the magnitude of treatment benefit in these subgroups.

Similar results were observed for subgroups according to assessments of disease severity, including categories of baseline fecal calprotectin, Mayo score category, and disease localization.

3.1.1.2.1.5 Applicant's Analyses of the Secondary Efficacy Endpoints

Secondary efficacy endpoints are:

- Clinical remission at Week 6
- Mucosal healing at Week 6

The CMH chi-square test was performed with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to TNFα antagonists and/or concomitant immunomodulator use (yes/no).

3.1.1.2.1.5.1 Clinical Remission at Week 6

Clinical remission is defined as complete Mayo score of ≤ 2 points and no individual subscore > 1 point.

The results from analysis of the proportion of patients with clinical remission at Week 6 are given below.

Table 6 Clinical Remission at Week 6 by Cohort

Clinical response	Placebo (n=149)	VDZ (n=225)	VDZ open-label (n=521)
N (%) achieve clinical remission at week 6	8 (5.4)	38 (16.9)	100 (19.2)
95% CI	(1.7, 9.0)	(12.0, 21.8)	(15.8,22.6)
Difference from placebo		11.5	
95% CI for difference from placebo		(4.7, 18.3)	
p-value for difference from placebo		0.0009	

Compiled by this reviewer from Table 19, CSR and Table 14.3.1.14 H.

As seen from the table above, significantly more patients in the vedolizumab achieved clinical remission at Week 6, as compared to the placebo group.

3.1.1.2.1.5.2 Mucosal Healing at Week 6

Mucosal healing is defined as Mayo endoscopic subscore of ≤ 1.

The results from analysis of the proportion of patients with mucosal healing at Week 6 are given below.

Table 7 Mucosal Healing at Week 6 by Cohort

Clinical response	Placebo (n=149)	VDZ (n=225)	VDZ open-label (n=521)
N (%) achieve mucosal healing at week 6	37 (24.8)	92 (40.9)	191 (36.7)
95% CI	(17.9, 31.8)	(34.5, 47.3)	(32.5, 40.8)
Difference from placebo		16.1	
95% CI for difference from placebo		(6.4, 25.9)	
p-value for difference from placebo		0.0012	

Compiled by this reviewer from Table 20, CSR and Table 14.3.1.14 H.

As seen from the table above, the percentage of patients who achieved mucosal healing at Week 6 was significantly greater for patients who received vedolizumab compared with those who received placebo.

3.1.1.2.2 Maintenance Phase

The Maintenance Phase began at Week 6, included study drug dosing at Week 6. The Maintenance Phase included three groups of patients who were assigned to treatment groups based on their induction treatment assignment and response to the study therapy.

Vedolizumab-treated patients from both Cohort 1 and Cohort 2 who demonstrated a clinical response according to protocol-specified criteria, as assessed by the investigator, were randomized in a 1:1:1 ratio to double-blind treatment with vedolizumab administered every 4 weeks (Q4W), vedolizumab administered every 8 weeks (Q8W), or placebo.

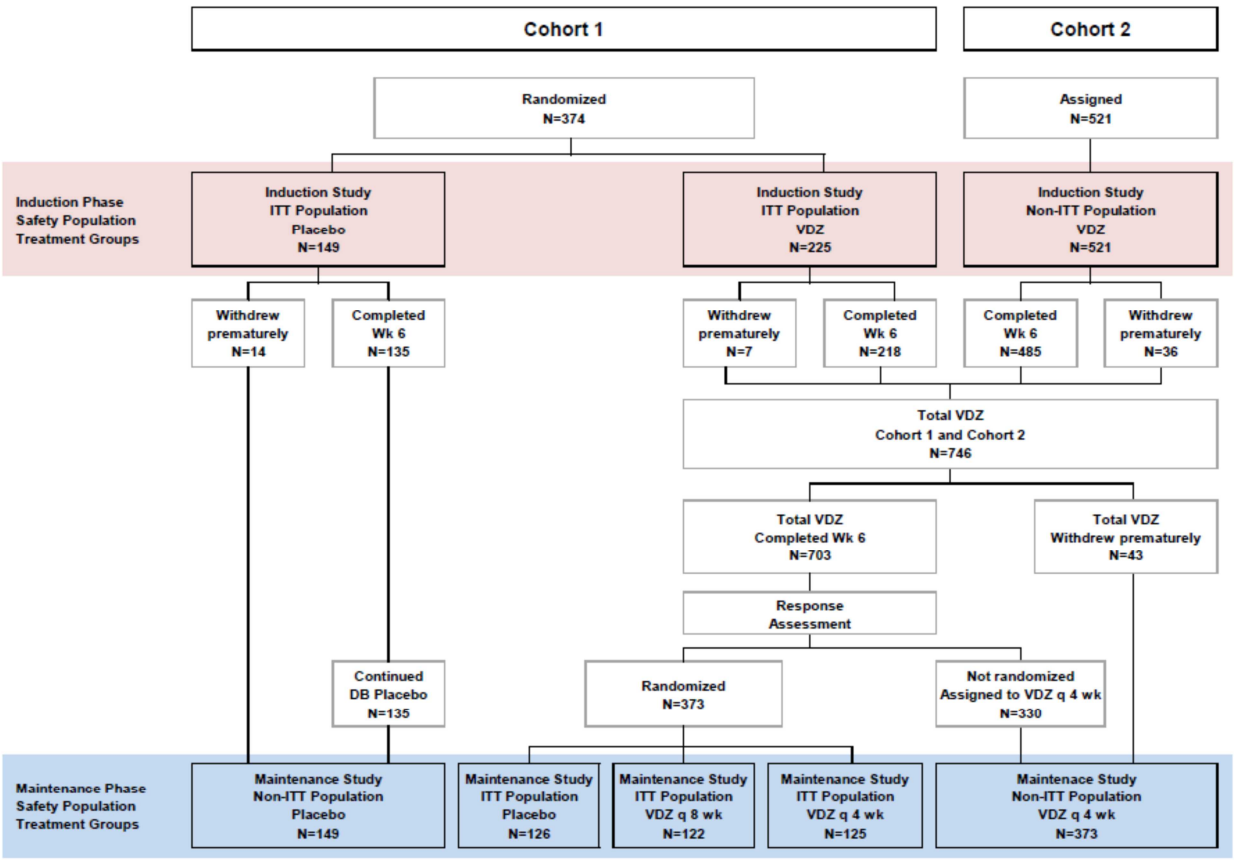
- Enrollment in Cohort 1 or Cohort 2 in the Induction Phase
- Concomitant use of oral corticosteroids
- Previous exposure to TNFα antagonists or concomitant immunomodulator use

As in the Induction Phase, the unblinded study pharmacist obtained the Maintenance Phase treatment assignment based on information provided by the Investigator; the Investigator remained blinded to the Induction Phase treatment (for those patients in Cohort 1 who had received the treatment in a double-blind manner) and there was no interruption of treatment between the two phases. These patients comprised the Maintenance Study ITT population, the primary efficacy population.

All patients who completed the Induction Phase entered the Maintenance Phase. Treatment assignments were based on the Induction Phase treatment and the Investigator-assessed treatment

response. The flow of patients from the Induction Phase into the Maintenance Phase treatment groups and the composition of the Maintenance Phase Safety Population treatment groups are summarized in the figure below.

Figure 1 Overview of Treatment Groups in Induction Phase and Maintenance Phase Safety Populations
Study C13006



Copied from Figure 11, page 174 CSR.

The Maintenance Study ITT Population included vedolizumab-treated patients who achieved clinical response at Week 6; at the start of the Maintenance Phase, these patients were randomized to one of two vedolizumab intravenous (IV) dosing regimens (300 mg every 4 weeks or every 8 weeks) or placebo. The Maintenance Study referred to the statistical analyses performed on efficacy variables in this population, the Maintenance Study ITT Population.

The Maintenance Non-ITT Population included two additional treatment groups: placebo and vedolizumab administered every 4 weeks. The non-ITT placebo group was comprised of those patients who were randomized to placebo in the Induction Phase; these patients remained on placebo in the Maintenance Phase, per the study design. The non-ITT vedolizumab group was

comprised of those patients who received vedolizumab in the Induction Phase and were assessed by the Investigator as not having achieved clinical response at Week 6; these patients received vedolizumab 300 mg every 4 weeks for the duration of the Maintenance Phase. These patients contributed to the safety analyses in the Maintenance Phase, and exploratory efficacy analyses were done for this population.

It should be noted that the safety analyses of patients in the Maintenance Phase included assessments from their participation during the entire study, starting at Week 0 of the Induction Phase. As such, information presented for the Non-ITT Population treatment groups included safety assessments from patients who withdrew from the treatment during the Induction Phase. Also of note, all patients in the Maintenance ITT Population randomized to the placebo treatment group in Maintenance Phase were treated with vedolizumab during the Induction Phase. Maintenance safety data for this group included safety assessments made while on vedolizumab treatment during the Induction Phase and on placebo during the Maintenance Phase.

3.1.1.2.2.1 Pre-specified Analyses

Primary and secondary efficacy assessments for the Maintenance Phases were based on the Mayo scores. Sigmoidoscopy was done 52 (or ET visit), and a complete Mayo score was calculated for these visits.

A baseline (Week 0) complete Mayo score, calculated by adding the screening endoscopy subscore to the partial Mayo score obtained on Day 1 (Week 0), was used for the comparison with the Week 6 complete Mayo score to determine the response and remission status at Week 6 and with the Week 52 complete Mayo score to determine the response and remission status at Week 52.

The Week 6 complete Mayo score was calculated by the investigator or designee and was recorded in the patient's source documents; this assessment was used to determine whether the patient had achieved clinical response at Week 6, and, therefore, was used to determine treatment assignment in the Maintenance Phase.

A partial Mayo score was to be derived at the visits where sigmoidoscopy was not performed (i.e., Weeks 0, 2, 4, 10, 14, 18, 22, 26, 30, 34, 38, 42, 46, and 50), and at any unscheduled visit(s) due to disease exacerbation. These scores were used to determine clinical response or disease worsening during the study.

Beginning at Week 6, patients receiving oral corticosteroids who had achieved a clinical response were to begin a corticosteroid tapering regimen. In addition, at Week 6, patients in Cohort 1 participating in the US sites who were taking concomitant azathioprine or

6-mercaptopurine during the Induction Phase were required to discontinue these medications.

After the Week 52 assessments, patients meeting protocol-defined criteria were eligible to enroll in the Long-term Safety Study C13008 to receive open-label vedolizumab treatment. Patients who withdraw early (prior to Week 52) due to sustained nonresponse, disease worsening, or the need for rescue medications might also have been eligible for Study C13008. Patients who did not enroll into Study C13008 completed a final on-study safety assessment at Week 66 (or Final Safety visit 16 weeks after the last dose) in the Maintenance Phase of Study C13006. In addition, after the end of the study, all patients who did not enroll in Study C13008 participated in a follow-up period in which they were contacted by telephone every 6 months for 2 years. The follow-up questionnaire administered at each time point collected information on events such as infections resulting in hospitalization (at 6 months only), pregnancy, colorectal dysplasia, cancer, IBD-related surgeries, and the development of Progressive Multifocal Leukoencephalopathy (PML).

The primary efficacy assessments were the differences in the proportions of patients who were in clinical remission at Week 52 between the vedolizumab every 4 weeks and placebo and between the vedolizumab every 8 weeks and placebo groups. For the two comparisons of the primary endpoint of clinical remission at 52 weeks, the Hochberg method was applied to control the overall Type I error rate at a 5% significance level. If both p values were ≤ 0.05, both dose regimens were to be declared significant. If 1 of the p values for the two dose comparisons was > 0.05, the other p value was to be tested at the 0.025 level and declared significant only if the p value was ≤ 0.025. If neither dose was declared significant for the primary endpoint, no further testing was to be conducted. If at least one of the dose regimens was significant, the sequential procedure was to be used to test the secondary endpoints for statistical significance.

For both assessments of the primary endpoint, the CMH chi-square test was used to compare the two treatment groups at the 5% level of significance with stratification according to the stratification factors (enrollment in Cohort 1 or 2 in the Induction Phase, concomitant use of oral corticosteroids, and previous exposure to TNFα antagonists or concomitant immunomodulator [6-mercaptopurine or azathioprine] use). The CMH chi-square p-values and the absolute treatment differences along with the 95% two-sided confidence intervals are provided. In addition, the relative risks are provided along with the 95% two-sided confidence intervals.

In addition to the primary comparisons, there were four key secondary assessments of clinical efficacy (durability of clinical response, mucosal healing, durability of clinical remission, and corticosteroid-free remission,), which compared treatment differences through closed testing procedures. To maintain the overall Type I error rate at 5% for the two dose regimen comparisons for each key secondary endpoint, the Hochberg method was used as described above for the primary comparisons. To further maintain the overall Type I error rate at 5%, the

key secondary assessments were also performed sequentially. The first key secondary endpoint was to be tested only if one or both of the primary comparisons were significant and the next key secondary endpoint was to be tested only if the previous key secondary endpoint was significant for at least one dose. The order of the key secondary objectives was specified in the statistical analysis plan before clinical database lock.

The differences in the proportions of patients who were in durable response, the differences in the proportions of patients with mucosal healing, the differences in the proportions of patients who were in durable clinical remission and the differences in the proportions of patients who discontinued corticosteroids by Week 52 and who were in clinical remission at Week 52 were analyzed in the same manner as the primary endpoint.

A sample size of 372 was planned to power the Maintenance Study primary and secondary efficacy endpoints. Assuming an induction response rate of 55% among patients receiving vedolizumab (either in Cohort 1 or 2), there would be approximately 372 patients on vedolizumab in the Induction Phase who achieved clinical response at Week 6. During Induction, the overall response rate and attrition rate for patients who were not willing to participate in the Maintenance Phase were evaluated periodically by the DSMB to assess study assumptions. This monitoring allowed for necessary adjustments to the number of patients enrolled into the second cohort to achieve the target Maintenance Study sample size of approximately 372 patients. A maximum of 100 additional patients could have been enrolled into the second cohort of the Induction Phase to achieve the target Maintenance Study sample size.

The sample size calculation for the Maintenance Study was based on the number of patients who had received vedolizumab (in either Cohort 1 or Cohort 2) in the Induction Phase and achieved clinical response at Week 6. Power estimates based on a total sample size of 372 patients (124 per arm) in the table below.

Table 8 Power Estimates for the Primary and Key Secondary Efficacy Analyses In Maintenance Study -Study C13006

Objective	Maintenance Period Endpoint at Week 52	Assumed Response Rate	Sample Size per Group[a]	Power
Primary	Remission for vedolizumab vs placebo	Placebo = 30% Vedolizumab = 50%	124	90%
Key secondary	Durable response for vedolizumab vs placebo	Placebo = 14% Vedolizumab = 37%	124	98%
	Mucosal healing for vedolizumab vs placebo	Placebo = 25% Vedolizumab = 50%	124	98%
	Durable remission for vedolizumab vs placebo	Placebo = 7% Vedolizumab=20%	124	85%
	Corticosteroid-free remission for vedolizumab vs placebo	Placebo = 9% Vedolizumab = 25%	68[b]	70%

Copied from Table 5, page 77 CSR.

3.1.1.2.2.2 Patient Disposition

The detailed patient disposition is given below.

Table 9 Patient Disposition– Maintenance Phase Study C13006

	Maintenance Study ITT [a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Non-ITT		Combined	
	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125	PLA [b] (from Week 0) N = 149	VDZ Q4W [c] (Week 6 non-responders) N = 373	PLA N = 275	VDZ N = 620
Completed induction treatment	126 (100)	122 (100)	125 (100)	135 (91)	330 (88)	261 (95)	577 (93)
Randomized	126 (100)	122 (100)	125 (100)	149 (100)	373 (100)	275 (100)	620 (100)
Randomized but not dosed	0	0	0	0	0	0	0
Safety population [d]	126 (100)	122 (100)	125 (100)	149 (100)	373 (100)	275 (100)	620 (100)
Maintenance Study Intent-to-treat population [a]	126 (100)	122 (100)	125 (100)	—	—	126 (46)	247 (40)
Maintenance Study ITT Per-Protocol population [e]	121 (96)	117 (96)	121 (97)	—	—	121 (44)	238 (38)
Completed Maintenance Phase [f]	48 (38)	77 (63)	84 (67)	30 (20)	135 (36)	78 (28)	296 (48)

Discontinued (reason) [g]	78 (62)	45 (37)	41 (33)	119 (80)	238 (64)	197 (72)	324 (52)
Adverse event	15 (12)	7 (6)	6 (5)	16 (11)	23 (6)	31 (11)	36 (6)
Protocol violation(s)	0	0	0	2 (1)	9 (2)	2 (< 1)	9 (1)
Lack of efficacy	61 (48)	31 (25)	33 (26)	88 (59)	171 (46)	149 (54)	235 (38)
Study terminated by sponsor	0	0	0	0	0	0	0
Withdrawal of consent	2 (2)	5 (4)	2 (2)	9 (6)	32 (9)	11 (4)	39 (6)
Lost to follow-up	0	2 (2)	0	4 (3)	3 (< 1)	4 (1)	5 (< 1)
Other	0	0	0	0	0	0	0
Enrolled into C13008	113 (90)	108 (89)	112 (90)	112 (75)	230 (62)	225 (82)	450 (73)

Copied from Table 41, page 175-176 CSR.

As seen from the table above, with regard to the ITT Population, a greater proportion of placebo-treated patients discontinued treatment than did vedolizumab-treated patients (62% placebo vs. 37% and 33% in the vedolizumab Q8W and Q4W dosing regimens, respectively). The most reported reason for discontinuation across all of the ITT Population treatment groups was lack of efficacy, which occurred in 48% of the placebo group and less frequently in the vedolizumab groups (25% and 26%, respectively). Discontinuations due to adverse events (AEs) were twice as in the placebo group (12% vs. 6% and 5% in the two vedolizumab groups, respectively). Most of the patients in the ITT population continued into the long term Study C13008.

In the non-ITT population, only 20% of the placebo patients completed 52 weeks, compared with 36% of patients in the vedolizumab population. The most reported reasons for treatment discontinuation were lack of efficacy, occurring in 59% of the placebo group and 46% of the vedolizumab group. Adverse events leading to discontinuation occurred in 11% of the placebo group and 6% of the vedolizumab group.

3.1.1.2.2.3 Treatment Group Comparability

Baseline demographic characteristics of the Maintenance Phase Safety population are summarized by treatment group in Appendix Tables 11 and 12.

As seen from Appendix Table 11 and 12, baseline demographic characteristics were similar in the ITT Population treatment groups, with the exception of geographic region. More patients in the Q8W vedolizumab dosing group were enrolled at sites in North America (40%, compared to 29% and 30% of patients in the placebo and vedolizumab Q4W groups, respectively) and fewer were enrolled at Asian, Australian, and African sites (15% vs. 27% and 22%, respectively).

Baseline demographics for patients in the non-ITT placebo and vedolizumab Q4W groups were similar to those of the ITT treatment groups, although both non-ITT treatment groups had a

slightly higher percentage of males (62% and 61%, respectively) compared with the maintenance ITT placebo, Q8W, and Q4W groups (55%, 57%, and 54% male, respectively).

Baseline UC disease characteristics of the Maintenance Phase Safety population are summarized in Appendix Table 13.

As seen from Appendix Table 13, in the ITT treatment groups, durations of UC were similar in the placebo and Q8W treatment groups (median 5.4 years in both) but slightly less in the Q4W group (median 5.0 years). Baseline disease activity, as assessed by mean Mayo score and category of Mayo score, was similar in the three groups, as was the category of baseline fecal calprotectin. There were treatment differences in mean baseline fecal calprotectin. However, as stated by the applicant, due to large variability in the values, it is unlikely that these represent actual differences. In general, baseline disease characteristics in the non-ITT treatment groups were similar to those in the ITT treatment groups.

Patients in the Maintenance Phase non-ITT vedolizumab group (non-responders at Week 6) had slightly higher disease activity at baseline (as assessed by the complete Mayo score) compared with patients in the other populations. This included baseline disease activity (median of 9.0 for non-ITT Q4W patients compared with 8.0 for each ITT group), and the percentage of patients with a complete Mayo score of 9 to 12 at baseline (56% of non-ITT Q4W patients compared with 42% to 45% of ITT patients).

The prior use of TNFα antagonists and treatment failure to UC therapies are summarized for the Maintenance Phase Safety population in Appendix Table 14.

As seen from Appendix Table 14, prior TNFα antagonist use was similar in the ITT population treatment groups, as was prior use of other UC treatments. The majority of patients had exposure to systemic corticosteroids and / or immunomodulators: 97% of combined placebo patients and 98% of combined vedolizumab patients had exposure to corticosteroids; 75% of combined placebo patients and 76% of combined vedolizumab patients had exposure to immunomodulators.

A hierarchical approach was used to categorize the nature of the prior treatment failure to TNFα antagonists, immunomodulators, and corticosteroids (i.e., inadequate response, loss of response, or intolerance). There were no notable differences among the treatment groups in these categories of failure.

The non-ITT Q4W treatment group (Week 6 non-responders) had a higher proportion of patients who had prior TNFα antagonist failure than patients in the ITT population (51%, compared to

30%, 36%, and 32% in the ITT placebo, vedolizumab Q8W, and vedolizumab Q4W groups, respectively).

Baseline UC therapy is summarized for the Maintenance Phase Safety population in Appendix Table 15.

As seen from Appendix Table 15, UC therapy use at baseline (Week 0) was similar between treatment groups in the Maintenance ITT Population.

Appendix Table 16 shows comparisons of selected baseline UC disease characteristics in the Maintenance Study ITT population treatment groups.

As seen from Appendix Table 16, there were no significant differences in baseline UC disease characteristics between treatment groups for duration of UC, baseline disease activity, use of concomitant therapies at baseline (corticosteroids or immunomodulators), or prior TNFα antagonist use.
Note that all patients randomized into the Maintenance Study ITT population were treated with vedolizumab during the Induction Phase and achieved clinical response at Week 6, as assessed by the investigator. Patients in the Maintenance Study ITT placebo treatment group received their first dose of placebo at Week 6.

3.1.1.2.2.4 Analysis Populations

The table below summarizes the analysis populations within the Maintenance Study ITT population (i.e., only patients who received vedolizumab during the Induction Phase, met the protocol definition of clinical response at Week 6, and then received any study drug in the Maintenance Phase.

Table 10 Summary of Maintenance Analysis Populations
Maintenance Study ITT Population
Study C13006

Data Set	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125
Randomized patients	126	122	125
Safety population [a]	126 (100)	122 (100)	125 (100)
Intent-to-treat population [b]	126 (100)	122 (100)	125 (100)
Modified ITT population [c]	112 (89)	111 (91)	116 (93)
Per-Protocol population [d]	121 (96)	117 (96)	121 (97)
Completers (Observed Case) population [e]	48 (38)	77 (63)	83 (66)

Copied from Table 52, page 196 CSR.

3.1.1.2.2.5 Applicant's Analyses of the Primary Efficacy Endpoint

The CMH chi-square test was performed with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to TNFα antagonists and/or concomitant immunomodulator use (yes/no); 3) enrollment in Cohort 1 or Cohort 2 in the Induction Phase.

The primary endpoint for the Maintenance Phase was the proportion of patients with clinical remission at Week 52. Summary of clinical remission at 52 is given in Table below.

Table 11 Clinical Remission at Week 52
Maintenance Study ITT Population
Study C13006

Clinical Remission [a]	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125
Number (%) achieving clinical remission	20 (15.9)	51 (41.8)	56 (44.8)
95% CI	(9.5, 22.3)	(33.1, 50.6)	(36.1, 53.5)
Difference from placebo [b]		26.1	29.1
95% CI for difference from placebo		(14.9, 37.2)	(17.9, 40.4)
P value for difference from placebo [c]		< 0.0001	< 0.0001
Relative risk [d]		2.7	2.8
95% CI for relative risk		(1.7, 4.2)	(1.8, 4.4)

Copied from Table 53, page 197 CSR.

As seen from the table above, both vedolizumab dosing treatment groups had significantly more patients achieving clinical remission at Week 52 as compared to the placebo treatment group.

Results of the primary endpoint for the per-protocol population are given in Appendix Table 17, and that of the completers (observed case) population in Appendix Table 18.

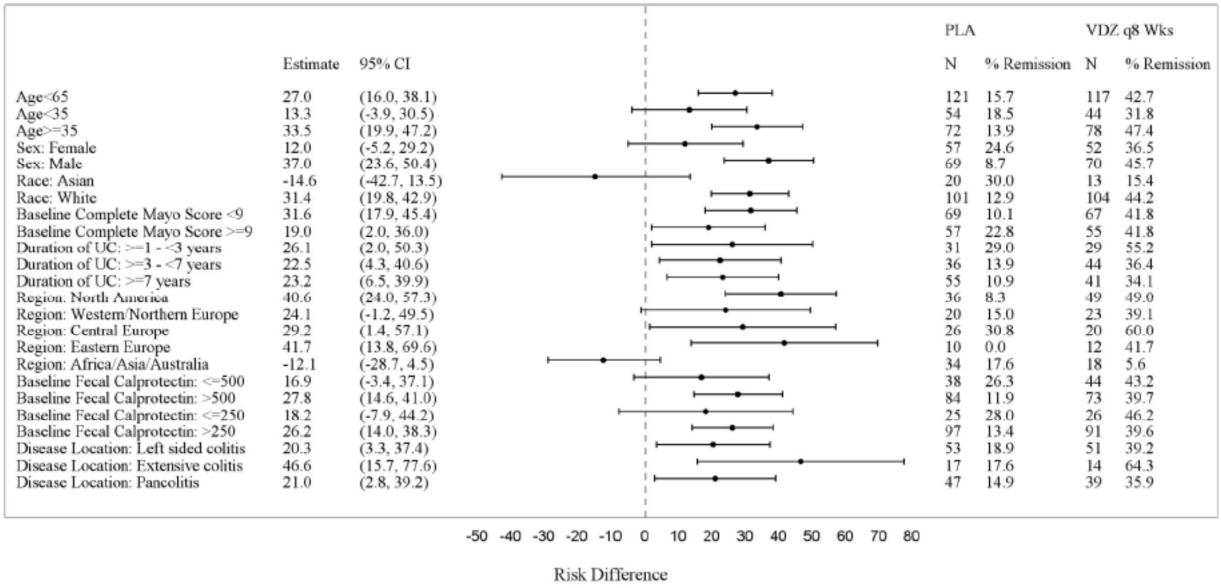
As seen from Appendix Tables 17 and 18, the results of these analyses were similar to those of the primary efficacy analyses; statistically significant treatment differences were observed in each population.

