

**CENTER FOR DRUG EVALUATION AND  
RESEARCH**

*APPLICATION NUMBER:*

**125370**

**STATISTICAL REVIEW(S)**



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 Number:** STN 125370-0000

**Drug Name:** Belimumab (BENLYSTA, Monoclonal Anti-BLyS Antibody)

**Indication(s):** Treatment of adult patients with active, autoantibody positive systemic lupus erythematosus (SLE) who are receiving standard therapy

**Applicant:** Human Genome Sciences, Inc.

**Date(s):** Stamp date: June 9, 2010  
PDUFA date: December 9, 2010

**Review Priority:** Priority

**Biometrics Division:** Division of Biometrics II

**Statistical Reviewer:** Ruthanna C. Davi *Ruthanna C. Davi 2/18/2011*

**Concurring Reviewers:** Joan Buenconsejo, Acting Statistical Team Leader *Joan Buenconsejo 2/18/2011*  
Tom Permutt, Division of Biometrics II Director *Tom Permutt 2/18/11*

**Medical Division:** Division of Pulmonary, Allergy, and Rheumatology Products

**Clinical Team:** Rosemarie Neuner, Medical Reviewer  
Sarah Okada, Medical Team Leader

**Project Manager:** Jessica Benjamin, Philantha Bowen

**Keywords:**  
BLA, clinical studies, missing data, subgroup analyses

# Table of Contents

<b>1. EXECUTIVE SUMMARY.....</b>	<b>3</b>
1.1 CONCLUSIONS AND RECOMMENDATIONS.....	3
1.2 BRIEF OVERVIEW OF CLINICAL STUDIES .....	3
1.3 STATISTICAL ISSUES AND FINDINGS.....	4
<b>2. INTRODUCTION.....</b>	<b>6</b>
2.1 OVERVIEW .....	6
2.2 DATA SOURCES.....	7
<b>3. STATISTICAL EVALUATION.....</b>	<b>7</b>
3.1 EVALUATION OF EFFICACY .....	7
3.2 EVALUATION OF SAFETY .....	18
<b>4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS.....</b>	<b>19</b>
<b>5. SUMMARY AND CONCLUSIONS.....</b>	<b>22</b>
5.1 STATISTICAL ISSUES AND COLLECTIVE EVIDENCE.....	22
5.2 CONCLUSIONS AND RECOMMENDATIONS.....	23

# 1. EXECUTIVE SUMMARY

## 1.1 Conclusions and Recommendations

Studies 1056 and 1057 adequately demonstrate that the proportion of subjects achieving SRI response at week 52 is higher with Belimumab 10 mg/kg than placebo. However, the magnitudes of the treatment effects observed in these studies were small and the results were somewhat dependent on the handling of missing data due to “medication failure”. Clinical interpretation regarding the importance of these two factors is needed. Subgroup efficacy analyses suggest that Belimumab 10 mg/kg could be harmful in black subjects. Labeling clearly communicating this data and further investigation of the effect of Belimumab in black subjects is needed.

Numerical results for the Belimumab 1 mg/kg to placebo comparison were similar to the 10 mg/kg Belimumab to placebo comparison but did not reach statistical significance in both studies.

Analyses of secondary efficacy endpoints (i.e., prednisone reduction by  $\geq 25\%$  from baseline to  $\leq 7.5$  mg/day during weeks 40 through 52 and severe flares) and the SRI response at 76 weeks were not consistently significantly better for either Belimumab dose relative to placebo in both studies 1056 and 1057.

## 1.2 Brief Overview of Clinical Studies

The sponsor has submitted the results of two similarly designed phase 3 pivotal studies to support the regulatory approval of Belimumab for treatment of adult patients with active, autoantibody positive systemic lupus erythematosus (SLE) who are receiving standard therapy.

The pivotal studies referred to as 1056 and 1057 are each titled, “A Phase 3, Multi-Center, Randomized, Double-blind, Placebo-Controlled, 76-Week [52-Week for 1057] Study to Evaluate the Efficacy and Safety of Belimumab (HGS1006, LymphoStat-B™), a fully Human Monoclonal Anti-BLyS Antibody, in Subjects with Systemic Lupus Erythematosus (SLE)”. As part of these studies, subjects were randomly assigned to one the following treatment groups in a 1:1:1 ratio: Belimumab 1 mg/kg, Belimumab 10 mg/kg, or placebo. Randomization was stratified by subjects’ screening SELENA SLEDAI score (6-9 vs  $\geq 10$ ), screening proteinuria level ( $< 2$  g/24 hour vs  $\geq 2$  g/24 hour equivalent) and race (African descent or indigenous-American descent vs other). The primary difference between studies 1056 and 1057 was the duration of the study. In study 1056 treatment was planned to be continued for 76 weeks while in study 1057 treatment was to planned to be 52 weeks. The primary time point for assessment of efficacy, however, was the same in both studies, 52 weeks. The primary efficacy objective of the studies was to demonstrate that for each Belimumab group a higher proportion of subjects achieved the primary efficacy composite endpoint, referred to as the SLE Responder Index (SRI), at week 52 compared to that in the placebo group. A success for the primary efficacy endpoint was defined as a subject who met the following criteria.