3.1.1.2.2.5.1 Subgroup Analysis

Subgroup analyses for clinical remission at Week 52 in the Maintenance Study ITT population were provided based on: age (< 35 vs. ≥ 35 years; < 65 vs. ≤ 65 years), gender, race, duration from UC diagnosis to first dose, geographic region, baseline (Week 0) disease activity, baseline (Week 0) fecal calprotectin (≤ 250 µg/g vs. > 250 µg/g; ≤ 500 µg/g vs. > 500 µg/g), and disease localization.

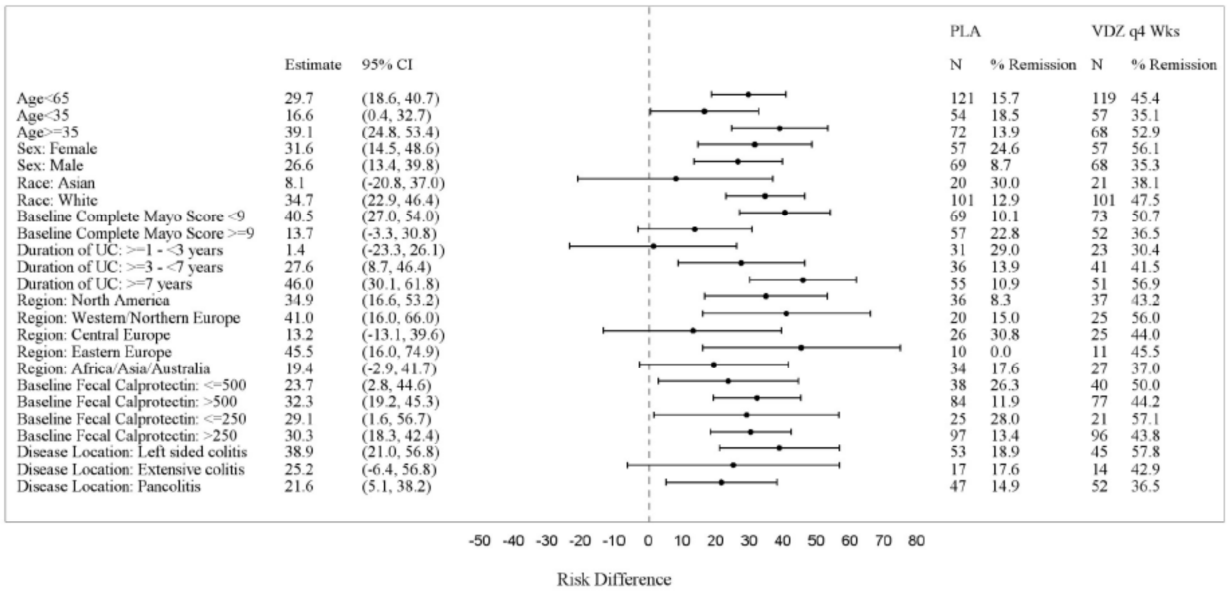
The risk differences compared with placebo for the primary endpoint, clinical response, in subgroups according to demographic characteristics and measures of disease severity are summarized in the figures below.

Figure 2 Risk Difference and 95% Confidence Interval for Subgroup Analyses of Clinical Remission at Week 52 for VDZ Q8W vs. Placebo
Maintenance Study ITT Population
Study C13006



Copied from Figure 12, page 199 CSR.

Figure 3 Risk Difference and 95% Confidence Interval for Subgroup Analyses of Clinical Remission at Week 52 for VDZ Q4W vs. Placebo
Maintenance Study ITT Population
Study C13006



Copied from Figure 13, page 200 CSR.

As seen from the figures above, for either dose (Q8W and Q4W), the treatment benefit of vedolizumab as measured by the primary endpoint was statistically significant for age ≥ 35 years and for males. For each of the other demographic subgroups (according to race and region), the risk differences consistently favored vedolizumab over placebo, although due to small sample sizes in some subgroups, not all of the treatment differences were significant. There were no apparent differences in the magnitude of treatment benefit in these subgroups.

Similar results were observed for subgroups according to assessments of disease severity, including categories of baseline fecal calprotectin, Mayo score category, and disease localization.

3.1.1.2.2.6 Applicant's Analyses of the Secondary Efficacy Endpoints

Secondary efficacy endpoints are:

- Durability of clinical response
- Mucosal healing at Week 52
- Durability of clinical remission
- Corticosteroid-free clinical remission at Week 52

The CMH chi-square test was performed with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to TNFα antagonists and/or concomitant immunomodulator use (yes/no); 3) enrollment in Cohort 1 or Cohort 2 in the Induction Phase.

3.1.1.2.2.6.1 Durable Clinical Response

Durable clinical response, defined as a clinical response at Weeks 6 and 52, was a key secondary endpoint of the Maintenance Study.

Table 13 Durable Clinical Response
Maintenance Study ITT Population
Study C13006

Durable Clinical Response [a]	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125
Number (%) achieving durable clinical response	30 (23.8)	69 (56.6)	65 (52.0)
95% CI	(16.4, 31.2)	(47.8, 65.4)	(43.2, 60.8)
Difference from placebo [b]		32.8	28.5
95% CI for difference from placebo		(20.8, 44.7)	(16.7, 40.3)
P value for difference from placebo [c]		< 0.0001	< 0.0001
Relative risk [d]		2.4	2.2
95% CI for relative risk		(1.7, 3.4)	(1.5, 3.1)

Copied from Table 54, page 203 CSR.

As seen from the table above, a significantly higher percentage of vedolizumab patients in both dosing regimen groups met this endpoint compared with patients who received placebo.

Per this reviewer's request, the applicant performed analysis of durable clinical response at Week 52 without imputation. Results of this analysis are presented in Appendix Table 19.

As seen from Appendix Table 19, the results of this analysis were similar to those of applicant's analysis for durable clinical response.

3.1.1.2.2.6.2 Mucosal Healing at Week 52

The number and proportion of patients with applicant's definition of mucosal healing at Week 52 in the Maintenance Study ITT population are summarized by treatment group in the table below.

Table 14 Mucosal Healing at Week 52
Maintenance Study ITT Population
Study C13006

Mucosal Healing [a]	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125
Number (%) achieving mucosal healing	25 (19.8)	63 (51.6)	70 (56.0)
95% CI	(12.9, 26.8)	(42.8, 60.5)	(47.3, 64.7)
Difference from placebo [b]		32.0	36.3
95% CI for difference from placebo		(20.3, 43.8)	(24.4, 48.3)
P value for difference from placebo [c]		< 0.0001	< 0.0001
Relative risk [d]		2.6	2.8
95% CI for relative risk		(1.8, 3.9)	(1.9, 4.2)

Copied from Table 55, page 204 CSR.

As seen from the table above, for mucosal healing at Week 52, both vedolizumab dosing regimens provided a significant benefit over placebo.

Per this reviewer's request, the applicant performed the "observed case" analysis of mucosal healing at Week 52. Results of this analysis are given Appendix Table 20.

As seen from Appendix Table 20, the results of these analyses were similar to those of applicant's analysis for mucosal healing at Week 52.

3.1.1.2.2.6.3 Durable Clinical Remission

The number and proportion of patients who experienced a durable clinical remission, defined as clinical remission at Week 6 and Week 52 in the Maintenance Study ITT population are summarized by treatment group in the table below

Table 15 Durable Clinical Remission
Maintenance Study ITT Population
Study C13006

Durable Clinical Remission [a]	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125
Number (%) achieving durable clinical remission	11 (8.7)	25 (20.5)	30 (24.0)
95% CI	(3.8, 13.7)	(13.3, 27.7)	(16.5, 31.5)
Difference from placebo [b]		11.8	15.3
95% CI for difference from placebo		(3.1, 20.5)	(6.2, 24.4)
P value for difference from placebo [c]		0.0079	0.0009
Relative risk [d]		2.4	2.8
95% CI for relative risk		(1.2, 4.6)	(1.4, 5.3)

Copied from Table 57, page 205 CSR.

As seen from the table above, the percentage of patients who experienced a durable clinical remission was significantly higher for patients in both the vedolizumab Q8W and Q4W dosing regimen groups compared with patients who received placebo.

Per this reviewer's request, the applicant performed an analysis of durable clinical remission at Week 52 without imputation. Results of this analysis are provided in Appendix Table 21.

As seen from Appendix Table 21, the results of this analysis were similar to those of applicant's analysis for durable clinical remission.

3.1.1.2.2.6.4 Corticosteroid-Free Clinical Remission at Week 52

All Maintenance Study ITT patients who were on corticosteroids at Week 6 were to begin a corticosteroid tapering regimen; approximately 58% of the ITT population was on corticosteroids at Week 6.

The number and proportion of patients in the Maintenance Study ITT population with corticosteroid-free remission at Week 52 are summarized by treatment group in the table below.

Table 16 Corticosteroid-free Remission at Week 52
Maintenance Study ITT Population, Patients on Corticosteroids at Baseline
Study C13006

Corticosteroid-free Clinical Remission [a]	PLA n = 72	VDZ Q8W n = 70	VDZ Q4W n = 73
Number (%) achieving corticosteroid-free clinical remission	10 (13.9)	22 (31.4)	33 (45.2)
95% CI	(5.9, 21.9)	(20.6, 42.3)	(33.8, 56.6)
Difference from placebo [b]		17.6	31.4
95% CI for difference from placebo		(3.9, 31.3)	(16.6, 46.2)
P value for difference from placebo [c]		0.0120	< 0.0001
Relative risk [d]		2.3	3.3
95% CI for relative risk		(1.2, 4.4)	(1.7, 6.1)

Copied from Table 58, page 206 CSR.

As seen from the table above, vedolizumab treatment was associated with significantly higher rates of corticosteroid-free remission at Week 52

Per this reviewer's request, the applicant performed the "observed case" analysis of corticosteroid-fee remission at Week 52. Results of this analysis are given in Appendix Table 22.

As seen from Appendix Table 22, the results of this analysis were similar to those of the applicant's analysis for corticosteroid-free remission at Week 52.

3.1.1.3 Reviewer's Comments and Evaluation

3.1.1.3.1 Induction Phase

3.1.1.3.1.1 Sensitivity Analyses for the Primary and Secondary Endpoint

Per the FDA's request to address the issue of missing data, the applicant performed the following sensitivity analyses for the primary endpoint and secondary endpoint for both Induction and Maintenance Phase of this Study:

- Observed case: exclude subjects from the analysis at a specific time point if the patients have insufficient data at that time point.
- Complete case: exclude subjects from the analysis at all-time points if they have insufficient data at any of the time points of the analysis.
- Worst case: (1) subjects with missing observations at any of the time points of analysis are assumed to be non-responders; (2) subjects receiving placebo with

missing observations at any of the time points of the analysis are assumed to be responders, and subjects receiving treatment with missing observations at any of the time points of analysis are assumed to be non-responders.

- LOCF (Last-Observation Carried Forward) analysis
- Multiple imputation

The applicant stated clarifying information relating to the five requests above as follows:

- Observed Case and Complete Case: The observed and complete case analyses are identical to analyses done without imputation. Only one set of analyses is provided with this response because observed case and complete case are identical analyses. In the observed case, insufficient data at a specific time point implies that there are no data at Week 6 or Week 52. In such cases, the observed case is equivalent to analyses without imputation. In the complete case, insufficient data at all analyses time points indicate that there are no data at Week 6 or Week 52.
- Worst Case 2: The requested analyses are provided in this response. Patients receiving placebo who had missing data were assumed to be responders and patients receiving vedolizumab who had missing data were assumed to be non-responders. There are limitations to this analysis, due to an imbalance in missing data between placebo and vedolizumab groups. This is because a larger numbers of the placebo patients reportedly failed treatment earlier and they were allowed to enroll in Study C13008. Thus, considering failure as a success for the placebo group may be biased against the vedolizumab group.
- LOCF analyses: The requested analyses are provided in this response. If a subject had missing data at a particular time point, then data from the prior time point was imputed. The applicant claimed unable to provide the requested analyses for the secondary endpoint of corticosteroid -free remission at Week 52 (based on the complete Mayo score and corticosteroid free status at Week 52), because the prior assessments of the corticosteroid-free remission are equivalent to Week 6 assessments, which is the baseline data (without tapering).
- Multiple imputations: The requested analyses are provided in this response. Multiple imputations were performed using SAS PROC MI. The number of iterations was set to 10. For the Induction Phase of Study C13006, stratification factors of concomitant use of oral corticosteroids, previous exposure to TNFα antagonist and/or concomitant immunomodulator use were used as adjusting factors. For the Maintenance Phase of Study C13006, stratification factors of concomitant use of oral corticosteroids, previous exposure to TNFα antagonist and/or concomitant immunomodulator use, and participation in Cohort 1 or Cohort 2 were used as adjusting factors.

Summary of the sensitivity analyses results for clinical response at Week 6 and clinical remission at Week 6 are given below.

Table 17 Sensitivity Analyses – Clinical Response at Week 6

Analysis	Placebo	VDZ	Difference	P-value
Primary	38/149 (25.5%)	106/225 (47.1%)	21.7%	<0.0001
Observed Case	38/137 (27.7%)	106/216 (49.1%)	21.5%	<0.0001
Per Protocol	38/138 (27.5%)	106/215 (49.3%)	21.8%	<0.0001
LOCF	39/149 (26.2%)	106/225 (47.1%)	21.0%	<0.0001
Worst Case 2	50/149 (33.6%)	106/225 (47.1%)	13..6%	0.0088
Multiple Imputation	43/149 (28.9%)	113/225 (50.2%)	21.4%	<0.0001

Compiled by this reviewer from Tables 18, CSR, and 14.3.1.2B, 14.3.1.2C 39.13.3.2A, 39.13.4.1A, and 39.13.5.1A.

Table 18 Sensitivity Analyses – Clinical Remission at Week 6

Analysis	Placebo	VDZ	Difference	P-value
Primary	8/149 (5.4%)	38/225 (16.9%)	11.5%	0.0009
Observed Case	8/137 (5.8%)	38/216 (17.6%)	11.7%	0.0014
ITT with Revised eDiary Requirements[a]	8/149 (5.4%)	38/225 (16.9%)	11.5%	0.0009
LOCF[b]	8/149 (5.4%)	38/225 (16.9%)	11.5%	0.0009
Worst Case 2[c]	20/149 (13.4%)	38/225 (16.9%)	3.5%	0.3631
Multiple Imputation	8/149 (5.4%)	38/225 (16.9%)	11.5%	0.0009

Compiled by this reviewer from Tables 19, CSR, 39.12.1.1B, 39.13.3.2B, 39.13.4.1B, and 39.13.5.1B.

[a] ITT population where patients with < 3 days of diary data within 7 days prior to Week 52 are classified as non-responders Table 39.12.2.1D
[b] Last Observation Carried Forward (LOCF) analysis imputed data from the prior time point, if a subject had missing data at a particular time point.
[c] Worst Case analysis assumed patients receiving placebo who had missing data to be responders and patients receiving vedolizumab who had missing data to be non-responders

As seen from the tables above, results from these sensitivity analyses favored consistently vedolizumab against the placebo for clinical response at Week six and clinical remission at Week 6.

3.1.1.3.1.2 Clinical Remission at Week 6 (Alternative Definition Proposed by the FDA

In this study, clinical remission, defined as a complete Mayo score of ≤ 2 points and no individual subscore > 1 point, was a secondary endpoint in the Induction Study (Week 6).

Additional analyses have been conducted based on an alternative definition of clinical remission proposed by the FDA. The new definition is defined as complete Mayo score of ≤ 2 points and no individual subscore > 1 point where rectal bleeding subscore = 0 and endoscopy subscore = 0.

The proportion of patients who achieved the alternative definition of clinical remission at Week 6 are summarized by treatment group for the Induction Study ITT Population in the table below.

Table 19 Clinical Remission (Alternative Definition Proposed by FDA) at Week 6 Induction Study ITT Population Study C13006

Clinical Remission,[a] n (%)	PLA N = 149	VDZ N = 225
Number (%) achieving clinical remission	4 (2.7)	10 (4.4)
95% CI	(0.1, 5.3)	(1.8, 7.1)
Difference from placebo[b]		1.8
95% CI for difference from placebo		(-2.1, 5.7)
P-value for difference from placebo[c]		0.3728
Relative risk[d]		1.7
95% CI for relative risk		(0.5, 5.2)

Copied from Table 3-6, page 21 FESA.

A seen from the table above, few patients in either treatment group achieved this alternative definition of clinical remission. Furthermore, no treatment difference was noted between the vedolizumab (4.4%) and placebo (2.7%) groups in the Induction Study ITT Population for the proportion of patients who achieved the alternative definition of clinical remission at Week 6.

Per the FDA's request to provide an explanation for why clinical remission at Week 6 results were different from those from the above analysis using the more stringent definition of clinical remission, the applicant stated that in the absence of prospective clinical studies testing these endpoints, a definitive explanation as to why results are different for the two definitions of clinical remission is not possible. However, as noted by FDA, the definition of clinical remission provided by FDA is more stringent, requiring achievement of a rectal bleeding score of 0 and endoscopy subscore of 0, which correspond clinically to no rectal bleeding at all and completely normal mucosa on endoscopy. Achieving this substantially more stringent definition of clinical remission results in total normalization of the colonic mucosa in patients with moderate to severe UC at baseline. It is possible that this more stringent endpoint requires treatment beyond six weeks for many patients, as suggested by the results of this study.

These results suggest that one plausible explanation for the decreased number of patients who met the more stringent definition of clinical remission was that patients might benefit from longer treatment durations beyond six weeks to achieve this endpoint.

3.1.1.3.1.3 Four Alternative Definitions for Clinical Remission

Per the FDA's request, the applicant provided exploratory analyses of clinical remission based on the following four alternative endpoint definitions for clinical remission:

a. Endoscopy subscore = 0, Rectal Bleeding subscore = 0, and Stool Frequency subscore decreases or no change from Baseline (all assessed at Week 6)
b. Endoscopy subscore ≤ 1, Rectal Bleeding subscore = 0, and Stool frequency subscore = 0 (all assessed at Week 6)
c. Endoscopy subscore ≤ 1, Rectal Bleeding subscore = 0, and Stool frequency subscore ≤ 1 (all assessed at Week 6)

d. Endoscopy subscore ≤ 1, Rectal Bleeding subscore = 0, Stool Frequency subscore decreases or no change from Baseline, and Total score ≤ 1 (all assessed at Week 6)

Results of the analyses based on these four alternative definitions for clinical remission at Week 6 are given in the table below.

Table 20 Clinical Remission at Week 6 – Four Alternative Definitions for Clinical Remission
Study C13006

Definition	Placebo	VDZ	Diff(VDZ-PLO)	p-value
Endo=0, RB=0,, SF decease or no change	4/149 (2.7%)	8/225 (3.6%)	0.9%	0.6329
Endo ≤1, RB=0,SF=0	6/149 (4.0%)	27/225 (12.0%)	8.0%	0.0077
Endo ≤1, RB=0, SF≤1	16/149 (10.7%)	60/225 (26.7%)	16.0	0.0002
Endo ≤1, RB=0, SF decease or no change, TS≤1	8/149 (5.4%)	28/225 (12.4%)	7.1%	0.0230

As seen from the table above, there was a trend favoring vedolizumab against placebo for these four alternative definitions of clinical remission at Week 6.

3.1.1.3.1.4 Mucosal Healing at Week 6

The applicant defined mucosal healing as Mayo endoscopic subscore of ≤ 1 point and provided no histologic data to support a mucosal healing claim.

Based on the applicant's definition, 40.9% of patients in the vedolizumab treatment group achieved mucosal healing, compared with 24.8% of the patients receiving placebo: a 16.1% treatment difference (95% CI: 1.2, 2.3; p = 0.0012) was observed.

However, when focusing on the subset of patients who had an endoscopy subscore of 0 at Week 6, indicating normal or inactive disease, there was no apparent treatment difference 0.9%; (95 CI: -3.4, 5.1; p=0.6956). However, this is a post-hoc subgroup analysis and the study was not properly powered to show treatment difference in this subpopulation.

3.1.1.3.1.5 Clinical Response at Week 6 in Subgroups based on Anti-TNF Status (Inadequate Response/Loss of Response)

Per this reviewer's request, the applicant performed a subgroup analysis of clinical response at Week 6 by Induction Phase Cohort in subgroups based on anti-TNF failure status (inadequate response/loss of response).

Results of this subgroup analysis are given in Appendix Table 23.

A summary of this subgroup analysis for combined Induction Phase Cohorts (Cohort1 and Cohort 2) is given below.

Table 21 Clinical Response at Week 6 – Evaluation in Subgroups Based on Anti-TNF Failure Status (Inadequate Response/Loss of Response) Study C13006

	Clinical response	Placebo (n=149)	VDZ (n=746)
Prior Anti-TNF Failure (Yes)	N (%) achieve clinical response at week 6	12/55 (21.8)	89/262 (34.0)
	95% CI	(10.9, 32.7)	(28.5, 40.0)
Prior Anti-TNF Failure (No)	N (%) achieve clinical response at week 6	26/94 (27.7)	248/484 (51.2)
	95% CI	(18.6, 36.7)	(48.8, 55.7)

Compiled by this reviewer from Table 39.15.1.1, page 160 Response to Agency Questions Received 8/9/2013.

As seen from the table above, the 95% confidence intervals overlapped for patients who had prior anti-TNF failures. For patients who did not have not prior anti-TNF failures, the 95% confidence intervals did not overlap.

3.1.1.3.2 Maintenance Phase

This reviewer found a discrepancy in the numbers of vedolizumab patients who were Week 6 responders in the Induction Phase (Cohort 1) given in Table 18 of Clinical CSR and in the Open Label study (Cohort 2) given in Table 39.17.1.1.

The applicant's response is provided below.

The cause of the discrepancy between the number of Week-6 responders for analyses and the number of patients randomized into the Maintenance Phase is attributable to differences in classifying patients as responders between the Applicant and the clinical sites, where randomization decisions were made. As shown in Figure 11 of the Study C13006 CSR, the clinical sites determined that a total of 373 vedolizumab patients from Cohort 1 and Cohort 2 were categorized as responders and randomized these patients into the Maintenance Phase.

Following a review by the Applicant, it was determined that 41 patients categorized as responders by the clinical sites, and therefore randomized into the Maintenance Phase, were in actuality non-responders. These 41 patients were excluded from the number of vedolizumab patients who were categorized as responders at Week 6 for the purpose of analyses. Conversely, there were five vedolizumab-treated patients who were categorized by the clinical sites as non-responders who the Applicant, upon later review, determined to be responders. These five patients were included in the number of vedolizumab patients who were categorized as responders at Week 6 for the purpose of analyses. These two discrepancies resulted in a difference of 36 vedolizumab-treated patients randomized into the Maintenance Phase (total of 373 patients) and the total number of vedolizumab patients categorized as Week-6 responders used for analyses (106 patients in Cohort 1 and 231 patients in Cohort 2 for a total of 337 patients). Please refer to the table below.

Table 1.a Treatment Group Calculations – Study C13006 CSR, Table 18 and Figure 11, and Response to Agency Table 39.17.1.1

		Incorrectly Randomized		
Cohort	Week 6 Clinical Response[a]	Randomized as Responders but Analyzed as Nonresponders[b]	Randomized as Nonresponders but Analyzed as Responders[c]	Figure 11[d]
Cohort 1	106	18	4	120
Cohort 2	231	23	1	253
Total	337	41	5	373

Source: Study C13006 CSR Table 18 and Figure 11; Response to Agency Table 39.17.1.1.
Abbreviations: CSR = clinical study report.
a Study C13006 CSR, Table 18, and Response to Agency, Table 39.17.1.1.
b Add to Study C13006 CSR, Table 18, and Response to Agency, Table 39.17.1.1.
c Subtract from Study C13006 CSR, Table 18, and Response to Agency, Table 39.17.1.1.
d Randomized to Maintenance.

A post-hoc sensitivity analysis was performed to assess the impact of inclusion of patients who were classified as responders by sites but were classified as non-responders by the Applicant; clinical remission was assessed for all patients in the ITT population who met the

protocol definition of clinical response at Week 6. The results of these analyses were similar to those of the primary efficacy analyses; statistically significant treatment differences were observed in each population (Study C13006 CSR Section 11.2.1-M, Table 14.3.1.27BM below).

Table 14.3.1.27BM
Clinical Remission at Week 52 For Patients Who Achieved Clinical Response at Week 6
Intent-to-Treat Population

Clinical Remission [a]	PLA N=110	VDZ q8 wks N=111	VDZ q4 wks N=111
Number (%) Achieving Clinical Remission	20 (18.2)	50 (45.0)	52 (46.8)
95% CI	(11.0, 25.4)	(35.8, 54.3)	(37.6, 56.1)
Difference from Placebo [b]		27.6	28.2
95% CI for Difference from Placebo		(15.5, 39.8)	(16.0, 40.4)
P-value for Difference from Placebo [c]		<0.0001	<0.0001
Relative Risk [d]		2.5	2.5
95% CI for Relative Risk		(1.6, 3.9)	(1.6, 4.0)

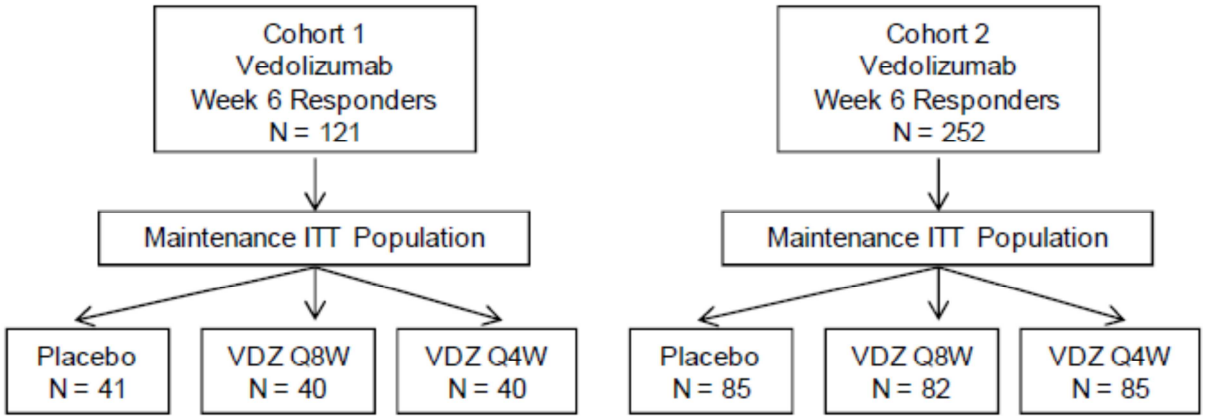
3.1.1.3.2.1 Analyses by Induction Cohort

To assess if the primary and secondary efficacy endpoints at Week 52 were affected by the patient's Induction Phase cohort, additional analyses were requested during a post-phase 3 Type C meeting held on July 24, 2012.

A total of 121 of 225 (54%) patients from Cohort 1 and 252 of 521 (48%) patients from Cohort 2 were randomized to the treatment and included in the Maintenance Study ITT Population.

The figure below displays the distribution of patients from each cohort randomized to the Maintenance Study ITT Population.

Figure 5 Overview of Patients in Cohort 1 and Cohort 2



Copied from Figure 3-1, page 13 FESA.

3.1.1.3.2.2 Clinical Remission at Week 52 by Induction Phase Cohort

The number and proportion of patients in the Maintenance Study ITT Population who achieved the protocol-specified definition of clinical remission at Week 52 are summarized by Induction Phase cohort in the table below.

Table 22 Clinical Remission at Week 52 by Induction Phase Cohort
Maintenance Study ITT Population
Study C13006

	Cohort 1			Cohort 2		
Clinical Remission,[a] n (%)	**PLA N = 41**	**VDZ Q8W N = 40**	**VDZ Q4W N = 40**	**PLA N = 85**	**VDZ Q8W N = 82**	**VDZ Q4W N = 85**
Number (%) achieving clinical remission	6 (14.6)	19 (47.5)	20 (50.0)	14 (16.5)	32 (39.0)	36 (42.4)
95% CI	(3.8, 25.5)	(32.0, 63.0)	(34.5, 65.5)	(8.6, 24.4)	(28.5, 49.6)	(31.8, 52.9)
Difference from placebo[b]		33.2	36.3		22.6	25.7
95% CI for difference from placebo		(13.1, 53.2)	(16.4, 56.2)		(9.2, 36.0)	(12.1, 39.3)
P-value for difference from placebo[c]		0.0012	0.0004		0.0009	0.0002
Relative risk[d]		3.3	3.5		2.4	2.6
95% CI for relative risk		(1.5, 7.4)	(1.6, 7.7)		(1.4, 4.1)	(1.5, 4.4)

Copied from Table 3-1, page 14 FESA.

As seen from the table above, greater proportions of patients treated with vedolizumab in the Q8W and Q4W treatment groups achieved clinical remission at Week 52 compared with patients

who received placebo, regardless of whether they were enrolled in Cohort 1 or 2 during the Induction Phase.

3.1.1.3.2.3 Durable Clinical Response by Induction Phase Cohort

The number and proportion of patients in the Maintenance Study ITT Population who achieved durable clinical response are summarized by Induction Phase cohort in the table below.

Table 23 Durable Clinical Response by Induction Phase Cohort
Maintenance Study ITT Population
Study C13006

	Cohort 1			Cohort 2		
Durable Clinical Response,[a] n (%)	**PLA N = 41**	**VDZ Q8W N = 40**	**VDZ Q4W N = 40**	**PLA N = 85**	**VDZ Q8W N = 82**	**VDZ Q4W N = 85**
Number (%) achieving durable clinical response	7 (17.1)	25 (62.5)	23 (57.5)	23 (27.1)	44 (53.7)	42 (49.4)
95% CI	(5.6, 28.6)	(47.5, 77.5)	(42.2, 72.8)	(17.6, 36.5)	(42.9, 64.5)	(38.8, 60.0)
Difference from placebo[b]		45.3	41.2		26.7	22.5
95% CI for difference from placebo		(24.1, 66.6)	(20.5, 61.9)		(12.3, 41.2)	(8.1, 36.8)
P-value for difference from placebo[c]		< 0.0001	< 0.0001		0.0003	0.0022
Relative risk[d]		3.7	3.4		2.0	1.8
95% CI for relative risk		(1.8, 7.5)	(1.7, 7.0)		(1.3, 2.9)	(1.2, 2.8)

Copied from Table 3-2, page 15 FESA.

As seen from Table above, greater proportions of patients treated with vedolizumab in the Q8W and Q4W treatment groups achieved durable clinical response compared with patients who received placebo, regardless of whether they were enrolled in Cohort 1 or 2 during the Induction Phase.

3.1.1.3.2.4 Mucosal Healing at Week 52 by Induction Phase Cohort

The number and proportion of patients in the Maintenance Study ITT Population who achieved mucosal healing at Week 52 are summarized by Induction Phase cohort in the table below.

Table 24 Mucosal Healing at Week 52 by Induction Phase Cohort
Maintenance Study ITT Population
Study C13006

	Cohort 1			Cohort 2		
Mucosal Healing,[a] n (%)	**PLA N = 41**	**VDZ Q8W N = 40**	**VDZ Q4W N = 40**	**PLA N = 85**	**VDZ Q8W N = 82**	**VDZ Q4W N = 85**
Number (%) achieving mucosal healing	6 (14.6)	24 (60.0)	24 (60.0)	19 (22.4)	39 (47.6)	46 (54.1)
95% CI	(3.8, 25.5)	(44.8, 75.2)	(44.8, 75.2)	(13.5, 31.2)	(36.8, 58.4)	(43.5, 64.7)
Difference from placebo[b]		45.8	45.9		25.4	31.7
95% CI for difference from placebo		(24.8, 66.7)	(25.1, 66.7)		(11.2, 39.6)	(17.2, 46.3)
P-value for difference from placebo [c]		< 0.0001	< 0.0001		0.0005	< 0.0001
Relative risk[d]		4.1	4.2		2.1	2.4
95% CI for relative risk		(1.9, 9.1)	(1.9, 9.1)		(1.4, 3.4)	(1.6, 3.8)

Copied from Table 3-3, page 16 FESA.

As seen from the table above, greater proportions of patients treated with vedolizumab in the Q8W and Q4W treatment groups achieved mucosal healing at Week 52 compared with patients who received placebo, regardless of whether they were enrolled in Cohort 1 or 2 during the Induction Phase.

3.1.1.3.2.5 Durable Clinical Remission by Induction Phase Cohort

The number and proportion of patients in the Maintenance Study ITT Population who achieved the protocol-specified definition of durable clinical remission are summarized by Induction Phase cohort in the table below.

Table 25 Durable Clinical Remission by Induction Phase Cohort
Maintenance Study ITT Population
Study C13006

Durable Clinical Remission,[a] n (%)	Cohort 1 PLA N = 41	Cohort 1 VDZ Q8W N = 40	Cohort 1 VDZ Q4W N = 40	Cohort 2 PLA N = 85	Cohort 2 VDZ Q8W N = 82	Cohort 2 VDZ Q4W N = 85
Number (%) achieving durable clinical remission	2 (4.9)	7 (17.5)	9 (22.5)	9 (10.6)	18 (22.0)	21 (24.7)
95% CI	(0.0, 11.5)	(5.7, 29.3)	(9.6, 35.4)	(4.0, 17.1)	(13.0, 30.9)	(15.5, 33.9)
Difference from placebo[b]		12.5	17.9		11.4	14.0
95% CI for difference from placebo		(-1.1, 26.1)	(3.1, 32.8)		(0.3, 22.6)	(2.7, 25.4)
P-value for difference from placebo[c]		0.0709	0.0176		0.0437	0.0155
Relative risk[d]		3.5	4.6		2.1	2.3
95% CI for relative risk		(0.8, 15.8)	(1.1, 19.7)		(1.0, 4.4)	(1.1, 4.8)

Copied from Table 3-4, page 18 FESA.

As seen from the table above, greater proportions of patients treated with vedolizumab in the Q8W and Q4W treatment groups achieved durable clinical remission compared with patients who received placebo, regardless of whether they were enrolled in Cohort 1 or 2 during the Induction Phase.

3.1.1.3.2.6 Corticosteroid-Free Clinical Remission at Week 52 by Induction Phase Cohort

The number and proportion of patients in the Maintenance Study ITT Population who achieved the protocol-specified definition of corticosteroid-free clinical remission at Week 52 are summarized by Induction Phase cohort in the table below.

Table 26 Corticosteroid-Free Clinical Remission at Week 52
Maintenance Study ITT Population
Study C13006

	Cohort 1			Cohort 2		
Corticosteroid-Free Clinical Remission,[a] n (%)	**PLA N = 25**	**VDZ Q8W N = 24**	**VDZ Q4W N = 24**	**PLA N = 47**	**VDZ Q8W N = 46**	**VDZ Q4W N = 49**
Number (%) achieving corticosteroid-free clinical remission	4 (16.0)	9 (37.5)	12 (50.0)	6 (12.8)	13 (28.3)	21 (42.9)
95% CI	(1.6, 30.4)	(18.1, 56.9)	(30.0, 70.0)	(3.2, 22.3)	(15.2, 41.3)	(29.0, 56.7)
Difference from placebo[b]		21.5	34.0		15.5	30.1
95% CI for difference from placebo		(-3.3, 46.2)	(7.7, 60.3)		(-0.9, 31.9)	(12.1, 48.0)
P-value for difference from placebo[c]		0.0887	0.0112		0.0633	0.0010
Relative risk[d]		2.3	3.1		2.2	3.4
95% CI for relative risk		(0.8, 6.7)	(1.2, 8.4)		(0.9, 5.3)	(1.5, 7.5)

Copied from Table 3-5, page 19 FESA.

As seen from the table above, greater proportions of patients treated with vedolizumab in the Q8W and Q4W treatment groups achieved corticosteroid-free clinical remission at Week 52 compared with patients who received placebo, regardless of whether they were enrolled treatment in Cohort 1 or 2 during the Induction Phase.

3.1.1.3.2.7 Sensitivity Analyses for the Primary Endpoint

Per the FDA's request, to address the issue of missing data, the applicant performed the sensitivity analyses (observed case, Per Protocol, LOCF, worst case, and multiple imputation), for the primary endpoint.

Per FDA's request, the applicant performed additional post-hoc sensitivity analysis assuming patients with less than 3 days of diary data within 7 days prior to study visit be classified as non-responders.

Summary of the sensitivity analyses results for clinical remission at Week 52 are given below.

Table 27 Clinical Remission at Week 52 – Sensitivity Analysis

Analysis Set	Placebo	VDZ 300 mg Q8W	Difference (VDZQ8w – placebo)	p-value	VDZ 300 mg Q4W	Difference (VDZQ4W – placebo)	p-value
Primary	20/126 (15.9%)	51/122 (41.8%)	26.1%	< 0.0001	56/125 (44.8%)	29.1%	< 0.0001
Per Protocol	20/121 (16.5%)	49/117 (41.9%)	25.7%	< 0.0001	55/121 (45.5%)	29.1%	< 0.0001
Observed Case	20/48 (41.7%)	51/77 (66.2%)	24.1%	0.0079	56/83 (67.5%)	25.0%	0.0042
ITT with Revised eDiary Requirements[a]	20/126 (15.9%)	48/122 (39.3%)	23.6%	<0.0001	56/125 (44.8%)	29.1%	<0.0001
LOCF[b]	27/126 (21.4%)	57/122 (46.7%)	25.4%	< 0.0001	58/125 (46.4%)	25.2%	< 0.0001
Worst Case[c]	98/126 (77.8%)	51/122 (41.8%)	- 36.0%	<0.0001	56/125 (44.8%)	- 32.9%	< 0.0001
Multiple Imputation	67/126 (53.2%)	91/122 (74.6%)	21.1%	0.0004	88/125 (70.4%)	17.1%	0.0039

Copied from Tables 53, 14.3.1.2CM, 14.3.1.2BM CSR and Tables 39.12.2.1D, 39.13.3.2D, 39.13.4.1D, and 39.13.5.1D Responses to Agency Questions (Questions Received 8/19/13).

[a] ITT population where patients with < 3 days of diary data within 7 days prior to Week 52 are classified as non-responders Table 39.12.2.1D

[b] Last Observation Carried Forward (LOCF) analysis imputed data from the prior time point, if a subject had missing data at a particular time point.

[c] Worst Case analysis assumed patients receiving placebo who had missing data to be responders and patients receiving vedolizumab who had missing data to be non-responders

As seen from the tables above, results from sensitivity analyses with exception of Worst Case analysis favored consistently vedolizumab against the placebo for clinical remission at Week 52.

3.1.1.3.2.8 Alternative Definition of Clinical Remission by the FDA

In Study C13006, clinical remission, defined as a complete Mayo score of ≤ 2 points and no individual subscore > 1 point, was the primary endpoint in the Maintenance Study (Week 52). Durable clinical remission was a secondary endpoint in the Maintenance Study, where patients met the definition of clinical remission at both Week 6 and Week 52.

Additional analyses had been conducted based on an alternative definition of clinical remission proposed by the FDA. This definition is defined as a complete Mayo score of ≤ 2 points and no individual subscore > 1 point where rectal bleeding subscore = 0 and endoscopy subscore = 0. Analyses results with the alternative definition of clinical remission are provided for Week 52 and for durable clinical remission (clinical remission at both Week 6 and Week 52). In addition,

results on corticosteroid-free clinical remission at Week 52 are also presented using the alternative definition.

3.1.1.3.2.8.1 Clinical Remission at Week 52

The proportion of patients who achieved the alternative definition of clinical remission at Week 52 is summarized by treatment group for the Maintenance Study ITT population in the table below.

Table 28 Clinical Remission (Alternative Definition Proposed by FDA) at Week 52
Maintenance Study ITT Population
Study C13006

Clinical Remission,[a] n (%)	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125
Number (%) achieving clinical remission	11 (8.7)	32 (26.2)	36 (28.8)
95% CI	(3.8, 13.7)	(18.4, 34.0)	(20.9, 36.7)
Difference from placebo[b]		17.6	20.2
95% CI for difference from placebo		(8.3, 27.0)	(10.7, 29.8)
P-value for difference from placebo[c]		0.0002	< 0.0001
Relative risk[d]		3.0	3.3
95% CI for relative risk		(1.6, 5.8)	(1.8, 6.3)

Copied from Table 3-7, page 22 FESA.

As seen from the table above, greater proportions of vedolizumab-treated patients in both the Q8W and Q4W treatment groups achieved the alternative definition of clinical remission at Week 52 compared with the placebo group.

3.1.1.3.2.8.2 Durable Clinical Remission

The proportion of patients who achieved the alternate definition of durable clinical remission is summarized by treatment group for the Maintenance Study ITT Population in the table below.

Table 29 Durable Clinical Remission (Alternative Definition Proposed by FDA)
Maintenance Study ITT Population
Study C13006

Durable Clinical Remission,[a] n (%)	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125
Number (%) achieving durable clinical remission	2 (1.6)	6 (4.9)	12 (9.6)
95% CI	(0.0, 3.8)	(1.1, 8.8)	(4.4, 14.8)
Difference from placebo[b]		4.8	12.6
95% CI for difference from placebo		(-1.4, 11.0)	(3.8, 21.4)
P-value for difference from placebo[c]		0.1305	0.0051
Relative risk[d]		3.1	6.0
95% CI for relative risk		(0.6, 15.2)	(1.4, 26.4)

Copied from Table 3-8, page 23 FESA.

As seen from the table above, a greater proportion of patients in the vedolizumab Q4W group achieved the alternate definition of durable clinical remission compared with the placebo group. No treatment difference was observed for the vedolizumab Q8W treatment group versus placebo.

3.1.1.3.2.8.3 Corticosteroid-Free Clinical Remission at Week 52

The number and proportion of patients in the Maintenance Study ITT Population who achieved corticosteroid-free clinical remission at Week 52, based on the alternative definition of clinical remission, are summarized by treatment group in the table below.

Table 30 Corticosteroid-Free Clinical Remission (Alternative Definition Proposed by FDA) at Week 52
Maintenance Study ITT Population
Study C13006

Corticosteroid-Free Clinical Remission,[a] n (%)	PLA N = 72	VDZ Q8W N = 70	VDZ Q4W N = 73
Number (%) achieving corticosteroid-free clinical remission	5 (6.9)	11 (15.7)	18 (24.7)
95% CI	(1.1, 12.8)	(7.2, 24.2)	(14.8, 34.5)
Difference from placebo[b]		8.8	17.7
95% CI for difference from placebo		(-1.6, 19.1)	(5.8, 29.6)
P-value for difference from placebo[c]		0.0981	0.0035
Relative risk[d]		2.3	3.6
95% CI for relative risk		(0.8, 6.3)	(1.4, 9.2)

Copied from Table 3-9, page 24 FESA.

As seen from the table above, greater proportions of patients treated with vedolizumab in the Q8W and Q4W treatment groups achieved corticosteroid-free clinical remission at Week 52 using the alternate definition compared with patients who received placebo. A trend was observed for the vedolizumab Q8W group.

3.1.1.3.2.9 Clinical Remission at Week 52 in Subgroups based on Anti-TNF Status (Inadequate Response/Loss of Response)

Per the medical officer's request, the applicant performed a subgroup analysis of clinical remission at Week 52 by subgroups based on anti-TNF failure status (inadequate response/loss of response).

Table 31 Clinical Response at Week 52 – Evaluation in Subgroups Based on Anti-TNF Failure Status (Inadequate Response/Loss of Response) Study C13006

Endpoint	Patients Without Prior Failure[a]			Patients with Prior Failure[a]		
	Placebo	VDZ 300 mg Q8W	VDZ 300 mg Q4W	Placebo	VDZ 300 mg Q8W	VDZ 300 mg Q4W
Patients who failed to respond to or lost response to TNFα agents						
N	94	90	93	32	32	32
Number (%) achieving remission	18 (19.1)	41 (45.6)	48 (51.6)	2 (6.3)	10 (31.3)	8 (25.0)
95% CI	(11.2, 27.1)	(35.3, 55.8)	(41.5, 61.8)	(0.8, 20.8)	(16.1, 50.0)	11.5 (43.4)
Difference from placebo		26.4	32.5		25.0	18.8
95% CI for Difference From Placebo		(13.4, 39.4)	(19.6, 45.4)		(-1.0, 48.6)	(-7.3, 43)

Copied from Table 39.20.1.6 page 8 Response to Agency Questions Received September 20, 2013.

[a] Protocol defined TNFα failure includes those patients who failed to respond to, lost response to, or were intolerant of TNFα agents

[b] Failure as defined by each of the subgroup analyses: 1). failed to respond to, lost response to, or become intolerant of TNFα agents or 2). failed to respond to or lost response to TNFα agents.

As seen from the table above, the 95% confidence intervals for difference from placebo included zero for patients who were prior anti-TNF failure. For patients who were not prior anti-TNF failure, the 95% confidence intervals excluded zero.

3.2 Evaluation of Safety

3.2.1 Induction Phase

The Induction Phase Safety population was analyzed by the number of completed infusions and by days on study. The start of the Induction Phase was the date of the first dose of study drug received, as recorded on the eCRF. The end of the Induction Phase was defined differently for patients who entered the Maintenance Phase and those who did not. For patients who continued into the Maintenance Phase, the end of the Induction Phase was on the calendar day before the date of the Week 6 visit. For patients who did not continue into the Maintenance Phase, the end of the Induction Phase was the calendar day of the Week 6 visit, or if they did not complete the Induction Phase, the date of early termination.

Table 32 Exposure to Study Medication during the Induction Phase
Induction Phase Safety Population
Study C13006

	Induction Cohort 1 ITT Population[a]		Induction Cohort 2[b] Open-label	
	PLA N = 149	VDZ Cohort 1 N = 225	VDZ Cohort 2 N = 521	VDZ Combined N = 746
Patients who received infusions[c], n (%)				
Patients who received 1 infusion	6 (4)	3 (1)	11 (2)	14 (2)
Patients who received 2 infusions	143 (96)	222 (99)	510 (98)	732 (98)
Days on study				
n	149	225	521	746
Mean (std dev)	40.2 (10.31)	42.5 (6.44)	41.1 (8.52)	41.5 (7.97)
Median	43.0	43.0	43.0	43.0
Minimum, maximum	1, 52	1, 67	1, 57	1, 67

Copied from Table 29, page 149 CSR.

An overall summary of AEs during the Induction Phase is presented for the Induction Phase Safety population in the table below.

Table 33 Overall Summary of Adverse Events during the Induction Phase
Induction Phase Safety Population -Study C13006

	Induction Cohort 1 ITT Population[a]		Induction Cohort 2[b] Open-label	
Adverse Event Category, n (%)	PLA N = 149	VDZ N = 225	VDZ N = 521	VDZ Combined N = 746
Any adverse event	69 (46)	90 (40)	247 (47)	337 (45)
Drug-related adverse events	25 (17)	35 (16)	102 (20)	137 (18)
Adverse event resulting in study discontinuation	4 (3)	0	8 (2)	8 (1)
Serious adverse event	10 (7)	5 (2)	20 (4)	25 (3)
Serious infection adverse events	3 (2)	1 (< 1)	3 (< 1)	4 (< 1)
Drug-related serious adverse event	3 (2)	1 (< 1)	3 (< 1)	4 (< 1)
Serious adverse event resulting in discontinuation	4 (3)	0	6 (1)	6 (< 1)
Death	0	0	1 (< 1)	1 (< 1)

Copied from Table 30, page 150 CSR.

As seen from the table above, the proportions of patients who experienced at least 1 AE during the study in the placebo, ITT vedolizumab, and Cohort 2 vedolizumab groups were similar. Drug-related AEs, as deemed by the investigator, were reported in similar proportions by both treatment groups in the Induction Study ITT population. A slightly higher percentage of drug-related AEs were reported among patients in the open-label Cohort 2 than in either ITT group. Adverse events that resulted in study discontinuation were reported by 12 patients (1%), 4 patients (3%) in the placebo group and 8 patients (2%) in the vedolizumab Cohort 2 group. No patients in the Cohort 1 vedolizumab group experienced AEs that led to study discontinuation.

The proportions of patients who experienced at least one serious adverse event (SAE) in the placebo, ITT vedolizumab, and Cohort 2 vedolizumab groups were 7%, 2%, and 4%, respectively. A total of 35 patients (4%) experienced at least one SAE; serious infection AEs and drug-related SAEs occurred in < 1% of patients in both vedolizumab treatment groups; SAEs were more frequent in the placebo group (7%) than in the vedolizumab ITT group (2%). Ten patients (1%) experienced an SAE that resulted in study discontinuation, 4 patients (3%) in the placebo group and 6 patients (1%) in the Cohort 2 group. No patients in Cohort 1 treated with vedolizumab experienced SAEs that led to study discontinuation.

One death was reported during the Induction Phase (Patient C13006-46007-608, vedolizumab Cohort 2).

3.2.2 Maintenance Phase

The Maintenance Phase Safety population was used for all safety analyses. The safety analyses in this section are cumulative, including data from both the Induction Phase and the Maintenance Phase. The Maintenance Phase Safety population includes safety data from all patients from Week 0 through study completion.

The table below summarized cumulative exposure to study medication from Week 0 through the last dose of the study drug for all patients as the number of completed infusions. The table also summarized the cumulative exposure in days from Week 0 to the end of the Maintenance Phase.

Table 34 Exposure to Study Medication – Number of Completed Infusion and Exposure in Days During Induction and Maintenance Phase
Maintenance Phase Safety Population
Study C13006

	Maintenance Study ITT[a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Maintenance Non-ITT		Combined	
	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125	PLA[b] (from Week 0) N = 149	VDZ Q4W[c] (Week 6 Nonresponders) N = 373	PLA N = 275	VDZ N = 620
Number of completed infusions							
≥ 1	126 (100)	122 (100)	125 (100)	149 (100)	373 (100)	275 (100)	620 (100)
≥ 2	126 (100)	122 (100)	125 (100)	144 (97)	361 (97)	270 (98)	608 (98)
≥ 3	126 (100)	122 (100)	125 (100)	135 (91)	327 (88)	261 (95)	574 (93)
≥ 4	120 (95)	121 (99)	121 (97)	119 (80)	298 (80)	239 (87)	540 (87)
≥ 5	110 (87)	113 (93)	118 (94)	93 (62)	264 (71)	203 (74)	495 (80)
≥ 6	99 (79)	101 (83)	108 (86)	65 (44)	214 (57)	164 (60)	423 (68)
≥ 7	87 (69)	91 (75)	101 (81)	55 (37)	198 (53)	142 (52)	390 (63)
≥ 8	77 (61)	87 (71)	98 (78)	50 (34)	183 (49)	127 (46)	368 (59)
≥ 9	70 (56)	85 (70)	94 (75)	43 (29)	170 (46)	113 (41)	349 (56)
≥ 10	64 (51)	81 (66)	89 (71)	41 (28)	157 (42)	105 (38)	327 (53)
≥ 11	61 (48)	81 (66)	88 (70)	37 (25)	148 (40)	98 (36)	317 (51)
≥ 12	54 (43)	78 (64)	86 (69)	32 (21)	145 (39)	86 (31)	309 (50)
≥ 13	50 (40)	76 (62)	85 (68)	32 (21)	141 (38)	82 (30)	302 (49)
≥ 14	46 (37)	71 (58)	82 (66)	31 (21)	134 (36)	77 (28)	287 (46)
Exposure (days)[d]							
Mean (std dev)	242.7 (113.56)	285.5 (110.70)	295.9 (107.41)	181.4 (118.40)	237.2 (119.02)	209.5 (119.96)	258.5 (117.98)
Median	255.0	351.0	351.0	127.0	213.0	180.0	345.0
Min, max	41, 470	40, 477	43, 477	39, 470	43, 474	39, 470	40, 477

Source: Table 14.1.1.9FM, Table 14.1.1.9BM

Copied from Table 72, page 248 CSR.