- $\geq 4$  point reduction from baseline in SELENA SLEDAI score, and

- No worsening (increase of <0.30 points from baseline) in physician's global assessment (PGA), and
- No new BILAG A organ domain score or 2 new BILAG B organ domain scores compared with baseline at the time of assessment

Numerous secondary efficacy endpoints were also examined as part of these studies. Among these, prednisone use, and lupus flares were selected by the FDA medical team as being of particular interest and thus are examined, along with the primary efficacy endpoint, in this review. In addition, primary efficacy response at week 76 was at the recommendation of the FDA pre-specified by the sponsor as a major secondary endpoint for study 1056 and is therefore examined in this review. For the statistical review of safety, mortality was highlighted by the FDA medical team as important for this application and thus is commented upon in this review.

### 1.3 Statistical Issues and Findings

The following statistical issues and their impact have been described in the context of the review. Please refer to the specified section for details. Issues that are of particular consequence are shown below in boldface type.

- In each study, approximately 20% of randomized subjects discontinued treatment before week 52. The most common reasons for early treatment discontinuation were subject request, adverse event, and lack of efficacy. Many but not all of these subjects also discontinued the study. Considering subjects who discontinued the study as failures for the primary efficacy analysis is likely a fair representation of the efficacy in these subjects in that the subjects' reasons for withdrawal from treatment indicate the study treatment could not be tolerated in exchange for whatever efficacy may have been being achieved. Therefore, the primary efficacy results in the mITT group likely remain reliable despite the fairly high early treatment discontinuation rate.
- No significant differences between treatment groups in the demographic and baseline characteristics in the mITT groups for studies 1056 or 1057 were noted. As would be expected due to the random treatment assignment, balance among the treatment groups in demographic and baseline characteristics appears adequate to allow by-treatment group differences in post-randomization outcomes to be attributed to treatment effects and not an artifact of an imbalance in pre-randomization characteristics.
- **While the magnitude of the differences between treatment groups were fairly small (observed difference between Belimumab 10 mg/kg and placebo of 9% and 14% for studies 1056 and 1057, respectively), in each study, the Belimumab 10 mg/kg had a statistically significantly higher SRI success rate than the placebo group (p=0.02 and p=0.0006 for studies 1056 and 1057, respectively). A statistically higher rate of SRI response for the Belimumab 1 mg/kg group as compared to placebo was demonstrated for only study 1057 (p=0.02). Reviewer analyses indicate that these conclusions are consistent in the face of slight variations in the logistic regression models employed.**
- The results from the analyses of the subcomponents of the SRI were generally consistent with those of the primary analysis. The proportions of subjects achieving success for each of the subcomponents of the SRI were numerically higher in the Belimumab groups than the placebo group in each study.

- Subjects who dropped out of the study early were considered failures for the primary efficacy analysis. The proportions of subjects who dropped out are approximately 16% in study 1056 and 12% in study 1057 and are fairly balanced across treatment groups within each study thus the impact of imputing dropouts as failures on the treatment effect in the primary analysis should be small.
- **Subjects who were “medication failures” were considered failures for the primary efficacy analysis. The rates of “medication failures” are not balanced across treatment groups (17%, 9%, and 10% for placebo, 1 mg/kg Belimumab, and 10 mg/kg Belimumab respectively in study 1056 and 11%, 7%, and 6% for the same in study 1057). Since medication failures are more frequent in the placebo groups than the Belimumab groups, imputing medication failures as efficacy failures could bias the treatment effect in the primary efficacy endpoint in favor of Belimumab (unless these subjects would truly have been unable to achieve success on the primary endpoint had they not taken the prohibited medication).**
- No multiplicity correction was planned for in the protocol for the secondary endpoints; however, two of the numerous secondary efficacy endpoints examined in these studies (i.e., prednisone reduction by  $\geq 25\%$  from baseline to  $\leq 7.5$  mg/day during weeks 40 through 52 and severe flares) were of particular interest to the FDA medical team for evaluation of the efficacy of Belimumab. The proportion of subjects who reduced their average prednisone dose by at least 25% to  $\leq 7.5$  mg/day during Weeks 40 through 52 were not consistently significantly different for either Belimumab dose relative to placebo in both studies. The risk of experiencing a severe flare was not consistently significantly reduced for either Belimumab dose relative to placebo for both studies 1056 and 1057.
- The SRI response at week 76 was at the recommendation of the FDA pre-specified by the sponsor as a major secondary endpoint for study 1056. There were no statistically significant differences between either Belimumab dose and placebo at week 76.
- **In study 1056, a significant treatment-by-race interaction for both Belimumab groups versus placebo suggest that there may be a reversal of the treatment effect (i.e., a qualitative interaction) in the AIA race category versus other race category ( $p=0.03$  and  $p=0.009$  for 1 mg/kg and 10 mg/kg, respectively). Response rates of subjects of AIA races were highest in the placebo group while response rates of subjects of other races were highest in the Belimumab groups. Exploratory subgroup analyses by race with a different categorization (black vs. white vs. Alaska Native or American Indian vs. other) also indicate that there may be a reversal of treatment effect in black subjects. This reversal of treatment effect in AIA subjects and/or black subjects could be due to the disproportionately high rate of dropouts in the 10 mg/kg group. This pattern of dropouts (i.e., an increasing rate of dropouts with increasing dose) is consistent with that of a drug with an undesirable or toxic effect at higher doses. Within the AIA stratum, the rates of dropouts due to subject request, AE, and lost to follow-up are numerically higher for the 10 mg/kg Belimumab group than the other groups.**