As seen from the table above, in the Maintenance ITT population, exposure to study medication was higher for the vedolizumab treatment groups than for the placebo treatment group when assessed by either cumulative number of infusions or exposure in days during the study which reflects the higher premature withdrawal rate in the placebo group. More than half of the patients in the vedolizumab treatment groups completed all infusions (Q8W: 58%; Q4W: 66%), while 37% of placebo patients completed all 14 planned infusions. These differences in exposure

between vedolizumab and placebo patients are also reflected in the median days on study, with a median of 351.0 days on study for both the vedolizumab Q8W and Q4W groups and 255.0 days on study for the placebo patients.

Cumulative exposure to study medication was similarly lower for patients in the non-ITT placebo group, who received placebo throughout the entire study, compared to patients in the combined vedolizumab group, who received vedolizumab throughout the entire study. Twenty-one percent (21%) of the non-ITT placebo patients completed all 14 planned infusions, compared with 46% of combined vedolizumab patients. The exposure in days from Week 0 to end of study was also lower for the non-ITT placebo group (median 127.0 days) compared to the combined vedolizumab group (median 345.0 days).

An overall summary of AEs is presented in the table below.

Table 35 Overall Summary of Adverse Events
Study C13006

	Maintenance Study ITT[a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Maintenance Non-ITT		Combined	
Adverse Event Category	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125	PLA[b] (from Week 0) N = 149	VDZ Q4W[c] (Week 6 Nonresponders) N = 373	PLA N = 275	VDZ N = 620
Any adverse event	106 (84)	100 (82)	101 (81)	114 (77)	296 (79)	220 (80)	497 (80)
Drug-related adverse events	40 (32)	37 (30)	37 (30)	38 (26)	126 (34)	78 (28)	200 (32)
Adverse event resulting in study discontinuation	15 (12)	7 (6)	6 (5)	16 (11)	23 (6)	31 (11)	36 (6)
Serious adverse event	20 (16)	10 (8)	11 (9)	17 (11)	56 (15)	37 (13)	77 (12)
Serious infection adverse events	4 (3)	3 (2)	2 (2)	4 (3)	7 (2)	8 (3)	12 (2)
Drug-related serious adverse event	4 (3)	3 (2)	1 (< 1)	3 (2)	9 (2)	7 (3)	13 (2)
Serious adverse event resulting in discontinuation	7 (6)	2 (2)	0	6 (4)	14 (4)	13 (5)	16 (3)
Deaths	0	0	0	0	1 (< 1)	0	1 (< 1)

Copied from Table 73, page 251 CSR.

As seen from the table above, in the Maintenance Study ITT population, one or more AEs were experienced by 84% of placebo patients, 82% of Q8W patients, and 81% of Q4W patients.

Drug-related AEs, as determined by the investigator, were reported with a similar incidence across the Maintenance Study ITT groups (placebo: 32%; 30% each from Q8W and Q4W). Rates of discontinuation from treatment due to AEs in the ITT placebo group were twice of those that were observed in the ITT vedolizumab-treatment groups (placebo: 12%; Q8W: 6%; Q4W: 5%).

In the Maintenance Study ITT population, SAEs were approximately twice as frequent in patients receiving placebo (16%) than among those enrolled in either vedolizumab dosing

regimen (Q8W: 8%; Q4W: 9%). Serious infection AEs, drug-related SAEs, and SAEs resulting in study discontinuation each occurred in ≤ 2% of vedolizumab-treated patients in the Maintenance ITT population. Serious AEs resulting in study discontinuation were more than twice as frequent among patients receiving placebo in the Maintenance Study ITT population (6%) than among those receiving vedolizumab (Q8W: 2%; Q4W:0%).

At least one AE was experienced by 80% of patients in the combined vedolizumab group (who received vedolizumab for the entire duration of the study) and by 77% of patients in the non- ITT placebo group (who received placebo treatment for the entire duration of the study). The rates of study discontinuation due to AEs were twice as high in the non-ITT placebo group as compared to those in the combined vedolizumab group (11% vs. 6%). Similar frequencies of SAEs were reported among patients in the non-ITT placebo group (11%) and in the combined vedolizumab group (12%). Serious infection AEs, drug-related SAEs, and SAEs resulting in study discontinuation also occurred with similar frequencies in the non-ITT placebo group and in the combined vedolizumab group.

One death was reported in a vedolizumab-treated patient in the Maintenance Study Safety population; this death occurred during the Induction Phase (Patient C13006-46007-608).

4 FINDINGS IN SPECIAL/SUBGROUP POPULATION

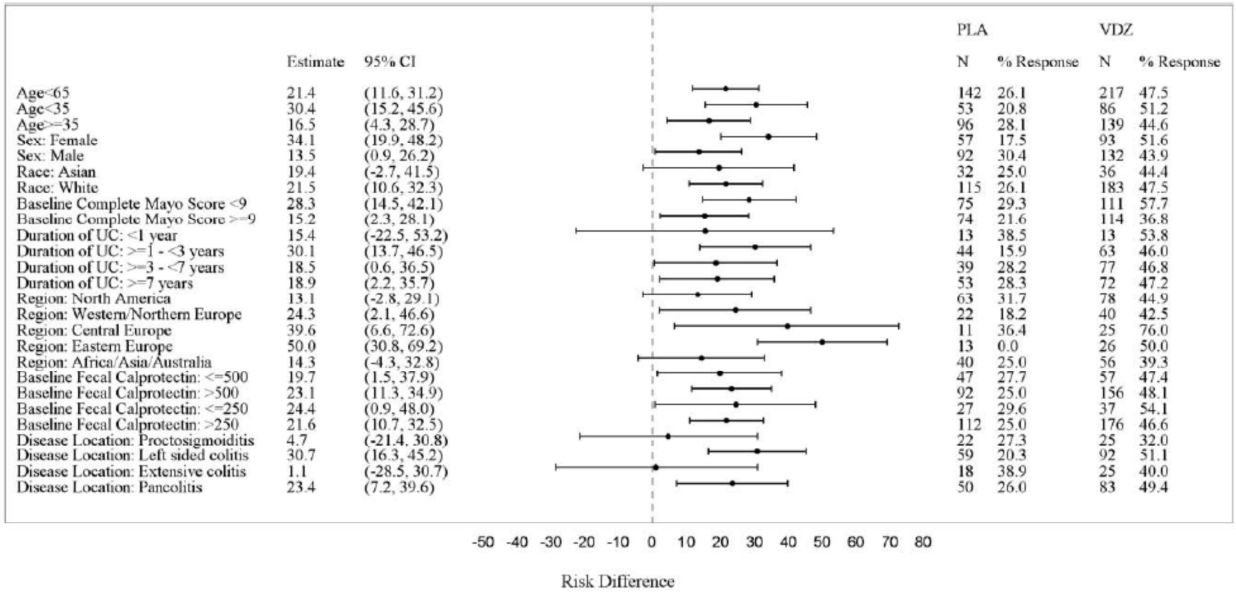
4.1 Gender, Race, Age, Other Special/Subgroup Population

Subgroup analyses for clinical response at Week 6 in the Induction Study ITT population and clinical remission at Week 52 in the Maintenance Study ITT population are provided based on: age (< 35 vs. ≥ 35 years; < 65 vs. ≤ 65 years), gender, race, duration from UC diagnosis to first dose, geographic region, baseline (Week 0) disease activity, baseline (Week 0) fecal calprotectin (≤ 250 µg/g vs. > 250 µg/g; ≤ 500 µg/g vs. > 500 µg/g), and disease localization.

4.1.1 Induction Phase

The risk differences compared with placebo for the primary endpoint, clinical response, in subgroups according to demographic characteristics and measures of disease severity are summarized in the figures below:

Figure 6 Risk Difference and 95% Confidence Interval for Subgroup Analyses of Clinical Response at Week 6
Induction Study ITT Population
Study C13006



Copied from Figure 6, page 124 CSR.

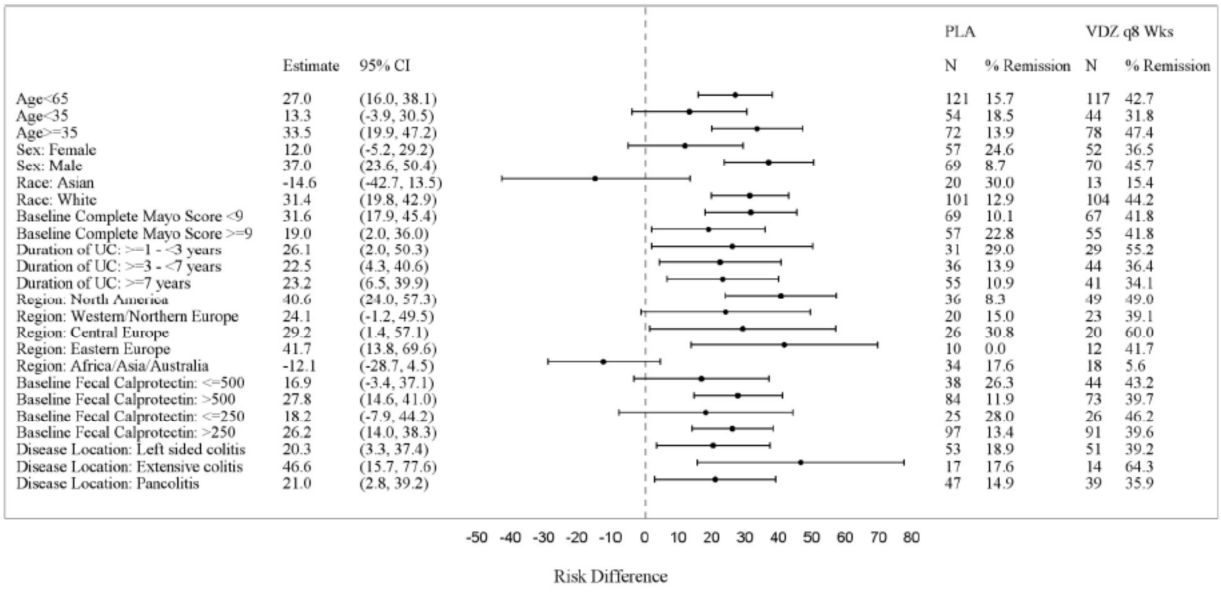
As seen from the figure above, the treatment benefit of vedolizumab as measured by the primary endpoint was statistically significant across age categories (< 35 and ≥ 35 years) and gender. For each of the other demographic subgroups (according to race and region), the risk differences consistently favored vedolizumab over placebo, although due to small sample sizes in some subgroups, not all of the treatment differences were significant. There were no apparent differences in the magnitude of treatment benefit in these subgroups.

Similar results were observed for subgroups according to assessments of disease severity, including categories of baseline fecal calprotectin, Mayo score category, and disease localization.

4.1.2 Maintenance Phase

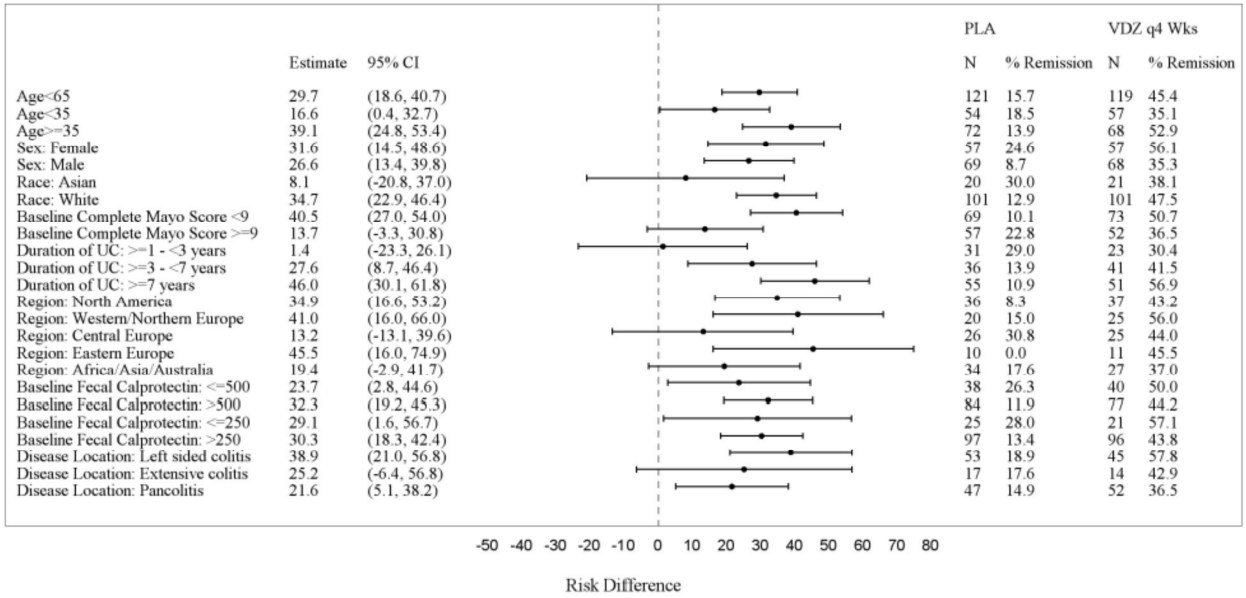
The risk differences compared with placebo for the primary endpoint, clinical response, in subgroups according to demographic characteristics and measures of disease severity are summarized in the figures below:

Figure 7 Risk Difference and 95% Confidence Interval for Subgroup Analyses of Clinical Remission at Week 52 for VDZ Q8W vs. Placebo
Maintenance Study ITT Population
Study C13006



Copied from Figure 12, page 199 CSR.

Figure 8 Risk Difference and 95% Confidence Interval for Subgroup Analyses of Clinical Remission at Week 52 for VDZ Q4W vs. Placebo
Maintenance Study ITT Population
Study C13006



Copied from Figure 13, page 200 CSR.

As seen from the figures above, for either dose regimen (Q8W and Q4W), the treatment benefit of vedolizumab as measured by the primary endpoint was statistically significant for age ≥ 35 years and for males. For each of the other demographic subgroups (according to race and region), the risk differences consistently favored vedolizumab over placebo, although due to small sample sizes in some subgroups, not all of the treatment differences were significant. There were no apparent differences in the magnitude of treatment benefit in these subgroups.

Similar results were observed for subgroups according to assessments of disease severity, including categories of baseline fecal calprotectin, Mayo score category, and disease localization.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

Induction Phase

Additional analyses have been conducted based on an alternative definition of clinical remission proposed by the FDA. The new definition is defined as a complete Mayo score of ≤ 2 points and no individual subscore > 1 point where rectal bleeding subscore = 0 and endoscopy subscore = 0. Results revealed that, fewer patients in either treatment group achieved this alternative definition of clinical remission. Furthermore, no treatment difference was noted between the vedolizumab (4.4%) and placebo (2.7%) groups in the Induction Study ITT Population for the proportion of patients who achieved the alternate definition of clinical remission at Week 6.

Per the FDA's request, for this study the applicant provided exploratory analyses of clinical remission based on the following four alternative endpoint definitions for clinical remission.

a. Endoscopy subscore = 0, Rectal Bleeding subscore = 0, and Stool Frequency subscore decreases or no change from Baseline (all assessed at Week 6)
b. Endoscopy subscore ≤ 1, Rectal Bleeding subscore = 0, and Stool frequency subscore = 0 (all assessed at Week 6)
c. Endoscopy subscore ≤ 1, Rectal Bleeding subscore = 0, and Stool frequency subscore ≤ 1 (all assessed at Week 6)
d. Endoscopy subscore ≤ 1, Rectal Bleeding subscore = 0, Stool Frequency subscore decreases or no change from Baseline, and Total score ≤ 1 (all assessed at Week 6)

Results showed that: there was a trend favoring vedolizumab over placebo for these four alternative definitions of clinical remission at Week 6.

The applicant defined mucosal healing as Mayo endoscopic subscore of ≤ 1 point and provided no histologic data to support a mucosal healing claim. Based on the applicant's definition, 40.9% of patients in the vedolizumab treatment group achieved mucosal healing, compared with 24.8% of patients receiving placebo, a 16.1% treatment difference (95% CI: 1.2, 2.3; p = 0.0012) was observed. When focusing only on the subset of patients who had an endoscopy subscore of 0 at Week 6, which indicates normal or inactive disease, there was no significant treatment difference observed 0.9%; (95 CI: -3.4, 5.1; p=0.6956) for the mucosal healing endpoint.

Per the medical officer's request, this review provided summary of this subgroup analysis for combined Induction Phase Cohorts (Cohort 1 and Cohort 2). Results show that the 95% confidence intervals overlapped for patients who had prior anti-TNF failures. For patients who did not have not prior anti-TNF failures, the 95% confidence intervals did not overlap.

All the analyses noted above were post-hoc sensitivity or subgroup analysis analyses. In general, the results from the sensitivity and subgroup analyses are consistently in favor of vedolizumab.

Maintenance Phase

To assess if the primary and secondary efficacy endpoints at Week 52 were affected by the Induction Phase Cohorts, additional analyses were requested during a Type C meeting held on July 24, 2012, after the phase 3 studies were completed.

Results from these analyses show that for clinical remission at Week 52 and durable clinical response at Week 52, vedolizumab in the Q8W and Q4W treatment groups showed a treatment effect compared to placebo, regardless of whether patients were enrolled in Cohort 1 or 2 during the Induction Phase.

Per the medical officer's request, the applicant performed a subgroup analysis of clinical remission at Week 52 in subgroups based on anti-TNF failure status (inadequate response/loss of response). Results revealed that the 95% confidence intervals for the treatment differences of each of vedolizumab dose regimen from placebo included zero for patients who were prior anti-TNF failures. For patients who were not prior anti-TNF failures, the 95% confidence intervals excluded zero.

It should be noted that with more than 60% of the data missing at Week 6 for placebo and more than 30% data missing at Week 52 for vedolizumab, the observed treatment difference might be overestimated when imputing all missing as non-responders. However, most of the missing data were due to lack of efficacy or adverse event.

5.2 Conclusion and Recommendation

Only one pivotal study, Study C13006, was conducted to support Ulcerative Colitis (UC) for induction for vedolizumab 300 mg intravenous (IV) infusions. This study showed that in the Induction Phase, vedolizumab was statistically significant better than placebo in clinical response at Week 6 and clinical remission at Week 6 with treatment difference of 22% and 12%, respectively. But trend was observed for clinical remission at Week 6 for more stringent definition of clinical remission (including endoscopy subscore=0).

In Maintenance Phase, both vedolizumab dose regimens demonstrated statistically significant benefit compared to placebo in clinical remission at Week 52 and durable clinical response (clinical responses at both Weeks 6 and 52). However, with more than 60% data missing for placebo and more than 30% data missing for vedolizumab, treatment effect might be overstated.

6 APPENDIX

Table 1 Baseline Demographics – Induction Phase Safety Population –Study C13006

	Induction Study ITT[a]		Non-ITT		
Parameter	PLA N = 149	VDZ Cohort 1 N = 225	VDZ Cohort 2[b] N = 521	VDZ Combined N = 746	Total N = 895
Gender, n (%)					
Male	92 (62)	132 (59)	301 (58)	433 (58)	525 (59)
Female	57 (38)	93 (41)	220 (42)	313 (42)	370 (41)
Race, n (%)					
White	115 (77)	183 (81)	436 (84)	619 (83)	734 (82)
Black	2 (1)	5 (2)	5 (< 1)	10 (1)	12 (1)
Asian	32 (21)	36 (16)	67 (13)	103 (14)	135 (15)
Other	0	1 (< 1)	13 (2)	14 (2)	14 (2)
Ethnicity, n (%)					
Hispanic or Latino	5 (3)	10 (4)	31 (6)	41 (5)	46 (5)
Not Hispanic or Latino	140 (94)	211 (94)	481 (92)	692 (93)	832 (93)
Not reported	4 (3)	4 (2)	9 (2)	13 (2)	17 (2)
Age (years)[c]					
n	149	225	521	746	895
Mean (std dev)	41.2 (12.50)	40.1 (13.11)	40.1 (13.27)	40.1 (13.21)	40.3 (13.09)
Median	40.7	39.4	38.4	38.7	39.4
Minimum, maximum	19, 76	18, 73	18, 78	18, 78	18, 78

Table 1 Baseline Demographics – Induction Phase Safety Population (continued) –Study C13006

Parameter	Induction Study ITT[a]		Non-ITT		
	PLA N = 149	VDZ Cohort 1 N = 225	VDZ Cohort 2[b] N = 521	VDZ Combined N = 746	Total N = 895
Age (years), n (%)					
< 35	53 (36)	86 (38)	214 (41)	300 (40)	353 (39)
≥ 35	96 (64)	139 (62)	307 (59)	446 (60)	542 (61)
Age (years), n (%)					
< 65	142 (95)	217 (96)	503 (97)	720 (97)	862 (96)
≥ 65	7 (5)	8 (4)	18 (3)	26 (3)	33 (4)
Body weight (kg)					
n	148	225	521	746	894
Mean (std dev)	72.4 (17.65)	72.4 (17.11)	74.2 (19.32)	73.6 (18.68)	73.4 (18.51)
Median	71.1	72.0	71.0	71.0	71.0
Minimum, maximum	40, 125	33, 118	32, 174	32, 174	32, 174
BMI (kg/m^2)					
n	148	225	521	746	894
Mean (std dev)	24.6 (5.11)	24.9 (4.85)	25.3 (6.05)	25.2 (5.72)	25.1 (5.62)
Median	23.9	24.3	24.1	24.1	24.1
Minimum, maximum	15, 44	15, 39	14, 61	14, 61	14, 61
Geographic region, n (%)					
North America	63 (42)	78 (35)	189 (36)	267 (36)	330 (37)
Western/Northern Europe	22 (15)	40 (18)	112 (21)	152 (20)	174 (19)
Central Europe	11 (7)	25 (11)	83 (16)	108 (14)	119 (13)
Eastern Europe	13 (9)	26 (12)	37 (7)	63 (8)	76 (8)
Asia/Australia/Africa	40 (27)	56 (25)	100 (19)	156 (21)	196 (22)

Source: Table 14.1.1.5CP.

Abbreviations: BMI = body mass index; PLA = placebo; std dev = standard deviation; VDZ = vedolizumab;.

a All patients enrolled in Cohort 1 who were randomized to blinded induction treatment with vedolizumab or placebo.

b All patients enrolled in Cohort 2 who received open-label vedolizumab induction treatment.

c Age is defined as (1 + first dose date - birth date) / 365.25.

Table 2 Comparison by Treatment Arm of Selected Baseline Demographics Characteristics – Induction Study ITT Population – Study C13006

Parameter	PLA N = 149	VDZ N = 225	P value[b]
Gender, n (%)			
Male	92 (62)	132 (59)	0.5521
Female	57 (38)	93 (41)	
Race, n (%)			
White	115 (77)	183 (81)	0.3286
Other	34 (23)	42 (19)	
Age (years) [a]			
Mean (std dev)	41.2 (12.50)	40.1 (13.11)	0.3758
Body weight (kg)			
Mean (std dev)	72.4 (17.65)	72.4 (17.11)	0.8926
Geographic region, n (%)			
North America	63 (42)	78 (35)	0.1530
Europe (Western, Central and Eastern)	46 (31)	91 (40)	
Asia/Australia/Africa	40 (27)	56 (25)	

Source: Table 14.1.1.5B.

Abbreviations: ITT = intent-to-treat; PLA = placebo; std dev = standard deviation; VDZ = vedolizumab.

a Age is defined as (1 + first dose date – birth date)/ 365.25.

b P values for categorical variables are from chi-square test and for continuous variables are from Kruskal Wallis test.

Table 3 Baseline Ulcerative Colitis Disease Characteristics – Induction Phase Safety Population – Study 13006

Parameter	Induction Cohort 1 ITT Population[a]		Induction Cohort 2[b] Open-label	VDZ	
	PLA N = 149	VDZ N = 225	VDZ N = 521	Combined N = 746	Total N = 895
Duration of ulcerative colitis (years) [c]					
n	149	225	519	744	893
Mean (std dev)	7.1 (7.25)	6.1 (5.08)	7.2 (6.61)	6.8 (6.20)	6.9 (6.39)
Median	4.5	4.6	5.1	5.0	4.9
Minimum, maximum	0.5, 38.5	0.5, 25.8	0.5, 37.5	0.5, 37.5	0.5, 38.5
Categorical duration of ulcerative colitis, n (%)					
< 1 year	13 (9)	13 (6)	38 (7)	51 (7)	64 (7)
≥ 1 - < 3 years	44 (30)	63 (28)	121 (23)	184 (25)	228 (25)
≥ 3 - < 7 years	39 (26)	77 (34)	163 (31)	240 (32)	279 (31)
≥ 7 years	53 (36)	72 (32)	197 (38)	269 (36)	322 (36)
Missing	0	0	2	2	2
Baseline disease activity [d]					
n	149	225	521	746	895
Mean (std dev)	8.6 (1.68)	8.5 (1.78)	8.6 (1.76)	8.6 (1.76)	8.6 (1.75)
Median	8.0	9.0	8.0	8.0	8.0
Minimum, maximum	3, 12	5, 12	3, 12	3, 12	3, 12
Baseline disease activity, n (%)					
Complete Mayo score < 6	5 (3)	6 (3)	14 (3)	20 (3)	25 (3)
Complete Mayo score of 6 to 8 (inclusive)	70 (47)	105 (47)	249 (48)	354 (47)	424 (47)
Complete Mayo score of 9 to 12 (inclusive)	74 (50)	114 (51)	258 (50)	372 (50)	446 (50)
Baseline fecal calprotectin (µg/g)					
n	139	213	505	718	857
Mean (std dev)	2369.9 (3258.82)	2552.2 (3800.36)	1442.7 (1855.61)	1771.8 (2635.91)	1868.8 (2753.28)
Median	1005.5	1111.9	782.3	867.9	898.9
Minimum, maximum	23.8, 16443.8	23.8, 20000.0	23.8, 12333.0	23.8, 20000.0	23.8, 20000.0
Categorical baseline fecal calprotectin					
≤ 250 µg/g	27 (18)	37 (16)	94 (18)	131 (18)	158 (18)

Table 3 Baseline Ulcerative Colitis Disease Characteristics – Induction Phase Safety Population (continued)– Study 13006

Parameter	Induction Cohort 1 ITT Population[a] PLA N = 149	Induction Cohort 1 ITT Population[a] VDZ N = 225	Induction Cohort 2[b] Open-label VDZ N = 521	VDZ Combined N = 746	Total N = 895
> 250 to ≤ 500 µg/g	20 (13)	20 (9)	82 (16)	102 (14)	122 (14)
> 500 µg/g	92 (62)	156 (69)	329 (63)	485 (65)	577 (64)
Missing	10	12	16	28	38
Disease localization, n (%)					
Proctosigmoiditis	22 (15)	25 (11)	69 (13)	94 (13)	116 (13)
Left-sided colitis	59 (40)	92 (41)	188 (36)	280 (38)	339 (38)
Extensive colitis	18 (12)	25 (11)	66 (13)	91 (12)	109 (12)
Pancolitis	50 (34)	83 (37)	198 (38)	281 (38)	331 (37)
Smoking status, n (%)					
Current smoker	11 (7)	12 (5)	32 (6)	44 (6)	55 (6)
Nonsmoker	88 (59)	145 (64)	322 (62)	467 (63)	555 (62)
Former smoker	50 (34)	68 (30)	167 (32)	235 (32)	285 (32)
History of extraintestinal manifestations					
Yes	44 (30)	74 (33)	180 (35)	254 (34)	298 (33)
No	105 (70)	151 (67)	341 (65)	492 (66)	597 (67)

Source: Table 14.1.1.6CP.

Abbreviations: PLA = placebo; std dev = standard deviation; VDZ = vedolizumab.

a All patients enrolled in Cohort 1 who were randomized to blinded induction treatment with vedolizumab or placebo.

b All patients enrolled in Cohort 2 who received open-label vedolizumab induction treatment.

c Duration of ulcerative colitis is defined as (1 + first dose date - diagnosis date) / 365.25.

d Baseline disease activity represents the baseline complete Mayo score.

Table 4 Categorization of Patients by Prior TNTα Antagonist Use and Worst Prior Treatment Failure, Induction Phase Safety Population – Study C13006

Medication Use/ Failure Failure Category	Induction Cohort 1 ITT Population [a] PLA N = 149	 VDZ N = 225	Induction Cohort 2 [b] Open-label VDZ N = 521	VDZ (Combined) N = 746	 Total N = 895
Prior TNFα antagonist use,[c] n (%)	73 (49)	95 (42)	263 (50)	358 (48)	431 (48)
No prior TNFα antagonist use, n (%)	76 (51)	130 (58)	258 (50)	388 (52)	464 (52)
Any prior TNFα antagonist failure,[d] n (%)	63 (42)	82 (36)	222 (43)	304 (41)	367 (41)
Inadequate response[e]	29 (46)	44 (54)	103 (46)	147 (48)	176 (48)
Loss of response[f]	26 (41)	32 (39)	83 (37)	115 (38)	141 (38)
Intolerance[g]	8 (13)	6 (7)	36 (16)	42 (14)	50 (14)
Prior immunomodulator failure but no TNFα antagonist failure, n (%)	55 (38)	96 (44)	209 (41)	305 (42)	360 (41)

Table 4 Categorization of Patients by Prior TNTα Antagonist Use and Worst Prior Treatment Failure, Induction Phase Safety Population (continued) – Study C13006

Medication Use/ Failure Failure Category	Induction Cohort 1 ITT Population [a] PLA N = 149	 VDZ N = 225	Induction Cohort 2 [b] Open-label VDZ N = 521	VDZ (Combined) N = 746	 Total N = 895
Inadequate response[e]	40 (73)	68 (71)	144 (69)	212 (70)	252 (70)
Intolerance[g]	15 (27)	28 (29)	65 (31)	93 (30)	108 (30)
Prior corticosteroid failure only, n %	25 (17)	42 (19)	78 (15)	120 (16)	145 (17)
Inadequate response[e]	23 (92)	36 (86)	68 (87)	104 (87)	127 (88)
Intolerance[g]	2 (8)	6 (14)	10 (13)	16 (13)	18 (12)

Source: Table 14.1.1.6CP, Table 14.1.1.12BP.

Abbreviations: IVRS = interactive voice response system; PLA = placebo; TNFα = tumor necrosis factor alpha; VDZ = vedolizumab.

Each patient is counted in only 1 medication class with their worst outcome being counted according to the following hierarchy: Inadequate response is considered worse than lost response; lost response is considered worse than intolerance.

a All patients enrolled in Cohort 1 who were randomized to blinded induction treatment with vedolizumab or placebo.

b All patients enrolled in Cohort 2 who received open-label vedolizumab induction treatment.

c Data for prior TNFα antagonist use were obtained from the IVRS, refers to patients with prior use, regardless of prior response.

d Data for prior treatment failure, including category of failure, were obtained from the UCRX case report form. Percent of patients with prior treatment failure to each medication class is based on number of patients in each treatment group. Percent of patients with category of failure is based on total number of patients with failure to that class.

e Inadequate response to TNFα antagonists is defined as persistently active disease despite induction treatment (as listed in the Study Definitions, Section 4.2) with specified medications. For immunomodulators and corticosteroids, inadequate response includes patients who had an inadequate response, lost response (immunomodulators only), or who were being treated with these agents at time of study entry and had active disease.

f Loss of response to TNFα antagonists is defined as recurrence of symptoms during maintenance dosing following prior clinical benefit.

g Intolerance is defined as occurrence of treatment-related toxicities (as defined by the protocol).

Table 5 Ulcerative Colitis Therapy Use at Baseline – Induction Phase Safety Population – Study C13006

Parameter	Induction Cohort 1 ITT Population[a] PLA N = 149	Induction Cohort 1 ITT Population[a] VDZ N = 225	Induction Cohort 2[b] Open-label VDZ N = 521	VDZ Combined N = 746	Total N = 895
Corticosteroids[c], n (%)	84 (56)	126 (56)	271 (52)	397 (53)	481 (54)
Immunomodulators[d], n (%)	44 (30)	75 (33)	189 (36)	264 (35)	308 (34)
Corticosteroids only, n (%)					
Yes	58 (39)	79 (35)	195 (37)	274 (37)	332 (37)
No	91 (61)	146 (65)	326 (63)	472 (63)	563 (63)
Corticosteroids and immunomodulators, n (%)					
Yes	26 (17)	47 (21)	76 (15)	123 (16)	149 (17)
No	123 (83)	178 (79)	445 (85)	623 (84)	746 (83)
Immunomodulators only, n (%)					
Yes	18 (12)	28 (12)	113 (22)	141 (19)	159 (18)
No	131 (88)	197 (88)	408 (78)	605 (81)	736 (82)
No corticosteroids or immunomodulators, n (%)					
Yes	47 (32)	71 (32)	137 (26)	208 (28)	255 (28)
No	102 (68)	154 (68)	384 (74)	538 (72)	640 (72)

Source: Table 14.1.1.6CP.

Abbreviations: IVRS = interactive voice response system; PLA = placebo; VDZ = vedolizumab.

Baseline UC medication use data were obtained from the IVRS for purposes of randomization stratification.

a All patients enrolled in Cohort 1 who were randomized to blinded induction treatment with vedolizumab or placebo.

b All patients enrolled in Cohort 2 who received open-label vedolizumab induction treatment.

c Corticosteroid use, with and without immunomodulator use

d Immunomodulator use, with and without corticosteroid use

Table 6 Comparison by Treatment Arm of Selected Baseline Ulcerative Colitis Disease Characteristics – Induction Study ITT Population – Study C13006

Disease Characteristic	PLA N = 149	VDZ N = 225	P value[b]
Duration of ulcerative colitis (years) [a]			
Mean (std dev)	7.1 (7.25)	6.1 (5.08)	0.8432
Baseline disease activity [c]			
Mean (std dev)	8.6 (1.68)	8.5 (1.78)	0.7276
Corticosteroid use at baseline, n (%)			
Yes	84 (56)	126 (56)	0.9428
No	65 (44)	99 (44)	
Immunomodulator use at baseline, n (%)			
Yes	44 (30)	75 (33)	0.4395
No	105 (70)	150 (67)	
Prior TNFα antagonist use[d], n (%)			
Yes	73 (49)	95 (42)	0.1975
No	76 (51)	130 (58)	
Prior TNFα antagonist failure[e], n (%)			
Yes	63 (42)	82 (36)	0.2567
No	86 (58)	143 (64)	

Source: Table 14.1.1.6B.

Abbreviations: ITT = intent-to-treat; IVRS = interactive voice response system; PLA = placebo; TNFα = tumor necrosis factor alpha; VDZ = vedolizumab.

a Duration of ulcerative colitis is defined as (1 + first dose date - diagnosis date)/ 365.25.

b P values for categorical variables are from chi-square test and for continuous variables are from Kruskal Wallis test.

c Baseline disease activity represents the baseline complete Mayo score.

d Data for prior TNFα antagonist use were obtained from the IVRS.

e Data for prior TNFα antagonist failure were obtained from the UCRX case report form.

Table 7 Clinical Response at Week 6 – Per Protocol Population –Study C13006

Table 14.3.1.2C
Clinical Response at Week 6
Per Protocol Population

Clinical Response [a]	PLA N=138	VDZ N=215
Number (%) Achieving Clinical Response	38 (27.5)	106 (49.3)
95% CI	(20.1, 35.0)	(42.6, 56.0)
Difference from Placebo [b]		21.8
95% CI for Difference from Placebo		(11.4, 32.3)
P-value for Difference from Placebo [c]		<0.0001
Relative Risk [d]		1.8
95% CI for Relative Risk		(1.3, 2.4)

(a) Clinical response is defined as reduction in complete Mayo score of >= 3 points and >=30% from baseline with an accompanying decrease in rectal bleeding subscore of >=1 point or absolute rectal bleeding subscore of <= 1 point
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI.
(c) P-values are based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no)
(d) Adjusted Relative Risk and its 95% CI.

Table 8 Clinical Response at Week 6 – Based on Patients Who Had Baseline and Week 6 (Observed Case) – Study C13006

Table 14.3.1.2B
Clinical Response at Week 6 - Based on Patients Who Had Baseline and Week 6 (Observed Case)

Clinical Response[a]	PLA N=137	VDZ N=216
Number (%) Achieving Clinical Response	38 (27.7)	106 (49.1)
95% CI	(20.2, 35.2)	(42.4, 55.7)
Difference from Placebo [b]		21.5
95% CI for Difference from Placebo		(11.0, 32.0)
P-value for Difference from Placebo [c]		<0.0001
Relative Risk [d]		1.8
95% CI for Relative Risk		(1.3, 2.4)

(a) Clinical response is defined as reduction in complete Mayo score of >= 3 points and >=30% from baseline with an accompanying decrease in rectal bleeding subscore of >=1 point or absolute rectal bleeding subscore of <= 1 point
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI.
(c) P-values are based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no)
(d) Adjusted Relative Risk and its 95% CI.

Table 9 Clinical Remission at Week 6 – Based on Patients Who Had Baseline and Week 6 Mayo Score (Observed Case) –Study C13006

Table 39.12.1.1B
Clinical Remission at Week 6, Based on Patients Who Had Baseline and Week 6 Mayo Scores (Observed Case)
Intent-to-Treat Population

Clinical Remission [a]	PLA N=137	VDZ N=216
Number (%) Achieving Clinical Remission	8 (5.8)	38 (17.6)
95% CI	(1.9, 9.8)	(12.5, 22.7)
Difference from Placebo [b]		11.7
95% CI for Difference from Placebo		(4.5, 18.9)
P-value for Difference from Placebo [c]		0.0014
Relative Risk [d]		3.0
95% CI for Relative Risk		(1.4, 6.3)

(a) Clinical remission is defined as complete Mayo score of <= 2 points and no individual subscore > 1 point
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI.
(c) P-values are based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no)
(d) Adjusted Relative Risk and its 95% CI.

Table 10 Mucosal Healing at Week 6 – Based on Patients Who Had Baseline and Week 6 Endoscopic Score (Observed Case) – Study C13006

Table 39.12.1.1C
Mucosal Healing at Week 6, Based on Patients Who had Baseline and Week 6 Endoscopic Scores (Observed Case)
Intent-to-Treat Population

Mucosal Healing [a]	PLA N=138	VDZ N=216
Number (%) Achieving Mucosal Healing	37 (26.8)	92 (42.6)
95% CI	(19.4, 34.2)	(36.0, 49.2)
Difference from Placebo [b]		16.1
95% CI for Difference from Placebo		(6.0, 26.3)
P-value for Difference from Placebo [c]		0.0019
Relative Risk [d]		1.6
95% CI for Relative Risk		(1.2, 2.2)

(a) Mucosal healing is defined as Mayo Endoscopic subscore of <=1 point
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI.
(c) P-values are based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with stratification according to: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no)
(d) Adjusted Relative Risk and its 95% CI.

Table 11 Baseline Demographics – Maintenance Phase Safety Population – Study C13006

	Maintenance Study ITT [a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Non-ITT		Combined	
Parameter	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125	PLA [b] (from Week 0) N = 149	VDZ Q4W[c] (Week 6 nonresponders) N = 373	PLA N = 275	VDZ N = 620
Gender, n (%)							
Male	69 (55)	70 (57)	68 (54)	92 (62)	226 (61)	161 (59)	364 (59)
Female	57 (45)	52 (43)	57 (46)	57 (38)	147 (39)	114 (41)	256 (41)
Race, n (%)							
White	101 (80)	104 (85)	101 (81)	115 (77)	313 (84)	216 (79)	518 (84)
Black	2 (2)	4 (3)	1 (< 1)	2 (1)	3 (< 1)	4 (1)	8 (1)
Asian	20 (16)	13 (11)	21 (17)	32 (21)	49 (13)	52 (19)	83 (13)
Other	3 (2)	1 (< 1)	2 (2)	0	8 (2)	3 (1)	11 (2)
Ethnicity, n (%)							
Hispanic or Latino	12 (10)	9 (7)	4 (3)	5 (3)	16 (4)	17 (6)	29 (5)
Not Hispanic or Latino	111 (88)	112 (92)	119 (95)	140 (94)	350 (94)	251 (91)	581 (94)
Not reported	3 (2)	1 (< 1)	2 (2)	4 (3)	7 (2)	7 (3)	10 (2)
Age (years) [d]							
n	126	122	125	149	373	275	620
Mean (std dev)	40.3 (13.92)	41.0 (12.85)	38.6 (14.21)	41.2 (12.50)	40.3 (12.73)	40.8 (13.15)	40.1 (13.07)
Median	39.9	39.7	35.7	40.7	38.7	40.4	38.5
Minimum, maximum	18, 74	19, 78	19, 76	19, 76	19, 75	18, 76	19, 78
Age (years), n (%)							
< 35	54 (43)	44 (36)	57 (46)	53 (36)	145 (39)	107 (39)	246 (40)
≥ 35	72 (57)	78 (64)	68 (54)	96 (64)	228 (61)	168 (61)	374 (60)

Table 11 Baseline Demographics – Maintenance Phase Safety Population (continued) – Study C13006

Parameter	Maintenance Study ITT [a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Non-ITT		Combined	
	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125	PLA [b] (from Week 0) N = 149	VDZ Q4W [c] (Week 6 nonresponders) N = 373	PLA N = 275	VDZ N = 620
Age (years), n (%)							
< 65	121 (96)	117 (96)	119 (95)	142 (95)	363 (97)	263 (96)	599 (97)
≥ 65	5 (4)	5 (4)	6 (5)	7 (5)	10 (3)	12 (4)	21 (3)
Body weight (kg)							
n	126	122	125	148	373	274	620
Mean (std dev)	74.7 (20.42)	78.2 (18.76)	71.8 (16.71)	72.4 (17.65)	72.4 (18.48)	73.5 (18.97)	73.4 (18.32)
Median	74.8	77.6	69.6	71.1	70.0	71.9	71.0
Minimum, maximum	36, 160	43, 133	38, 119	40, 125	32, 174	36, 160	32, 174
BMI (kg/m^2)							
n	126	122	125	148	373	274	620
Mean (std dev)	25.8 (6.06)	26.8 (6.26)	24.5 (4.70)	24.6 (5.11)	24.7 (5.62)	25.2 (5.59)	25.1 (5.64)
Median	25.1	26.1	23.6	23.9	23.7	24.6	24.0
Minimum, maximum	15, 52	16, 48	16, 40	15, 44	14, 61	15, 52	14, 61
Geographic region, n (%)							
North America	36 (29)	49 (40)	37 (30)	63 (42)	145 (39)	99 (36)	231 (37)
Western/Northern Europe	20 (16)	23 (19)	25 (20)	22 (15)	84 (23)	42 (15)	132 (21)
Central Europe	26 (21)	20 (16)	25 (20)	11 (7)	37 (10)	37 (13)	82 (13)
Eastern Europe	10 (8)	12 (10)	11 (9)	13 (9)	30 (8)	23 (8)	53 (9)
Asia/Australia/Africa	34 (27)	18 (15)	27 (22)	40 (27)	77 (21)	74 (27)	122 (20)

Source: Table 14.1.1.5AM.
Baseline refers to Week 0.

Abbreviations: BMI = body mass index; ITT = intent-to-treat; PLA = placebo; Std Dev = standard deviation; Q4W = dosing every 4 weeks; Q8W = dosing every 8 weeks; VDZ = vedolizumab.

a Patients who were randomized into the ITT population during the Maintenance Phase.

b Patients who received placebo during the Induction Phase and continued to receive placebo during the Maintenance Phase.

c Patients who received vedolizumab in the Induction Phase but did not achieve clinical response at Week 6 and continued to receive vedolizumab every 4 weeks during the Maintenance Phase.

d Age is defined as (1 + first dose date - birth date) / 365.25.

Table 12 Comparison by Treatment Arm of Selected Baseline Demographics Characteristics – Maintenance Study ITT Population – Study C13006

Parameter	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125	P value[b]
Gender, n (%)				
Male	69 (55)	70 (57)	68 (54)	0.8767
Female	57 (45)	52 (43)	57 (46)	
Race, n (%)				
White	101 (80)	104 (85)	101 (81)	0.5263
Other	25 (20)	18 (15)	24 (19)	
Age (years)[a]				
Mean (std dev)	40.3 (13.92)	41.0 (12.85)	38.6 (14.21)	0.2156
Body weight (kg)				
Mean (std dev)	74.7 (20.42)	78.2 (18.76)	71.8 (16.71)	0.0293
Geographic region, n (%)				
North America	36 (29)	49 (40)	37 (30)	0.0994
Europe (Western/Northern, Central, and Eastern)	56 (44)	55 (45)	61 (49)	
Asia/Australia/Africa	34 (27)	18 (15)	27 (22)	

Source: Table 14.1.1.5BM.

Abbreviations: ITT = intent-to-treat; PLA = placebo; Std Dev = standard deviation; Q4W = dosing every 4 weeks; Q8W = dosing every 8 weeks; VDZ = vedolizumab.

Baseline refers to Week 0.

a Age is defined as (1 + first dose date – birth date) / 365.25.

b P values for categorical variables are from Chi-Square Test and for continuous variables are from Kruskal Wallis Test.

Table 13 Baseline Ulcerative Colitis Disease Characteristics – Maintenance Phase Safety Population – Study C13006

	Maintenance Study ITT [a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Non-ITT		Combined	
	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125	PLA[b] N = 149	VDZ Q4W[c] (Week 6 nonresponders) N = 373	PLA (from Week 0) N = 275	VDZ N = 620
Duration of ulcerative colitis (years) [d]							
n	126	122	125	149	371	275	618
Mean (std dev)	7.8 (6.88)	6.2 (4.76)	7.6 (7.02)	7.1 (7.25)	6.5 (6.05)	7.4 (7.08)	6.7 (6.04)
Median	5.4	5.4	5.0	4.5	4.6	4.8	5.0
Minimum, maximum	0.5, 29.7	0.7, 26.3	0.5, 37.5	0.5, 38.5	0.5, 35.4	0.5, 38.5	0.5, 37.5
Categorical duration of ulcerative colitis, n (%)							
< 1 year	4 (3)	8 (7)	10 (8)	13 (9)	29 (8)	17 (6)	47 (8)
≥ 1 to < 3 years	31 (25)	29 (24)	23 (18)	44 (30)	101 (27)	75 (27)	153 (25)
≥ 3 to < 7 years	36 (29)	44 (36)	41 (33)	39 (26)	119 (32)	75 (27)	204 (33)
≥ 7 years	55 (44)	41 (34)	51 (41)	53 (36)	122 (33)	108 (39)	214 (35)
Missing	0	0	0	0	2	0	2
Baseline disease activity, n [e]							
n	126	122	125	149	373	275	620
Mean (std dev)	8.4 (1.75)	8.4 (1.80)	8.3 (1.66)	8.6 (1.68)	8.7 (1.77)	8.5 (1.72)	8.6 (1.76)
Median	8.0	8.0	8.0	8.0	9.0	8.0	9.0
Minimum, maximum	5, 12	4, 12	5, 12	3, 12	3, 12	3, 12	3, 12
Baseline disease activity, n (%)							
Complete Mayo score < 6	6 (5)	5 (4)	1 (< 1)	5 (3)	8 (2)	11 (4)	14 (2)
Complete Mayo score of 6 to 8 (inclusive)	63 (50)	62 (51)	72 (58)	70 (47)	157 (42)	133 (48)	291 (47)
Complete Mayo score of 9 to 12 (inclusive)	57 (45)	55 (45)	52 (42)	74 (50)	208 (56)	131 (48)	315 (51)

Table 13 Baseline Ulcerative Colitis Disease Characteristics – Maintenance Phase Safety Population (continued) – Study C13006

	Maintenance Study ITT [a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Non-ITT		Combined	
	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125	PLA[b] N = 149	VDZ Q4W[c] (Week 6 nonresponders) N = 373	PLA (from Week 0) N = 275	VDZ N = 620
Baseline fecal calprotectin							
n	122	117	117	139	362	261	596
Mean (std dev)	2055.6 (2935.39)	1686.3 (2609.01)	1782.9 (2918.17)	2369.9 (3258.82)	1700.3 (2439.94)	2223.0 (3109.89)	1713.7 (2569.14)
Median	1070.9	863.7	793.0	1005.5	852.0	1012.4	844.5
Minimum, maximum	23.8, 20000.0	34.9, 18061.8	23.8, 20000.0	23.8, 16443.8	23.8, 20000.0	23.8, 20000.0	23.8, 20000.0
Categorical baseline fecal calprotectin							
≤ 250 µg/g	25 (20)	26 (21)	21 (17)	27 (18)	59 (16)	52 (19)	106 (17)
> 250 to ≤ 500 µg/g	13 (10)	18 (15)	19 (15)	20 (13)	52 (14)	33 (12)	89 (14)
> 500 µg/g	84 (67)	73 (60)	77 (62)	92 (62)	251 (67)	176 (64)	401 (65)
Missing	4	5	8	10	11	14	24
Disease localization, n (%)							
Proctosigmoiditis	9 (7)	18 (15)	14 (11)	22 (15)	53 (14)	31 (11)	85 (14)
Left sided colitis	53 (42)	51 (42)	45 (36)	59 (40)	131 (35)	112 (41)	227 (37)
Extensive colitis	17 (13)	14 (11)	14 (11)	18 (12)	46 (12)	35 (13)	74 (12)
Pancolitis	47 (37)	39 (32)	52 (42)	50 (34)	143 (38)	97 (35)	234 (38)
Smoking status, n (%)							
Current smoker	8 (6)	7 (6)	8 (6)	11 (7)	21 (6)	19 (7)	36 (6)
Nonsmoker	87 (69)	82 (67)	82 (66)	88 (59)	216 (58)	175 (64)	380 (61)
Former smoker	31 (25)	33 (27)	35 (28)	50 (34)	136 (36)	81 (29)	204 (33)
History of extraintestinal manifestations							
Yes	39 (31)	46 (38)	48 (38)	44 (30)	121 (32)	83 (30)	215 (35)
No	87 (69)	76 (62)	77 (62)	105 (70)	252 (68)	192 (70)	405 (65)

Source: Table 14.1.1.6AM

Baseline refers to Week 0.

Abbreviations: ITT = intent-to-treat; PLA = placebo; std dev = standard deviation; Q4W = dosing every 4 weeks; Q8W = dosing every 8 weeks; VDZ = vedolizumab.

a Patients who were randomized into the ITT population during the Maintenance Phase.

b Patients who received placebo during the Induction Phase and continued to receive placebo during the Maintenance Phase.

c Patients who received vedolizumab in the Induction Phase but did not achieve clinical response at Week 6 and continued to receive vedolizumab every 4 weeks during the Maintenance Phase.

d Duration of Ulcerative Colitis is defined as (1+first dose date – diagnosis date)/ 365.25.

e Baseline disease activity represents the baseline complete Mayo score.

Table 14 Ulcerative Colitis Prior Therapy History – Maintenance Safety Population – Study C13006

	Maintenance Study ITT[a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Non-ITT		Combined	
	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125	PLA[b] N = 149	VDZ Q4W[c] (Week 6 nonresponders) N = 373	PLA (from Week 0) N = 275	VDZ N = 620
Prior TNFα antagonist use,[d] n (%)	47 (37)	50 (41)	52 (42)	73 (49)	209 (56)	120 (44)	311 (50)
No prior TNFα antagonist use,[d] n (%)	79 (63)	72 (59)	73 (58)	76 (51)	164 (44)	155 (56)	309 (50)
Any prior TNFα antagonist failure,[e] n (%)	38 (30)	43 (36)	40 (32)	63 (44)	183 (51)	101 (38)	266 (44)
Inadequate response [f]	19 (50)	16 (37)	17 (43)	29 (46)	95 (52)	48 (48)	128 (48)
Loss of response [g]	13 (34)	16 (37)	15 (38)	26 (41)	71 (39)	39 (39)	102 (38)
Intolerance [h]	6 (16)	11 (26)	8 (20)	8 (13)	17 (9)	14 (14)	36 (14)
Prior immunomodulator failure but no TNFα antagonist failure, [e] n (%)	61 (49)	56 (47)	60 (48)	55 (38)	128 (35)	116 (43)	244 (40)
Inadequate response[f]	43 (70)	43 (77)	40 (67)	40 (73)	86 (67)	83 (72)	169 (69)
Intolerance[h]	18 (30)	13 (23)	20 (33)	15 (27)	42 (33)	33 (28)	75 (31)
Prior corticosteroid failure only, [e] n %	26 (21)	19 (16)	25 (20)	25 (17)	50 (14)	51 (19)	94 (16)
Inadequate response [f]	21 (81)	19 (100)	20 (80)	23 (92)	44 (88)	44 (86)	83 (88)
Intolerance[h]	5 (19)	0	5 (20)	2 (8)	6 (12)	7 (14)	11 (12)

Sources: Table 14.1.1.6AM, Table 14.1.1.12AM.

Abbreviations: ITT = intent-to-treat; IVRS = interactive voice response system; PLA = placebo; std dev = standard deviation; TNFα = tumor necrosis factor alpha; Q4W = dosing every 4 weeks; Q8W = dosing every 8 weeks; VDZ = vedolizumab.

Each patient is counted in only 1 medication class with their worst outcome being counted according to the following hierarchy: Inadequate response is considered worse than lost response; lost response is considered worse than intolerance.

a Patients who were randomized into the ITT population during the Maintenance Phase.

b Patients who received placebo during the Induction Phase and continued to receive placebo during the Maintenance Phase.

c Patients who received vedolizumab in the Induction Phase but did not achieve clinical response at Week 6 and continued to receive vedolizumab every 4 weeks during the Maintenance Phase.

d Data for prior TNFα antagonist use were obtained from the IVRS, refers to patients with prior use, regardless of prior response.

e Data for prior treatment failure, including category of failure, were obtained from the UCRX case report form. Percent of patients with prior treatment failure to each medication class is based on number of patients in each treatment group. Percent of patients with category of failure is based on total number of patients with failure to that class.

f Inadequate response to TNFα antagonists is defined as persistently active disease despite induction treatment (as defined in the Study Definitions, Section 4.2) with specified medications . For immunomodulators and corticosteroids, inadequate response includes patients who had an inadequate response, lost response (immunomodulators only), or who were being treated with these agents at time of study entry and had active disease.

g Loss of response to TNFα antagonists is defined as recurrence of symptoms during maintenance dosing following prior clinical benefit.

h Intolerance is defined as occurrence of treatment-related toxicities (as defined by the protocol).

Table 15 Ulcerative Colitis Therapy Use at Baseline – Maintenance Phase Safety Population – Study C13006

	Maintenance Study ITT[a] (Responders to VDZ induction, randomized to Maint. Tmt. at Week 6)			Non-ITT		Combined	
	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125	PLA[b] N = 149	VDZ Q4W[c] (Week 6 nonresponders) N = 373	PLA (from Week 0) N = 275	VDZ N = 620
Corticosteroids, n (%)	72 (57)	70 (57)	73 (58)	84 (56)	182 (49)	156 (57)	325 (52)
Immunomodulators, n (%)	51 (40)	43 (35)	45 (36)	44 (30)	125 (34)	95 (35)	213 (34)
Corticosteroids only, n (%)							
Yes	48 (38)	48 (39)	48 (38)	58 (39)	130 (35)	106 (39)	226 (36)
No	78 (62)	74 (61)	77 (62)	91 (61)	243 (65)	169 (61)	394 (64)
Corticosteroids and immunomodulators, n (%)							
Yes	24 (19)	22 (18)	25 (20)	26 (17)	52 (14)	50 (18)	99 (16)
No	102 (81)	100 (82)	100 (80)	123 (83)	321 (86)	225 (82)	521 (84)
Immunomodulators only, n (%)							
Yes	27 (21)	21 (17)	20 (16)	18 (12)	73 (20)	45 (16)	114 (18)
No	99 (79)	101 (83)	105 (84)	131 (88)	300 (80)	230 (84)	506 (82)
No corticosteroids or immunomodulators, n (%)							
Yes	27 (21)	31 (25)	32 (26)	47 (32)	118 (32)	74 (27)	181 (29)
No	99 (79)	91 (75)	93 (74)	102 (68)	255 (68)	201 (73)	439 (71)

Sources: Table 14.1.1.6AM, Table 14.1.1.6BM.

Baseline refers to Week 0.

Abbreviations: ITT = intent-to-treat; PLA = placebo; std dev = standard deviation; Q4W = dosing every 4 weeks; Q8W = dosing every 8 weeks; VDZ = vedolizumab.

a Patients who were randomized into the ITT population during the Maintenance Phase.

b Patients who received placebo during the Induction Phase and continued to receive placebo during the Maintenance Phase.

c Patients who received vedolizumab in the Induction Phase but did not achieve clinical response at Week 6 and continued to receive vedolizumab every 4 weeks during the Maintenance Phase.

Table 16 Comparison by Treatment Arm of Selected Baseline Ulcerative Colitis Disease Characteristics – Induction Study ITT Population – Study C13006

Disease Characteristic	PLA N = 126	VDZ Q8W N = 122	VDZ Q4W N = 125	P value[b]
Duration of ulcerative colitis (years) [a]				
Mean (std dev)	7.8 (6.88)	6.2 (4.76)	7.6 (7.02)	0.5242
Baseline disease activity [c]				
Mean (std dev)	8.4 (1.75)	8.4 (1.80)	8.3 (1.66)	0.7635
Corticosteroid use at randomization, n (%)				
Yes	72 (57)	70 (57)	73 (58)	0.9774
No	54 (43)	52 (43)	52 (42)	
Immunomodulator use at randomization, n (%)				
Yes	51 (40)	43 (35)	45 (36)	0.6524
No	75 (60)	79 (65)	80 (64)	
Prior TNFα antagonist use, n (%)				
Yes	47 (37)	50 (41)	52 (42)	0.7540
No	79 (63)	72 (59)	73 (58)	
Prior TNFα antagonist failure, n (%)				
Yes	38 (30)	43 (35)	40 (32)	0.6878
No	88 (70)	79 (65)	85 (68)	

Source: Table 14.1.1.6BM.

Abbreviations: ITT = intent-to-treat; PLA = placebo; Std Dev = standard deviation; TNFα = tumor necrosis factor alpha; Q4W = dosing every 4 weeks; Q8W = dosing every 8 weeks; VDZ = vedolizumab.

a Duration of ulcerative colitis is defined as (1 + first dose date – diagnosis date) / 365.25.

b P values for categorical variables are from Chi-Square Test and for continuous variables are from Kruskal Wallis Test.

c Baseline (Week 0) disease activity represents the baseline (Week 0) complete Mayo score.

Table 17 Clinical Remission at Week 52 – Per Protocol Population – Study C13006

Table 14.3.1.2CM
Clinical Remission at Week 52
Per Protocol Population

Clinical Remission[a]	PLA N=121	VDZ q8 wks N=117	VDZ q4 wks N=121
Number (%) Achieving Clinical Remission	20 (16.5)	49 (41.9)	55 (45.5)
95% CI	(9.9, 23.1)	(32.9, 50.8)	(36.6, 54.3)
Difference from Placebo [b]		25.7	29.1
95% CI for Difference from Placebo		(14.2, 37.1)	(17.6, 40.7)
P-value for Difference from Placebo [c]		<0.0001	<0.0001
Relative Risk [d]		2.6	2.8
95% CI for Relative Risk		(1.6, 4.0)	(1.8, 4.3)

(a) Clinical remission is defined as complete Mayo score of <= 2 points and no individual subscore > 1 point at Week 52
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI.
(c) P-values are based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with 3 stratification factors: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no); 3) enrollment in cohort 1 or cohort 2 in the induction phase
(d) Adjusted Relative Risk and its 95% CI.

Table 18 Clinical Remission at Week 52 – Based on Patients Who Had Baseline and Week 52 (Observed Case) - Study C 13006

Table 14.3.1.2BM
Clinical Remission at Week 52 - Based on Patients Who Had Baseline and Week 52 Visit (Observed Case)

Clinical Remission[a]	PLA N=48	VDZ q8 wks N=77	VDZ q4 wks N=83
Number (%) Achieving Clinical Remission	20 (41.7)	51 (66.2)	56 (67.5)
95% CI	(27.7, 55.6)	(55.7, 76.8)	(57.4, 77.5)
Difference from Placebo [b]		24.1	25.0
95% CI for Difference from Placebo		(6.3, 41.9)	(7.9, 42.2)
P-value for Difference from Placebo [c]		0.0079	0.0042
Relative Risk [d]		1.6	1.6
95% CI for Relative Risk		(1.1, 2.3)	(1.1, 2.3)

(a) Clinical remission is defined as complete Mayo score of <= 2 points and no individual subscore > 1 point at Week 52
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI.
(c) P-values are based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with 3 stratification factors: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no); 3) enrollment in cohort 1 or cohort 2 in the induction phase
(d) Adjusted Relative Risk and its 95% CI.

Table 19 Durable Clinical Response at Week 52 – Without Imputation –Study C13006

Table 39.12.1.1E
Durable Clinical Response at Week 52 – Without Imputation
Intent-to-Treat Population

Durable Clinical Response[a]	PLA N=48	VDZ q8 wks N=77	VDZ q4 wks N=82
Number (%) Achieving Durable Clinical Response	30 (62.5)	69 (89.6)	65 (79.3)
95% CI	(48.8, 76.2)	(82.8, 96.4)	(70.5, 88.0)
Difference from Placebo [b]		28.4	18.6
95% CI for Difference from Placebo		(13.3, 43.5)	(2.9, 34.3)
P-value for Difference from Placebo [c]		0.0002	0.0203
Relative Risk [d]		1.4	1.3
95% CI for Relative Risk		(1.1, 1.8)	(1.0, 1.7)

(a) Durable clinical response is defined as reduction in complete Mayo score of >= 3 points and >= 30% from baseline with an accompanying decrease in rectal bleeding subscore of >= 1 point or absolute rectal bleeding subscore of <= 1 point at both Weeks 6 and 52
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI.
(c) P-values are based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with 3 stratification factors: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no); 3) enrollment in cohort 1 or cohort 2 in the induction phase
(d) Adjusted Relative Risk and its 95% CI.

Table 20 Mucosal Healing at Week 52 – Based on Patients Who Had Baseline and Week 52 Endoscopic Score (Observed Case) – Study C13006

Table 39.12.1.1F
Mucosal Healing at Week 52, Based on Patients Who had Baseline and Week 52 Endoscopic Scores (Observed Case)
Intent-to-Treat Population

Mucosal Healing[a]	PLA N=48	VDZ q8 wks N=77	VDZ q4 wks N=83
Number (%) Achieving Mucosal Healing	25 (52.1)	63 (81.8)	70 (84.3)
95% CI	(38.0, 66.2)	(73.2, 90.4)	(76.5, 92.2)
Difference from Placebo [b]		29.9	31.8
95% CI for Difference from Placebo		(13.4, 46.3)	(16.0, 47.7)
P-value for Difference from Placebo [c]		0.0004	<0.0001
Relative Risk [d]		1.6	1.6
95% CI for Relative Risk		(1.2, 2.1)	(1.2, 2.1)

(a) Mucosal Healing is defined as Mayo Endoscopic subscore of <= 1
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI.
(c) P-values are based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with 3 stratification factors: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no); 3) enrollment in cohort 1 or cohort 2 in the induction phase
(d) Adjusted Relative Risk and its 95% CI.

Table 21 Durable Clinical Remission at Week 52 – Without Imputation –Study C13006

Table 39.12.1.1G
Durable Clinical Remission at Week 52 – Without Imputation
Intent-to-Treat Population

Durable Clinical Remission[a]	PLA N=48	VDZ q8 wks N=77	VDZ q4 wks N=82
Number (%) Achieving Durable Clinical Remission	11 (22.9)	25 (32.5)	30 (36.6)
95% CI	(11.0, 34.8)	(22.0, 42.9)	(26.2, 47.0)
Difference from Placebo [b]		9.9	14.0
95% CI for Difference from Placebo		(-6.4, 26.1)	(-2.5, 30.5)
P-value for Difference from Placebo [c]		0.2336	0.0959
Relative Risk [d]		1.4	1.6
95% CI for Relative Risk		(0.8, 2.6)	(0.9, 2.9)

(a) Durable clinical remission is defined as complete Mayo score of <= 2 points and no individual subscore > 1 point at both Weeks 6 and 52
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI.
(c) P-values are based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with 3 stratification factors: 1) concomitant use of oral corticosteroids (yes/no); 2) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no); 3) enrollment in cohort 1 or cohort 2 in the induction phase
(d) Adjusted Relative Risk and its 95% CI.

Table 22 Corticosteroid Free Remission at Week 52 Based on Patients who had Baseline and Week 52 Mayo Scores (Observed) – Study C13006

Table 39.12.1.1H
Corticosteroid Free Clinical Remission at Week 52,Based on Patients Who Had Baseline and Week 52 Mayo Scores (Observed Case)
Intent-to-Treat Population

Corticosteroid Free Clinical Remission[a]	PLA N=24	VDZ q8 wks N=40	VDZ q4 wks N=46
Number (%) Achieving Corticosteroid Free Clinical Remission	10 (41.7)	22 (55.0)	33 (71.7)
95% CI	(21.9, 61.4)	(39.6, 70.4)	(58.7, 84.8)
Difference from Placebo [b]		11.2	30.2
95% CI for Difference from Placebo		(-13.6, 36.0)	(6.0, 54.3)
P-value for Difference from Placebo [c]		0.3755	0.0144
Relative Risk [d]		1.3	1.7
95% CI for Relative Risk		(0.7, 2.2)	(1.0, 2.8)

(a) Corticosteroid free clinical remission is defined as patients using oral corticosteroids at baseline who have discontinued corticosteroids and are in clinical remission at Week 52
(b) Difference and 95% CI: adjusted percent vedolizumab - adjusted percent placebo and its 95% CI.
(c) P-values are based on the Cochran-Mantel-Haenszel (CMH) chi-square test, with 2 stratification factors: 1) previous exposure to anti-TNF and/or concomitant immunomodulator use (yes/no); 2) enrollment in cohort 1 or cohort 2 in the induction phase
(d) Adjusted Relative Risk and its 95% CI.
N represents patients who were on corticosteroids at baseline

Table 23 Clinical Responses at Week 6 in Cohort 1 and Cohort 2 – Evaluation in – Subgroups Based on Anti-TNF Failure Status (Inadequate Response/Loss of Response) –Study C13006

Table 39.15.1.1
Clinical Response at Week 6 in Cohort 1 and Cohort 2 - Evaluation in Subgroups Based on Anti-TNF Failure Status (Inadequate Response/Loss of Response)
Safety Population

Clinical Response [a]	PLA N=149	VDZ Cohort 1 N=225	VDZ Cohort 2 N=521
Prior Anti-TNF Failure (Yes)	55	76	186
Number (%) Achieving Clinical Response	12 (21.8)	29 (38.2)	60 (32.3)
95% CI	(10.9, 32.7)	(27.2, 49.1)	(25.5, 39.0)
Prior Anti-TNF Failure (No)	94	149	335
Number (%) Achieving Clinical Response	26 (27.7)	77 (51.7)	171 (51.0)
95% CI	(18.6, 36.7)	(43.7, 59.7)	(45.7, 56.4)

(a) Clinical response is defined as reduction in complete Mayo score of >= 3 points and >=30% from baseline with an accompanying decrease in rectal bleeding subscore of >=1 point or absolute rectal bleeding subscore of <= 1 point
Confidence intervals for categorical data with numerators less than or equal to five are from the exact method, otherwise from the normal approximation

---

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

MILTON C FAN
05/15/2014

FREDA COONER
05/15/2014
See Statistical Team Leader Memorandum

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW

Vedolizumab interim results

BACKGROUND INFORMATION FOR ADVISORY COMMITTEE ON VEDOLIZUMAB

Statistical briefing material for the Joint Meeting of the Gastrointestinal Drugs Advisory Committee (GIDAC) and the Drug Safety and Risk Management Advisory Committee (DSaRM) Meeting, December 9, 2013.

John Stephen Yap, PhD
Clara Kim, PhD
Aloka Chakravarty, PhD

Division of Biometrics 7
Office of Biostatistics
Office of Translational Sciences
Center for Drug Evaluation and Research
U.S. Food and Drug Administration

Document Date: November 22, 2013

1 Introduction

This document presents the statistical safety perspective on the size of the safety database of BLA 125476 for vedolizumab. No cases of PML were identified in the vedolizumab safety database as of the June 28, 2013 cutoff date (safety update). In this document, we assess whether the vedolizumab safety database included sufficient number of patients with adequate exposure time to observe PML cases, given the rarity, and duration of exposure time needed to potentially develop PML. We estimate the worst case scenario for the PML incidence rate given no PML cases, and discuss a formal comparison of PML risks among vedolizumab and natalizumab patients.

2 Duration of Exposure and Estimating the Incidence Rate of PML

Although direct comparison of vedolizumab to natalizumab is infeasible, the total number of patients and exposure time of vedolizumab is compared to that of natalizumab when the first three PML cases on it were identified. In clinical trials of natalizumab, two PML cases were identified in 1,869 multiple sclerosis (MS) patients and one PML case in 1,043 Crohn's disease (CD) patients. The overall mean duration of exposure to natalizumab was approximately 18 months.

The vedolizumab safety database (as of the June 28, 2013 cutoff date) includes 3,326 patients exposed to at least one dose of vedolizumab. Among these patients, 2,830 (85%) patients filled out at least one subjective checklist as part of the RAMP program. The summary statistics (mean and median) for exposure data is shown in the following table (Table 1):

Table 1: Vedolizumab Patient Exposure (as of June 28, 2013)

		All Patients Exposed to Vedolizumab (n=3,326)	**All Patients Exposed to Vedolizumab w/ RAMP (n=2,830)**
No. of Infusions	Mean (SD)	16.8 (15.5)	19.5 (15.2)
	Median (Min-Max)	12.0 (1.0 – 65.0)	14.0 (1.0 – 65.0)
No. of Infusions with > 28 days FU	Mean (SD)	16.4 (15.2)	19.1 (15.0)
	Median (Min-Max)	11.0 (0.0 – 64.0)	13.0 (0.0 – 64.0)
No. of Months Exposure	Mean (SD)	14.9 (14.5)	17.4 (14.3)
	Median (Min-Max)	10.1 (0.0 – 65.0)	12.0 (0.0 – 65.0)

FU=follow-up; SD=standard deviation; Min=minimum value; Max=maximum value

The summary statistics indicates that the mean duration of exposure of all exposed patients (14.9 months) was shorter than the natalizumab mean exposure time of 18 months with a mean of 16.4 vedolizumab infusions. When limiting all exposed patients to those who have been assessed under the RAMP (2,830 patients), the mean exposure was 17.4 months with a mean of 19.1 vedolizumab infusions. Therefore, the size of the vedolizumab safety database and duration of

patient exposure is roughly similar to the natalizumab safety database when the first three PML cases were observed.

The point estimate of the incidence rate was 0 cases per 1,000 patients. The worst possible scenario, upper bound of the incidence rate can be calculated. The "rule of three" states that 3/n is an upper 95% confidence bound of the incidence rate when in a sample size of n, no cases occur[1]. For example, if no PML cases are observed in a study with 3,000 patients, then the true rate of PML will be lower than 3/3,000 (or 1 cases in 1,000 patients) with 95% confidence. Using the rule of three, the 95% upper bound of the PML incidence rate for vedolizumab was 0.9 and 1.1 in 1,000 patients, in all patients exposed and all patients with RAMP assessments, respectively.

3 Formal Comparison of PML Risk for Vedolizumab to Natalizumab

The applicant discussed a theoretical approach of evaluating the risk of PML in patients exposed to vedolizumab by using the estimated risk of PML among MS patients treated with natalizumab (Bloomgren, et al) [2] (Table 2). The applicant assumed that the PML risk for vedolizumab users was similar to that for natalizumab users, then used PML incidence rates for natalizumab on the current vedolizumab safety database to estimate the expected number of PML cases in the vedolizumab safety database.

Table 2: Natalizumab Estimated PML Incidence Stratified by Risk Factors

	Anti-JCV Antibody Positive	
Natalizumab Exposure	**No Prior IS Use**	**Prior IS Use**
1-24 months	0.56/1000	1.6/1000
25-48 months	4.6/1000	11.1/1000

JCV=John Cunningham virus; PML=progressive multifocal leukoencephalopathy; IS=immunosuppressant

Table 3-3 of the applicant's document titled, "Progressive Multifocal Leukoencepalopathy Risk Assessment for Vedolizumab" (page 5).

The applicant's approach is described as follows. First, all 3,326 patients exposed to at least one vedolizumab dose were stratified by the three known natalizumab PML risk factors: longer duration of treatment (beyond 24 months), prior immunosuppressant use, and positive anti-JCV antibody (Table 3). Approximately 80% of vedolizumab patients had prior immunosuppressant use. For anti-JCV antibody status, the applicant used published rates in the literature, and assumed that approximately 50% of patients to be JC virus antibody positive. Note in Table 3 the second row is a subset of the first row.

Second, Bloomgren's stratified PML rates in Table 2 were multiplied with the corresponding number of patients in Table 3 and the products were summed up to yield an expected number of PML cases of 6.75 for vedolizumab. Finally, the applicant assumed that PML occurrence among

the 3,326 vedolizumab-exposed patients followed a Poisson distribution with a mean of 6.75. Under this assumption, the probability of no PML cases was very low (~0.1%).

Table 3: Vedolizumab Patient Exposure Stratified by Natalizumab Risk Factors (as of June 28, 2013)

	Anti-JCV Antibody Positive			
Vedolizumab Exposure	**No Prior IS Use**	**Prior IS Use**	**Anti-JCV Ab Neg**	**Total**
1-24 months	333	1330	1663	3326
25-48 months	91	362	453	906

Ab Neg=antibody negative; JCV=John Cunningham virus; PML=progressive multifocal leukoencephalopathy; IS=immunosuppressant

Table 3-4 of the applicant's document titled, "Progressive Multifocal Leukoencepalopathy Risk Assessment for Vedolizumab" (page 6).

The applicant concluded that if the risk of PML among vedolizumab users were similar to natalizumab users, it would be almost certain that a PML case would occur. Because no PML cases were observed in the vedolizumab safety database, the PML risk is, therefore, lower for vedolizumab than for natalizumab.

The applicant's approach is appealing because the three risk factors for PML are all accounted for in the calculations. However, the comparison of the risk of PML for natalizumab and vedolizumab should be considered crude and interpreted with caution. Natalizumab and vedolizumab are treatments that are intended for distinctly different populations; Bloomgren et al.'s study was primarily on natalizumab exposed MS patients while vedolizumab is intended for UC and CD patients. Moreover, the MS patients in Bloomgren's article were from clinical trials, observation studies (including a Swedish registry), and also from spontaneous reports, while vedolizumab UC and CD patients were from clinical trials only. To compare the risk of PML between natalizumab and vedolizumab, one should conduct a study where enrolled patients are randomized to either natalizumab or vedolizumab. Also, it is not clear whether the three PML risk factors among natalizumab users are also risk factors for vedolizumab users and whether they interact amongst each other in the same manner as in natalizumab in modifying the PML incidence rate.

4 Statistics Summary and Conclusion

Assessment of the risk of PML in vedolizumab-treated patients is an important issue. In clinical trials of natalizumab, three PML cases were observed in around 3,000 patients with approximately 18 months of mean duration of exposure. The vedolizumab safety database includes roughly 3,000 patients with slightly less than 18 months average duration of exposure (and slightly less than 18 vedolizumab infusions). No PML cases have been identified in the current vedolizumab safety database. However, using the rule of three, one can estimate the true rate of PML for vedolizumab to be less than 0.9–1.1 cases per 1,000 patients, in all patients exposed and all patients with RAMP assessments, respectively. Although the vedolizumab and

natalizumab safety databases are somewhat comparable in exposure time and number of patients, and no PML cases for vedolizumab were identified, this does not imply that the PML risk for vedolizumab is lower than the PML risk for natalizumab.

From the clinical trials of natalizumab, with only 3 PML cases in 3,000 patients, it was not possible to determine PML risk factors. It was only after several years, when close to 100,000 patients were exposed to natalizumab for over 200,000 patient-years were observed that three risk factors were confirmed and consequently possible to estimate the stratified PML incidence rates. Thus, with no PML case in the safety database with roughly 3,000 patients treated with vedolizumab, it is unclear whether these same PML risk factors for natalizumab would be actual risk factors for vedolizumab and yield similar estimated stratified PML incidence rates. In order to assess PML risk for vedolizumab, a larger number of vedolizumab-treated patients will need to be studied. If the risk of PML for vedolizumab were to be directly compared with natalizumab, then both treatments will need to be included in the same study.

5 References

[1]Jovanovic, B.D. and Levy, P.S. A Look at the Rule of Three. *The American Statistician*, vol. 51, no. 2 (May 1997), pp. 137-139.

[2]Bloomgren G, Richman S, Hotermans C, Subramanyam M, Goelz S, Natarajan A, et al. Risk of natalizumab-associated progressive multifocal leukoencephalopathy. *N Engl J Med* 2012; 366(20):1870-80.

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

JOHN S YAP
11/22/2013

CLARA KIM
11/22/2013
I concur with this briefing document.

ALOKA G CHAKRAVARTY
11/22/2013

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	BLA 125476 (Ulcerative Colitis)/125507 (Crohn's Disease)
Drug Name:	Entyvio (vedolizumab) 300 mg of lyophilized in a single vial for intravenous infusion
Indication(s):	Treatment of moderate or severe ulcerative colitis and Crohn's disease in adults
Applicant:	Takeda Pharmaceuticals, Inc.
Date(s):	6/20/2013 (Receipt)
	2/18/2014 (PDUFA)
Review Priority:	Standard
Biometrics Division:	Division of Biometrics VII
Statistical Reviewer:	John Stephen Yap, PhD
Concurring Reviewers:	Clara Kim, PhD (Team Leader)
	Aloka Chakravarty, PhD (Division Director)
Medical Division:	Division of Gastroenterology and Inborn Errors Products (DGIEP)
Clinical Team:	Laurie Muldowney, MD
	Anil Rajpal, MD (Team Leader)
Project Manager:	Kevin Bugin, MS, RAC

Keywords: PML, integrin receptor antagonist, ulcerative colitis, Crohn's disease, rule of three

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

Takeda Pharmaceuticals, Inc. submitted a BLA for vedolizumab with proposed indications for ulcerative colitis (UC) and Crohn's disease (CD) on 6/20/2013. Vedolizumab belongs to the class of integrin receptor antagonists which includes Tysabri (natalizumab). Natalizumab is approved for the treatment of multiple sclerosis (MS) and CD, and has a black-box warning for the risk of progressive multifocal leucoencephalopathy (PML) — an opportunistic viral infection of the brain — that usually leads to death or severe disability. Because of the severity and potential risk of PML among vedolizumab patients, the FDA stated that the applicant's, "…safety database at the time of original BLA submission must include data on at least 900 patients that received ≥24 infusions (with a minimum of 4 weeks of follow-up after the last infusion)". This statistical safety review includes assessment of the risk of PML among vedolizumab patients.

No PML cases were confirmed in the vedolizumab safety database which contained roughly the same number of patients (~3,000 patients) and mean exposure duration (~18 months) as the natalizumab safety database when three PML cases were observed. Using the rule of three (Jovanovic and Levy, 1997), the upper bound of the 95% confidence interval of the PML incidence rate of vedolizumab was 0.9 in 1,000 patients. This risk estimate did not take duration of exposure into consideration and included all subjects who received at least one dose of vedolizumab. When limited to patients with at least 24 infusions and at least 4 weeks of follow up, the upper bound was 3.0 in 1,000 patients.

The applicant estimated the expected number of PML cases in the vedolizumab safety database using PML risk estimates among multiple sclerosis patients treated with natalizumab, stratified by PML risk factors. The expected number of PML cases was 6.8. Assuming a Poisson distribution and similar risk to natalizumab, the probability of observing zero PML cases in the vedolizumab safety database was approximately 0.1%.

The applicant concluded that the risk of PML is lower among vedolizumab patients compared to natalizumab patients because zero PML cases were identified and the probability of observing zero PML cases would be very low if the risk were similar to natalizumab.

This analysis, however, should be considered as crude and interpreted with caution because vedolizumab and natalizumab are intended for different patient populations and the patient data for vedolizumab were limited to clinical trials while those for natalizumab were from clinical trials, observational studies, and spontaneous reports (with a much larger number of patients and longer average durations of exposure).

To rule out the risk of PML among vedolizumab users, a larger number of patients with longer duration of exposure would need to be observed. If the risk of PML for vedolizumab were to be directly compared with natalizumab, both treatments would have to be included in the same study.

2 INTRODUCTION

2.1 Overview

Vedolizumab is a humanized immunoglobulin monoclonal antibody developed by Takeda Pharmaceuticals, Inc. The proposed indications are as follows:

- **Ulcerative colitis (UC)**: Reducing signs and symptoms, inducing and maintaining clinical response and remission, and mucosal healing, and achieving corticosteroid-free remission in adult patients with moderately to severely active ulcerative colitis who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy or a tumor necrosis factor-alpha (TNFα) antagonist.
- **Crohn's disease (CD)**: Reducing signs and symptoms, inducing and maintaining clinical response and remission, and achieving corticosteroid-free remission in adult patients with moderately to severely active Crohn's disease who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy or a tumor necrosis factor-alpha (TNFα) antagonist.

The proposed dosage and administration are 300 mg infused intravenously over approximately 30 minutes at 0, 2, and 6 weeks, and then every 8 weeks thereafter. The proposed dosage form and strength are 300 mg lyophilized vedolizumab in a single-use vial for intravenous infusion.

Vedolizumab belongs to the class of integrin receptor antagonists which includes Tysabri (natalizumab). Natalizumab is approved for the treatment of multiple sclerosis (MS) and Crohn's disease (CD) and has a black-box warning for the risk of progressive multifocal leucoencephalopathy (PML) – an opportunistic viral infection of the brain that usually leads to death or severe disability. PML is caused by reactivation of the John Cunningham (JC) virus. Increased risk of PML in natalizumab-treated patients has been shown to be associated with treatment duration (at least 24 months of exposure), prior immunosuppressant use, and the presence of anti-JC virus antibodies (Bloomgren et al., 2012).

On July 20, 2011, the FDA held a closed-session Joint Meeting of the Gastrointestinal Drugs Advisory Committee and the Drug Safety and Risk Management Advisory Committee (AC) for vedolizumab. Recommendations sought from the AC included the number of patients and duration of exposure needed to rule out the risk of PML associated with vedolizumab. The AC indicated that patient exposure should be at least 24 months, however, the AC did not specify the total number of patients that would be required in the vedolizumab safety database at the time of BLA submission.

In a previous meeting with the applicant (dated November 6, 2012), the FDA recommended that the safety database must include data on at least 900 patients that received ≥ 24 infusions (with a minimum of 4 weeks of follow-up after the last infusion) at the time of original BLA submission.

This review will assess all studies in the Integrated Summary of Safety (ISS) of the BLA, which includes all patients who received at least one dose of study drug. Table 1 lists all studies included in the ISS by the type of patients and study phase.

Table 1: Studies in ISS

	Study Phase		
Type of Patients	**1**	**2**	**3**
Healthy	L297-007 C13001 C13005 C13009 C13010 C13012		
UC	L297-005 L297-006 M200-021	M200-022 C13002	C13006
CD		L299-016	C13007 C13011
UC and CD		C13004	C13008

The pivotal studies for this application were studies C13006, C13007, and C13011. Studies C13006 and C13007 consisted of induction and maintenance phases while study C13011 consisted only of an induction phase. Figure 1 shows the design for study C13006. The design for study C13007 was similar.

For studies C13006 and C13007, in the induction phase, patients were initially randomized 3:2 to receive vedolizumab (300 mg dose) or placebo at Week 0 and Week 2 (cohort 1). After enrollment in cohort 1 was completed, patients were enrolled into cohort 2 and received vedolizumab at Week 0 and Week 2. Patients in cohorts 1 and 2 were assessed at Week 6. Patients who received vedolizumab at induction and achieved clinical response were randomized 1:1:1 to receive vedolizumab every 4 weeks, vedolizumab every 8 weeks, or placebo, for an additional 44 weeks. Patients who received vedolizumab at induction and did not achieve clinical response continued to receive vedolizumab every 4 weeks. Patients who received placebo at induction continued on placebo. Patients were stratified by concomitant use of oral corticosteroids and previous exposure to TNFα antagonists and concomitant immunomodulators during the induction and maintenance phases, and by cohorts during the maintenance phases.

Studies C13004 and C13008 were uncontrolled long-term safety studies. Patients who were enrolled in study C13004 were either *de novo* patients (or patients who have not previously received any study drug) or patients who were previously enrolled in study C13002. Patients who were enrolled in study C13008 were either *de novo* patients or patients who were previously enrolled in studies C13004, C13006, C13007, and C13011.

Figure 1: Study C13006 Design

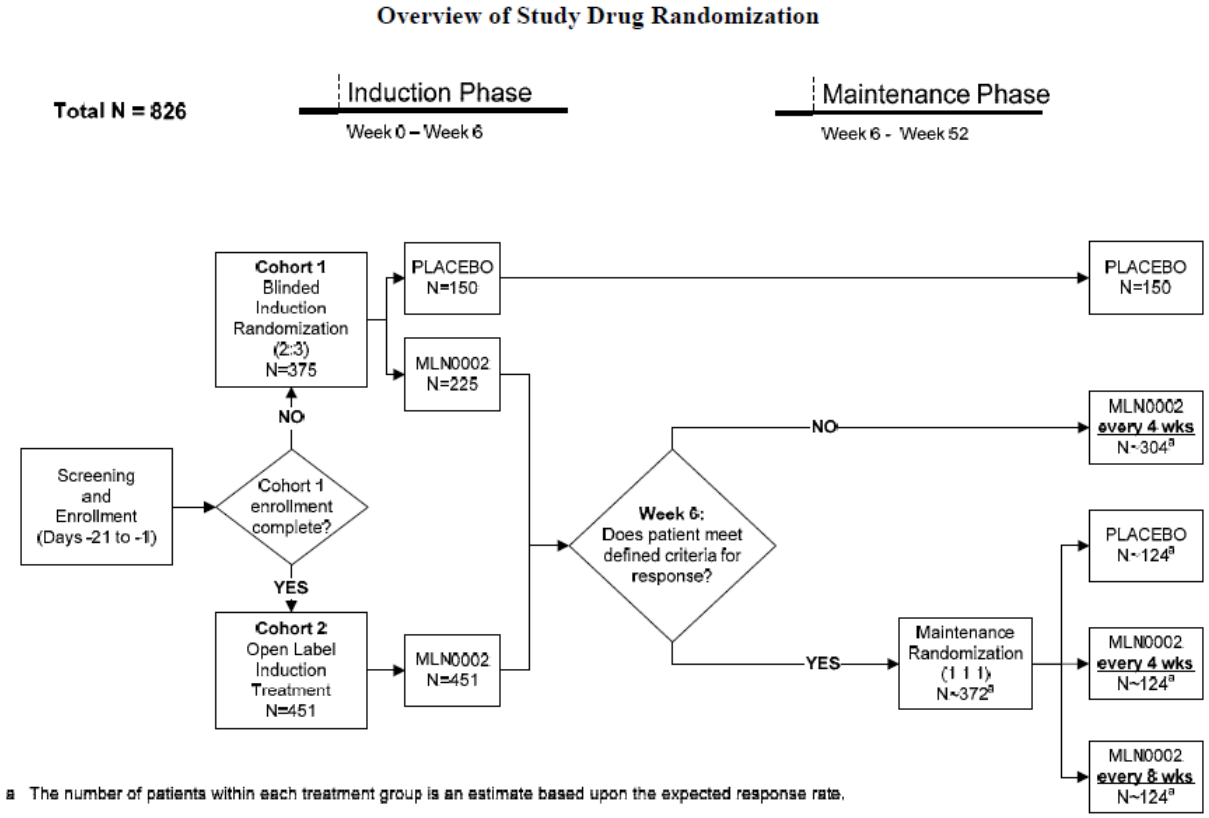


Figure titled, "Overview of Study Drug Randomization" from Study C13006 protocol, page 8

The applicant stated two studies (C13013 and CPH-001), and patients in Study C13004 who did not roll-over into Study C13008, were not included in the ISS.

2.2 Data Sources

The following applicant materials were reviewed:

1. Study Reports
   - Integrated Summary of Safety
   - ISS Statistical Analysis Plan (dated March 13, 2013)
   - Progressive Multifocal Leucoencephalopathy Risk Assessment for Vedolizumab

2. Data Sets:
   - ADSL (ISS analysis dataset for subject-level)
   - ADEXCMB (ISS analysis dataset for infusion summary)
   - ADSLCMB (ISS analysis dataset for VEDOLIZUMAB exposure)
   - ADCF (ISS analysis dataset for clinical findings)

The cut-off date for the datasets was March 14, 2013. The 120-day safety update cut-off date was June 28, 2013. No software codes were submitted by the applicant. All materials reviewed can be found in the following link:

\\cdsesub1\bla\ectd_submissions\stn125476\125476.enx

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

Analysis and Study Data Tabulation Model datasets were available. The applicant provided sufficient documentation for the datasets and also included a statistical analysis plan for the ISS analyses.

3.2 Evaluation of Efficacy

This review is focused entirely on safety.

3.3 Evaluation of Safety

3.3.1 Safety Analysis Population(s) and Endpoint(s)

The applicant defined the safety population as all patients who received any amount of study drug based on what they actually received.

Analyses were performed by study and across studies, and by indication (UC, CD, or combined) and by study design (induction or induction and maintenance).

The primary endpoint was the occurrence of PML. To assess PML, the applicant implemented a Risk Assessment and Minimization for PML (RAMP) program. PML assessment in the RAMP consisted of (1) a subjective checklist, a questionnaire inquiring patients about the presence of specific neurologic signs and symptoms of PML (administered prior to entering the study, at each visit prior study drug administration, and at the final safety visit), (2) an objective checklist filled out by the investigator, (3) an assessment by a site's study neurologist, and (4) assessment by an Independent Adjudication Committee, an external committee of academic experts (including neurologists, neuroradiologists, and virologists) who provided independent assessment. For more details about the RAMP, see the clinical review by Dr. Laurie Muldowney.

3.3.2 Data Quality

The reviewer did not identify any data quality issues.

3.3.3 Statistical Methods

The applicant's primary analyses consisted of summary tables for vedolizumab exposure of all patients in the safety population, and estimated the 95% confidence interval upper bound of the PML incidence risk. Additionally, the applicant estimated the expected number of PML cases in the vedolizumab safety database using PML risk estimates among multiple sclerosis patients treated with natalizumab, stratified by PML risk factors

3.3.4 Results and Conclusions

This section covers the applicant's and reviewer's analyses of the following topics: safety population, exposure, prior and concomitant immunosuppressant use, RAMP, and PML risk assessment. If not indicated as the reviewer's analysis results, the results discussed in this section were of the applicant's, and replicated by the reviewer. All exposure (by number of infusions and duration) and RAMP results include data as of June 28, 2013 (120-day safety update). All other results were from data as of March 14, 2013.

3.3.4.1. Safety Population

Table 2 summarizes the enrollment and formulation in each study included in the ISS.

Table 2: Summary of Studies Included in ISS

Study (Formulation)	Phase	Number Dosed	
		Placebo	Vedolizumab
Healthy Subjects			
L297-007 (A)	1	5	14
C13001 (B)	1	10	39
C13005 (B)	1	0	26
C13009 (B and C)	1	25	62
C13010 (C)	1	0	42
C13012 (C)	1	0	14
	Total	40	197
Ulcerative Colitis			
L297-005 (A)	1b/2a	5	9
L297-006 (A)	1b/2a	8	21
M200-021 (A)	1/2	6	24
M200-022 (A)	2	63	118
C13002 (B)	2	9	37
C13006 (C)	3	149	746
	Total	240	955
Crohn's Disease			
L299-016 (A)	2	58	127
C13007 (C)	3	148	967
C13011 (C)	3	207	209
	Total	413	1,303
Ulcerative Colitis and Crohn's Disease			
C13004 (B)	2	0	53/19*

C13008 (C)	3	0	
Rollover Patients			704/1118**
De Novo Patients			190/231**

* Of 72 enrolled patients, 53 had UC and 19 had CD.
**Of 1822 patients enrolled from previous studies, 704 had UC and 1118 had CD. Of 421 de novo patients, 190 had UC and 231 had CD.
A= (b) (4)
B= (b) (4)
C=powder for solution for infusion: a lyophilized formulation in 2 strengths, 180 mg/vial and 300 mg/vial, used for IV infusion

Table 1-1, page 25, of ISS

Reviewer's Comment: The total number of patients in Table 2, after excluding the long-term, uncontrolled studies C13004 and C13008 are summarized in the following table.

	Placebo	***Vedolizumab***	***Total***
Healthy Subjects	*40 (6)*	*197 (8)*	*237 (8)*
UC	*240 (35)*	*955 (39)*	*1,195 (38)*
CD	*413 (59)*	*1,303 (53)*	*1,716 (54)*
Total	*693*	*2,455*	*3,148*

The clinical development program included 3,326 patients who were exposed to at least one dose of vedolizumab. Among these were 197 healthy subjects and 3129 UC and CD patients (1279 with UC and 1850 with CD). The number of patients in phase 1 studies and patients in controlled phase 1b/2/3 studies are shown in Figure 2.

Reviewer's Comment: Table 2 is stratified by types of subjects/indication whereas Figure 1 is stratified by study phase.

There were 2 long-term, uncontrolled studies: C13004 (phase 2) and C13008 (phase 3). Study C13004 included 41 patients who had not previously received vedolizumab (7 roll-over patients from phase 2 study C13002 who were on placebo and 34 *de novo* patients). Study C13008 included 830 patients who had not previously received vedolizumab:

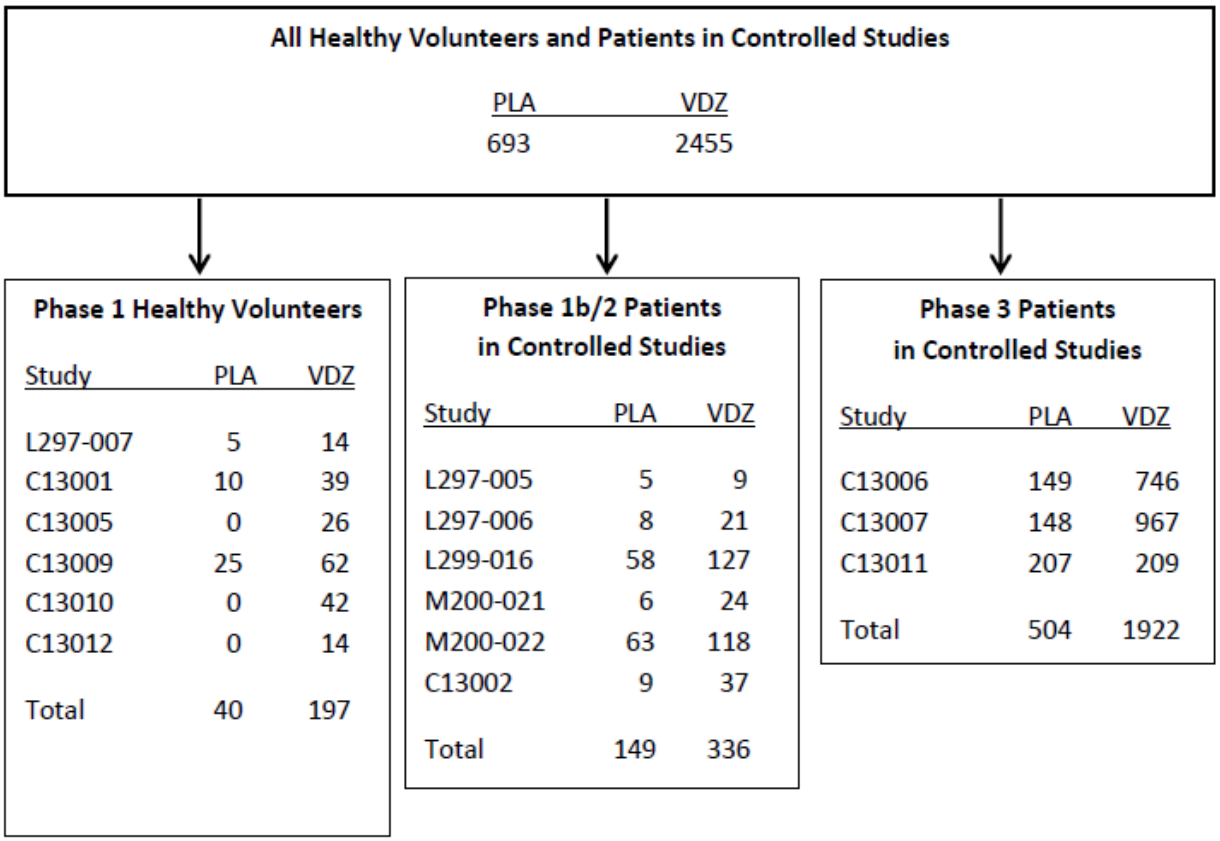
- 409 roll-over patients from studies C13006, C13007, and C13011 who were on placebo
- 421 *de novo* patients

Reviewer's Comments: From the applicant's analyses it can be inferred that

- ***Among the safety population, 277 patients were not exposed to VDZ (40 healthy volunteers, 121 UC patients, and 116 CD patients).***
- ***Three patients (Unique subject IDs 22-003-001, C13002-212-001, and C13007-28011-701) were not included in the safety population.***

The reviewer performed additional analysis to determine the numbers of patients who were rolled-over into the long-term, uncontrolled studies C13004 and C13008. The results of this analysis are presented in the appendix.

Figure 2: Number of Healthy Subjects in Phase 1 Studies and Patients in Controlled Phase 1b, 2, and 3 Studies



PLA=placebo; VDZ=vedolizumab
Figure 2-1, page 85 of ISS.

3.3.4.2. Exposure

Exposure to vedolizumab was summarized by number of doses, number of months, and by the number of infusions with and without at least 4 weeks of follow-up. In phase 1 and 2 studies, a dose was defined as the administration of any amount of vedolizumab. In phase 3 studies, a dose was defined as the administration of at least 75% of the infusion by volume. Exposure in days was converted to months by dividing the number of days by 30.4.

Patients who received only placebo throughout any of the studies had vedolizumab exposure calculated as 0 days. The total vedolizumab exposure for patients enrolled in multiple studies was calculated as the sum of all exposures in each study the patient was enrolled.

Tables 3 and 4 (calculated by the reviewer) summarize the durations of vedolizumab exposure in months and number of infusions, respectively, across all studies, as of June 28, 2013. In phase 2 and 3 trials combined, patients were exposed to vedolizumab for a mean of 480.6 days while in phase 3 trials combined, patients were exposed to vedolizumab for a mean of 532.0 days.

Table 3: Duration of Exposure to Vedolizumab in Months (as of June 28, 2013)

Duration	Healthy Subjects	UC	CD	Total
At least 1 dose	197	1279	1850	3326
Months of Exposure				
≥6	0	855	1167	2022
≥12	0	588	830	1418
≥18	0	485	677	1162
≥24	0	428	478	906
≥36	0	198	209	407
≥48	0	30	10	40

Table 4 shows that the BLA includes 1004 patients with at least 24 vedolizumab infusions and 532 patients with at least 36 infusions (with a minimum of 4 weeks follow-up after the last infusion).

Table 4: Duration of Exposure to Vedolizumab by Number of Infusions (as of June 28, 2013)

Duration	Healthy Subjects	UC	CD	Total
At least 1 dose	197	1279	1850	3326
Number of Infusions*				
≥1	193	1261	1826	3280
≥6	0	913	1283	2196
≥12	0	673	916	1589
≥18	0	498	730	1228
≥24	0	444	560	1004
≥36	0	254	278	532
≥48	0	63	53	116

*with 4-week follow-up

The reviewer analyzed exposure for all vedolizumab-exposed patients and a subset of these who filled out at least one subjective checklist as part of the RAMP (see section **3.3.4.4. RAMP** below). The mean and median number of infusions and duration of exposure are shown in Table 5.

The mean duration of exposure of all exposed patients was 14.9 months with a mean of 16.4 vedolizumab infusions. When limiting the analysis to those who have been assessed at least one subjective checklist under the RAMP (2,830 patients), the mean exposure was 17.4 months with a mean of 19.1 vedolizumab infusions. The number of infusions and duration of exposure were largely driven by the ongoing uncontrolled, long-term safety study C13008 (2,243 [67%] patients).

Reviewer's Comment: In clinical trials of natalizumab, two PML cases were identified in 1,869 MS patients and one PML case in 1,043 CD patients (3 PML cases in almost 3,000 patients). The overall mean duration of exposure to natalizumab was approximately 18 months. The reviewer's analysis shows that the number of patients and mean durations of

exposure between the vedolizumab safety database and the natalizumab safety database (when the 3 PML cases were observed) are similar.

Table 5: Vedolizumab Patient Exposure (as of June 28, 2013)

		All Patients Exposed to Vedolizumab (N =3,326)	All Patients Exposed to Vedolizumab w/ RAMP* (N =2,830)
No. of Infusions	Mean (SD)	16.8 (15.5)	19.5 (15.2)
	Median (Min-Max)	12.0 (1.0 – 65.0)	14.0 (1.0 – 65.0)
No. of Infusions with > 28 days FU	Mean (SD)	16.4 (15.2)	19.1 (15.0)
	Median (Min-Max)	11.0 (0.0 – 64.0)	13.0 (0.0 – 64.0)
No. of Months Exposure	Mean (SD)	14.9 (14.5)	17.4 (14.3)
	Median (Min-Max)	10.1 (0.0 – 65.0)	12.0 (0.0 – 65.0)

FU=follow-up; SD=standard deviation; Min=minimum value; Max=maximum value

*Assessed with at least one subjective checklist in RAMP

3.3.4.3. Prior and Concomitant Use of Immunosuppressants

Prior medication was defined as a medication that stopped prior to the calendar day of the first dose of study medication. Concomitant medication was defined as any non-study medication that was taken between the first and last days of the study (inclusive).

Table 6 summarizes the total numbers of patients who had prior and concomitant immunosuppressant use, stratified by number of infusions.

Table 6: Exposure to Vedolizumab by Number of Infusions and by Prior and Concomitant Immunosuppresant Use (as of March 14, 2013)

	Number of Vedolizumab Infusions[a,b]				
Category	≥ 6 N = 2136	≥ 12 N = 1436	≥ 18 N = 1136	≥ 24 N = 869	≥ 36 N = 385
Prior Immunosuppressant Use					
Yes	1735 (81)	1155 (80)	900 (79)	678 (78)	296 (77)
No	401 (19)	281 (20)	236 (21)	191 (22)	89 (23)
Concomitant Immunosuppressant Use[c]					
Yes	596 (28)	440 (31)	349 (31)	261 (30)	103 (27)
No	1540 (72)	996 (69)	787 (69)	608 (70)	282 (73)

Source: Table 18.1.1.2C.

a Includes phase 3 Studies C13006, C13007, C13011, and C13008; because Studies C13002 and C13004 are not included, the exposure numbers will not match the overall exposure numbers.

b Patients had a minimum of 4 weeks of follow-up after the last infusion.

c US patients are classified as no concomitant immunosuppressant use.

Table 2-7, page 87 of the ISS.

Approximately 80% of patients had prior immunosuppressant use and approximately 30% of patients have concomitant immunosuppressant use. When using the number of months instead of number of infusions, the results were similar.

Reviewer's Comments:

- ***The results did not differ when stratified by "≥ 1 infusions" in Table 6. Approximately 80% of patients had prior immunosuppressant use and approximately 30% of patients had concomitant immunosuppressant use.***
- ***A number of patients in the United States (around 197 or more) had concomitant immunosuppressant use. The precise number of these patients cannot be verified from the applicant's data.***

3.3.4.4. RAMP

Findings from the PML checklist were not summarized for the healthy subject safety population because of the limited exposure to vedolizumab (1 dose) and the lack of concomitant risk factors in this population.

Table 7: RAMP Results Stratified by Number of Infusions (as of June 28, 2013)

PML Checklist	Ulcerative Colitis (n=1146)	Crohn's Disease (n=1781)	Total* (n=2,927)
No. of Patients			
Subjective Checklist Administered	1142 (>99)	1771 (>99)	2,913 (>99)
Positive Subjective Finding	97 (8)	193 (11)	290 (10)
Objective Checklist Admin.	97 (8)	193 (11)	290 (10)
Abnormal Finding on Obj. Chk.	17 (1)	47 (3)	64 (2)
Ramp Algorithm			
Referred to a Neurologist	24 (2)	61 (3)	85 (3)
MRI performed	15 (1)	43 (2)	58 (2)
IAC involved	24 (2)	62 (3)	86 (3)
Lumbar puncture	2 (<1)	3 (<1)	5 (<1)
CS fluid analyzed by PCR for JCV	2 (<1)	3 (<1)	5 (<1)
JCV DNA detected by PCR in CSF	0	0	0
Diagnosed with PML by IAC	0	0	0

*Received at least 1 vedolizumab infusion; Note: (%) in 3rd and 4th columns based on n=2,927 in column 2; inf=infusions

Table 7 summarizes the results of the RAMP algorithm by proposed indication. These results were calculated by the reviewer from data as of March 14, 2013 and applicant's updated RAMP tables as of June 28, 2013. No actual RAMP data as of June 28, 2013 were provided by the applicant. From among 2,927 patients included in the RAMP, at least 2,913 patients had been administered the subjective checklist. Among these patients, 290 had positive subjective findings and were also administered objective checklists. There were 85 patients who were referred to neurologists, 58 had MRI performed, 86 cases involved the Independent Adjudication Committee (IAC), 5 patients had lumbar punctures, and 5 patients had cerebrospinal fluid analyzed by PCR for JC virus. ***There were no PML cases that were confirmed by the IAC.*** For

a more detailed discussion of the RAMP results, see the clinical review by Dr. Laurie Muldowney.

The reviewer performed two additional analyses on the RAMP results. First, to determine whether assessments of PML symptoms were observed mostly in patients with ≥24 infusions, the RAMP results were stratified by the number of infusions. Secondly, a cross table of the number of subjective checklist administered versus number of infusions was created. The resulting cross table shows whether the subjective checklists were consistently administered to patients at each scheduled vedolizumab infusion session.

The results of the first analysis (based on data as of March 14, 2013) are in Table 8. Numbers in parentheses in columns 3 and 4 are percentages based on the total number of RAMP-assessed patients in column 2 (i.e. 2,927). Among patients who received at least 12 and 24 infusions, 172 (61%) and 113 (40%) patients, respectively, had positive subjective finding. This shows that around 40% of patients with positive objective finding had ≥24 infusions and 60% had <24 infusions. Among patients who received at least 12 and 24 infusions, 37 (60%) and 26 (42%) patients, respectively, had abnormal finding on objective checklist. In general, for each row in Table 8, the percentages in column three (around 50%-60%) were consistently much higher than those in column four (around 20%-40%). These results suggest that assessments of PML symptoms were not observed mostly in patients with ≥24 infusions.

Table 8: RAMP Results Stratified by Infusions (as of March 14, 2013)

PML Checklist	Total* (n=2,927)	No. of Patients with ≥12 inf	No. of Patients with ≥24 inf
No. of Patients			
Subjective Checklist Administered	2,913	1,488 (51)	903 (31)
Positive Subjective Finding	284	172 (61)	113 (40)
Objective Checklist Admin.	284	172 (61)	113 (40)
Abnormal Finding on Obj. Chk.	62	37 (60)	26 (42)
Ramp Algorithm			
Referred to a Neurologist	82	50 (61)	33 (40)
MRI performed	56	30 (54)	20 (36)
IAC involved	83	52 (63)	35 (42)
Lumbar puncture	5	3* (60)	1 (20)
CS fluid analyzed by PCR for JCV	5	3* (60)	1 (20)
JCV DNA detected by PCR in CSF	0	0	0
Diagnosed with PML by IAC	0	0	0

*Received at least 1 VDZ infusion; Note: (%) in 3rd and 4th columns based on row totals in 2nd column; inf=infusions

Table 9 is the cross table indicates where the number of subjective checklist administered (rows) was less than the number of infusions (columns).

Table 9: No. of Assessments versus No. of Infusions

Number of Subj. Check. Administered	Number of Infusions		
	12	30	43
10	1	0	0
26	0	1	0
42	0	0	1

Only three patients were not administered the subjective checklist consistently at each of the vedolizumab infusion sessions.

3.3.4.5 PML Risk Assessment

Section 8.5 of the ISS and Module 5.3.5.4 (titled, "Estimate of Risk of PML") discuss PML risk assessment. In these sections, the applicant discussed how to characterize the theoretical risk of PML using data for natalizumab, an integrin receptor antagonist that is in the same class of treatments as vedolizumab. Natalizumab is indicated for the treatment of MS and CD and is known to increase the risk of PML. The applicant's objective was to evaluate the risk of PML in patients exposed to vedolizumab by using the PML risk observed in the natalizumab-exposed population as a benchmark.

The applicant stated that, based on the publication by Bloomgren et al. (2012), as of February 29, 2012, a total of 99,571 patients have been exposed to natalizumab and 212 PML cases have been confirmed; yielding a PML incidence rate of 2.13/1000 patients.

The worst possible scenario, upper bound of the incidence rate can be calculated. The "rule of three" states that in a study where no events are observed, the 95% confidence upper bound for the true event rate is approximately 3/n, where n is the study sample size (Jovanovic and Levy, 1997). For example, if no PML cases are observed in a study with 3,000 patients, then the true rate of PML will be lower than 3/3,000 (or 1 cases in 1,000 patients) with 95% confidence. The vedolizumab safety database included a total of 3,326 patients exposed to at least one dose and zero confirmed PML cases. Using the rule of three, the applicant concluded that the upper bound of the 95% confidence interval (CI) of the risk estimate for PML in vedolizumab-exposed patients was 0.90/1000 patients.

The reviewer used the rule of three and calculated the 95% upper bound of the PML incidence rate for vedolizumab in all vedolizumab-exposed patients and all vedolizumab-exposed patients with RAMP assessments, as 0.9 and 1.1 in 1,000 patients, respectively (based on the number of patients in Table 5, section **3.3.4.2. Exposure**). When limited to patients with at least 24 infusions and at least 4 weeks of follow-up, the upper bound is approximately 3.0 in 1,000 patients.

Table 10 compares the upper bound of the 95% CI of PML rates in patients treated with natalizumab and in that of the vedolizumab safety database, stratified by minimum duration of exposure. The estimates of the natalizumab PML rates were derived by the applicant using the Bloomgren publication. The applicant stated that they demonstrated a lower risk of PML among

vedolizumab patients compared to that among natalizumab patient, when considering duration of exposure.

Copyright Material

Source: Bloomgren, Richman et al. 2012.[3]

Abbreviations: CI = confidence interval; PML = progressive multifocal leukoencephalopathy.

Table 3-1, page 4, of the applicant's document titled, "Progressive Multifocal Leukoencepalopathy Risk Assessment for Vedolizumab".

The applicant further discussed a theoretical approach of evaluating the risk of PML in patients exposed to vedolizumab by using the estimated risk of PML among MS patients treated with natalizumab, stratified by risk factors. The three known natalizumab PML risk factors are: longer duration of treatment (beyond 24 months), prior immunosuppressant use, and positive anti-JCV antibody. The applicant used stratified PML incidence rates for natalizumab (Table 11) on the current vedolizumab safety database to estimate the total expected number of PML cases in the vedolizumab safety database if the PML risk and risk factors for vedolizumab users were similar to that of natalizumab.

Copyright Material

Source: Bloomgren, Richman et al. 2012.[3]

Abbreviations: JCV = John Cunningham virus; PML = progressive multifocal leukoencephalopathy.

Source: Table 3-3 of the applicant's document titled, "Progressive Multifocal Leukoencepalopathy Risk Assessment for Vedolizumab" (page 5).

Table 12: Vedolizumab Patient Exposure Stratified by Natalizumab PML Risk Factors

Vedolizumab Exposure	Anti-JCV Antibody Positive*		Anti-JCV Ab Neg	Total
	No Prior Immunosuppressant Use	Prior Immunosuppressant Use		
1-24 months	333	1330	1663	3326
25-48 months	91	362	453	906

Abbreviations: Ab Neg = antibody negative; JCV = John Cunningham virus; PML = progressive multifocal leukoencephalopathy.

* Anti-JCV antibody status for vedolizumab is estimated.

Source: Table 3-4 of the applicant's document titled, "Progressive Multifocal Leukoencepalopathy Risk Assessment for Vedolizumab" (page 6).

The applicant's approach is described as follows. First, all 3,326 patients exposed to at least one vedolizumab dose were stratified by the three known natalizumba PML risk factors Approximately 80% of vedolizumab patients had prior immunosuppressant use. For anti-JCV antibody status, the applicant used published rates in the literature, and assumed that approximately 50% of patients to be JC virus antibody positive.

Table 13: Expected Number of PML Cases If Risk Similar to Natalizumab

Vedolizumab Exposure	Anti-JCV Antibody Positive	
	No Prior Immunosuppressant Use	**Prior Immunosuppressant Use**
1-24 months	0.19	2.13
25-48 months	0.42	4.02
	Total Expected Cases of PML:	**6.75**
	Probability of Observing Zero Cases:	**0.0012**

Abbreviations: JCV = John Cunningham virus; PML = progressive multifocal leukoencephalopathy.

Source: Table 3-5 of the applicant's document titled, "Progressive Multifocal Leukoencepalopathy Risk Assessment for Vedolizumab" (page 7).

Then the applicant applied the risk-stratified PML incidence rates for natalizumab to the vedolizumab safety database stratified by natalizumab PML risk factors. Bloomgren's stratified PML rates in Table 11 were multiplied with the corresponding number of patients in Table 12 to obtain the expected number of PML cases per stratification cell in Table 13. The total expected number of PML cases in the vedolizumab safety database was therefore 6.75 (=0.19+2.13+0.42+4.02). Lastly, the applicant assumed that PML occurrence among the 3,326 vedolizumab-exposed patients followed a Poisson distribution with a mean of 6.75. Figure 3 illustrates the probability of observing cases of PML under this assumption; the probability of observing zero PML cases in the current safety database was very low (~0.1%).

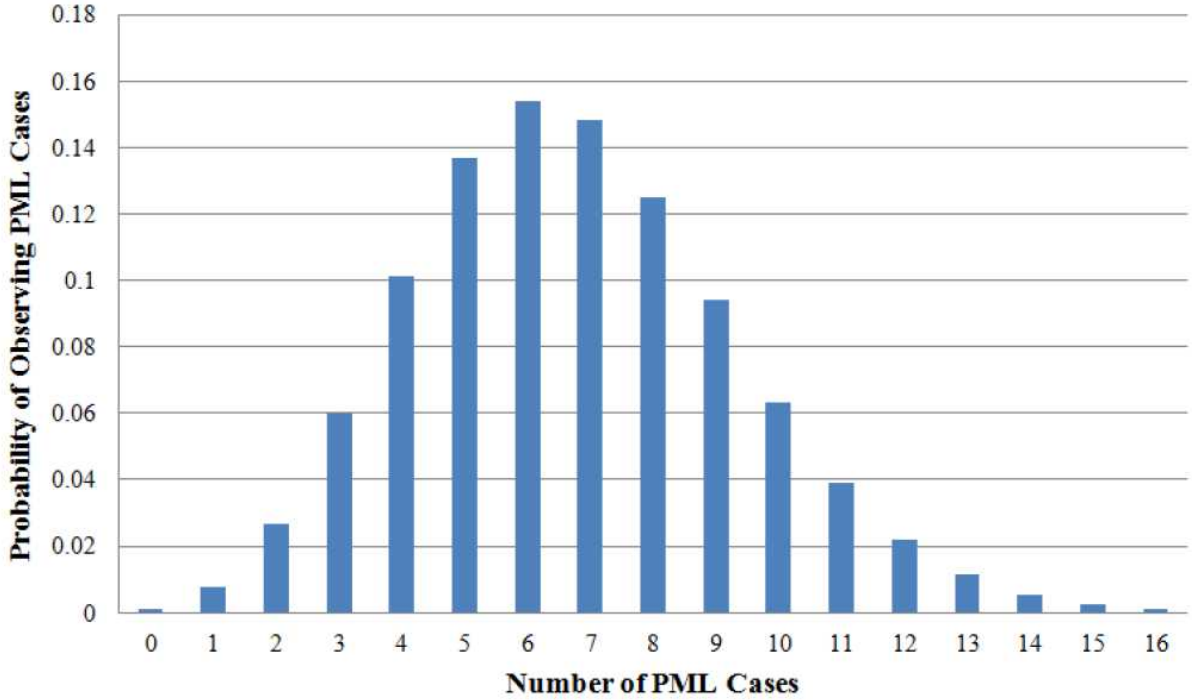
The applicant concluded that if the risk of PML among vedolizumab users were similar to natalizumab users, it would be almost certain that a PML case would occur. Because no PML cases were observed in the vedolizumab safety database, the PML risk is, therefore, lower for vedolizumab than for natalizumab.

Reviewer's Comment: The applicant's approach is appealing because the three risk factors for PML are all accounted for in the calculations. However, it implicitly involves a comparison of the risk of PML for natalizumab and vedolizumab and should, as discussed in the previous comment, be considered crude and interpreted with caution. Also, it is not clear whether the three PML risk factors among natalizumab users are also risk factors for vedolizumab users and whether they interact amongst each other in the same manner as in natalizumab in modifying the PML incidence rate.

Also, natalizumab and vedolizumab are treatments that are intended for distinctly different populations; Bloomgren et al.'s study was primarily on natalizumab exposed MS patients

while vedolizumab is intended for UC and CD patients. Moreover, the MS patients in Bloomgren's article were from clinical trials, observation studies (including a Swedish registry), and also from spontaneous reports, while vedolizumab UC and CD patients were from clinical trials only.

Figure 3: Poisson Probability Distribution of the Likelihood of Observing Cases of PML with Vedolizumab if Risk Similar to Natalizumab



Source: Figure 3-1 of the applicant's document titled, "Progressive Multifocal Leukoencepalopathy Risk Assessment for Vedolizumab" (page 7).

Although the vedolizumab and natalizumab safety databases are somewhat comparable in exposure time and number of patients, and no PML cases for vedolizumab were identified, this does not imply that the PML risk for vedolizumab is lower than the PML risk for natalizumab.

If the risk of PML for vedolizumab were to be directly compared with natalizumab, both treatments would have to be included in the same study.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

No analyses were performed across subgroups defined by gender, race, age, and geographic region.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

The reviewer did not identify any major statistical issues in the vedolizumab BLA except the applicant's proposed analysis evaluating the risk of PML in patients exposed to vedolizumab by using the natalizumab-exposed population as a benchmark for PML risk. As discussed in section **3.3.4.5 PML Risk Assessment**, the comparison of the risk of PML for natalizumab and vedolizumab should be considered crude and interpreted with caution because these treatments are intended for different patient populations and the PML risk data available are also different. For vedolizumab, the only available patient data were from clinical trials, whereas for natalizumab, patient data were available from clinical trials, observation studies (including a Swedish registry), and also from spontaneous reports. The available data for natalizumab has a much larger total number of patients and longer average durations of exposure than data for vedolizumab.

5.2 Collective Evidence

There were 3,603 patients in the safety population including 3,148 patients in controlled studies and 455 de novo patients in long-term, uncontrolled studies. There were 3,326 patients who were exposed to at least one dose of vedolizumab. The long-term, uncontrolled study C13008 was an ongoing study at the time of this review and enrolled 2,243 patients (or 67% of all vedolizumab-exposed patients). A large portion (81%) of patients in this study was rolled-over from previous studies (about 80% of patients came from studies C13006 and C13007 combined while the rest were from studies C13004 and C13011). As of June 28, 2013, the safety population included 1,004 patients with at least 24 vedolizumab infusions and at least 4 weeks of follow-up. The mean number of months of exposure was around 15 months (when considering all vedolizumab-exposed patients) or 17 months (when considering all vedolizumab-exposed patients who were assessed by the RAMP algorithm) and the mean number of infusions with at least 4 weeks of follow-up was around 18 infusions. The amount of vedolizumab exposure was mainly driven by the exposure from patients who were enrolled in study C13008.

Approximately 80% of patients had prior immunosuppressant use and approximately 30% of patients had concomitant immunosuppressant use.

The applicant's RAMP algorithm included subjective checklists that appeared to be consistently administered to patients prior to each vedolizumab infusion session. The algorithm resulted in about 10% (284 of 2,927) of patients with positive subjective findings and about 3% (82 of 2,927) of assessed patients referred to a neurologist. At most 3% of assessed patients had other further diagnostics performed. The assessments of PML symptoms in the RAMP did not appear to be observed mostly in patients with ≥24 infusions. There were no confirmed PML cases in the vedolizumab BLA.

5.3 Conclusions and Recommendations

Although the vedolizumab and natalizumab safety databases were somewhat comparable in exposure time and number of patients, and no PML cases were identified in the vedolizumab safety database, this does not imply that the PML risk for vedolizumab is lower than the PML risk for natalizumab.

Using the rule of three (Jovanovic and Levy, 1997), the upper bound of the 95% confidence interval of the PML incidence rate for vedolizumab was 0.9–1.1 cases in 1,000 patients, all and patient with RAMP assessments respectively. These risk estimates do not take duration of treatment into consideration and includes all exposed subjects who received at least one dose of vedolizumab. When limited to patients with at least 24 infusions and at least 4 weeks of follow-up, the upper bound was approximately 3.0 cases in 1,000 patients.

From the clinical trials of natalizumab, with only 3 PML cases in 3,000 patients, it was not possible to determine PML risk factors. It was only after several years, when close to 100,000 patients were exposed to natalizumab for over 200,000 patient-years were observed that three risk factors were confirmed and consequently possible to estimate the stratified PML incidence rates. Thus, with no PML case in the safety database with roughly 3,000 patients treated with vedolizumab, it is unclear whether these same PML risk factors for natalizumab would be actual risk factors for vedolizumab and yield similar estimated stratified PML incidence rates. In order to assess PML risk for vedolizumab, a larger number of patients with longer duration of exposure would need to be observed. If the risk of PML for vedolizumab were to be directly compared with natalizumab, both treatments would have to be included in the same study.

5.4 Labeling Recommendations

The reviewer has no disagreement with the applicant's proposed label.

REFERENCES

Bloomgren, G, Richman, S., Hotermans, C., et al. Risk of Natalizumab-Associated Progressive Multifocal Leukoencephalopathy. *N Engl J Med* 2012; 366(20):1870-80.

Jovanovic, B.D. and Levy, P.S. A Look at the Rule of Three. *The American Statistician*, vol. 51, no. 2 (May 1997), pp. 137-139.

APPENDIX

Table 14 summarizes the number of patients who were rolled-over into studies C13004 and C13008 from previous studies. There were 72 patients in study C13004 of which 38 patients were rolled-over from study C13002 and 34 were *de novo* patients. Among the 38 rolled-over patients, 31 patients were previously on VDZ and 7 patients were previously on placebo. Therefore, 41 patients (34 de novo and 7 previously on placebo) were exposed to VDZ in study C13004 for the first time.

In Study C13008, a large portion (81%) of all patients enrolled in this study were rolled-over from studies C13004, C13006, C13007, and C13011 while the rest (19%) were *de novo* patients. Among rolled-over and de novo patients, there were generally more CD than UC patients. *Note: The numbers of patients in study C13008 in Table 14 were also presented by the applicant in Table 2.* There were 830 patients (421 de novo and 409 previously on placebo) who were exposed to VDZ in study C13008 for the first time.

Table 14: Uncontrolled Long Term Safety Studies

Study	Number Dosed		Total
	UC	CD	
C13004	53	19	72
rollover	38	0	38 (53)
de novo	15	19	34 (47)
C13008	894	1,349	2,243
rollover	704	1118	1,822 (81)
de novo	190	231	421 (19)
Total	947	1,368	2,315

Table 15 summarizes the numbers of patients in study C13008 who were rolled-over from studies C13004, C13006, C13007, and C13011. Most patients were rolled-over from studies C13007 (40%) and C13006 (37%). *Note: The applicant provided the same numbers as in Table 15 (but with VDZ/VDZ Q4W and VDZ/VDZ Q8W combined in one column) on September 3, 2013 in a response to an information request.*

Table 15: Number of Patients Rolled-Over to Study C13008 from Previous Studies

Study	Study Group						Total
	PLA	VDZ	PLA/PLA	VDZ/PLA	VDZ/VDZ Q4W	VDZ/VDZ Q8W	
C13004	-	37	-	-	-	-	37 (2)
C13006	-	-	112	113	342	108	675 (37)
C13007	-	-	107	127	366	126	726 (40)
C13011	190	194	-	-	-	-	384 (21)
Total	190	231	219	240	708	234	1822

PLA/PLA=placebo at induction and maintenance; VDZ/PLA=VDZ at induction and placebo at maintenance; VDZ/VDZ=VDZ at induction and maintenance

---

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

JOHN S YAP
11/20/2013

CLARA KIM
11/20/2013
I concur with this review.

ALOKA G CHAKRAVARTY
11/20/2013

STATISTICAL SAFETY REVIEW AND EVALUATION
FILING REVIEW OF A NDA/BLA

NDA/BLA Number: 125476
NDA/BLA Type: BLA, Standard Review
Product Name: Entyvio (vedolizumab)
Indication Sought: Ulcerative colitis and Crohn's disease
Safety Issue: Progressive Multifocal Leucoencephalopathy (PML)
Applicant: Takeda Pharmaceuticals, Inc.
Received Date: 6/20/2013
Reviewer: John Stephen Yap

1. Brief Summary of Controlled Clinical Trials

Table 1: Summary of Trials to be Assessed in the Statistical Safety Review

Trial ID	Design	Treatment/ Sample Size	Endpoint/Analysis	Preliminary Findings
C13002	P2, R, PC, DB, MD	VDZ/37 Placebo/9	PK, PD,..., PML	no PML
C13004	P2, OL, MD, LT	VDZ/53 UC, 19 CD (72 total) Placebo/0	Efficacy,..., PML	no PML
C13006	P3, R, PC, DB	VDZ/746 Placebo/149	Efficacy,..., PML	no PML
C13007	P3, R, PC, DB	VDZ/967 Placebo/148	Efficacy,..., PML	no PML
C13008	P3, OL, LT	VDZ/704 UC, 1118 CD rollover, 190 UC, 231 CD de novo (894 UC, 1349 CD; 2243 total) Placebo/0	..., PML	no PML
C13011	P3, R, PC, DB	VDZ/209 Placebo/207	Efficacy,..., PML	no PML
P2=phase 2, R=randomized, PC=placebo-controlled, DB=double-blind, MD=multiple-dose, OL=open-label, LT=long-term; UC=ulcerative colitis; CD=Crohn's diseases;				

PML=progressive multifocal leucoencaphalopathy

2. Assessment of Protocols and Study Reports

Table 2: Summary of Information from Review of the Protocol(s) and the Study Report.

Content Parameter	Response/Comments
Designs utilized are appropriate for the indications requested.	NA; Safety assessment for PML will involve pooling of safety patients across all studies listed above (i.e. studies included in the ISS).
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.	NA; The ISS includes a statistical analysis plan that specifies the endpoints and method of analysis.
Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available.	NA
Appropriate references for novel statistical methodology (if present) are included.	NA
Safety data organized to permit analyses across clinical trials in the NDA/BLA.	Yes
Investigation of effect of missing data and discontinued follow-up on statistical analyses appears to be adequate.	NA

3. Electronic Data Assessment

Table 3: Information Regarding the Data

	Content Parameter	Response/Comments
1	Dataset location	\\cdsesub1\bla\ectd_submissions\stn125476\125476.enx
2	Dataset structure (e.g., SDTM or ADaM)	SDTM and ADAM
3	Based on the analysis datasets, can results of the primary endpoint(s) be reproduced? (Yes	Yes

	Content Parameter	Response/Comments
	or No)	
4	List the dataset(s) that contains the primary endpoint(s)	ADCF (analysis data clinical findings) includes PML assessment.
5	Are there any concerns about site(s) that could lead to inspection? If so, list of site(s) that needs inspection and rationale	NA
6	Are the define files sufficiently detailed?	Yes
7	Safety data are organized to permit analyses across clinical trials in the NDA/BLA.	Yes

4. Filing Issues

Table 4: Initial overview of the NDA/BLA application for refuse-to-file (RTF):

Content Parameter	Yes	No	NA	Comments
Index is sufficient to locate necessary reports, tables, data, etc.	✓			
ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)	✓			ISE not reviewed.
Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated.	✓			Efficacy not reviewed.
Data sets in EDR are accessible, sufficiently documented, and of sufficient quality (e.g., no meaningful data errors).	✓			

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? Yes

If the NDA/BLA is not fileable from the statistical perspective, state the reasons and provide comments to be sent to the Applicant.

Please identify and list any potential review issues to be forwarded to the Applicant for the 74-day letter.

Reviewing Statistician Date

Supervisor/Team Leader Date

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

JOHN S YAP
08/13/2013

CLARA KIM
08/13/2013
I concur with this review.