## 2. INTRODUCTION

### 2.1 Overview

The sponsor has submitted the results of two similarly designed phase 3 pivotal studies to support the regulatory approval of Belimumab for treatment of adult patients with active, autoantibody positive systemic lupus erythematosus (SLE) who are receiving standard therapy.

The pivotal studies referred to as 1056 and 1057 are each titled, “A Phase 3, Multi-Center, Randomized, Double-blind, Placebo-Controlled, 76-Week [52-Week for 1057] Study to Evaluate the Efficacy and Safety of Belimumab (HGS1006, LymphoStat-B™), a fully Human Monoclonal Anti-BLyS Antibody, in Subjects with Systemic Lupus Erythematosus (SLE)”. As part of these studies, subjects were randomly assigned to one the following treatment groups in a 1:1:1 ratio: Belimumab 1 mg/kg, Belimumab 10 mg/kg, or placebo. Randomization was stratified by subjects’ screening SELENA SLEDAI score (6-9 vs  $\geq 10$ ), screening proteinuria level ( $< 2$  g/24 hour vs  $\geq 2$  g/24 hour equivalent) and race (African descent or indigenous-American descent vs other). The primary difference between studies 1056 and 1057 was the duration of the study. In study 1056 treatment was planned to be continued for 76 weeks while in study 1057 treatment was to planned to be 52 weeks. The primary time point for assessment of efficacy, however, was the same in both studies, 52 weeks. The primary efficacy objective of the studies was to demonstrate that for each Belimumab group a higher proportion of subjects achieved the primary efficacy composite endpoint, referred to as the SLE Responder Index (SRI), at week 52 compared to that in the placebo group. A success for the primary efficacy endpoint was defined as a subject who met the following criteria.

- $\geq 4$  point reduction from baseline in SELENA SLEDAI score, and
- No worsening (increase of  $< 0.30$  points from baseline) in physician’s global assessment (PGA), and
- No new BILAG A organ domain score or 2 new BILAG B organ domain scores compared with baseline at the time of assessment

Numerous secondary efficacy endpoints were also examined as part of these studies. Among these, prednisone use, and lupus flares were selected by the FDA medical team as being of particular interest and thus are examined, along with the primary efficacy endpoint, in this review. In addition, primary efficacy response at week 76 was at the recommendation of the FDA pre-specified by the sponsor as a major secondary endpoint for study 1056 and is therefore examined in this review. For the statistical review of safety, mortality was highlighted by the FDA medical team as important for this application and thus is commented upon in this review.

Communication with the sponsor regarding these studies is documented under IND 9970. Pertinent parts of the statistical portion of those communications are summarized herein. The design and analysis of the phase 3 studies was discussed at the End-of-Phase 2 meeting held on April 26, 2006. The Division informally agreed with the sponsor’s proposal for the primary efficacy endpoint and to the 52-week time point for analysis. The Division also informally agreed to the statistical analysis plan for the primary efficacy endpoint and to the structure of the Data Monitoring Committee and frequency of data review. The Division did not agree with the sponsor’s proposal regarding which background medications should be controlled and suggested several additional medications that should be controlled. Negotiation regarding these protocols

continued over the next several months and finally, the Division's agreement to them was formally documented in response to the sponsor's request for a special protocol assessment (Division letter dated October 19, 2006).

## 2.2 Data Sources

The following data sets were submitted electronically and utilized in the review of this study.

R:\STN125370\0000\m5\datasets\c1056\analysis\comp.xpt  
R:\STN125370\0000\m5\datasets\c1056\analysis\respons.xpt  
R:\STN125370\0000\m5\datasets\c1056\analysis\rspwk52.xpt  
R:\STN125370\0000\m5\datasets\c1056\analysis\subjchar  
R:\STN125370\0000\m5\datasets\c1057\analysis\respons.xpt  
R:\STN125370\0000\m5\datasets\c1057\analysis\rspwk52.xpt  
R:\STN125370\0000\m5\datasets\c1057\analysis\subjchar

All submitted data sets were found to be adequately documented and organized.

## 3. STATISTICAL EVALUATION

### 3.1 Evaluation of Efficacy

#### 3.1.1 Study Design (Studies 1056 and 1057)

Studies 1056 and 1057 were multi-center, randomized, double-blind, parallel group, placebo-controlled studies with a primary efficacy objective of demonstrating superiority of each dose of Belimumab over placebo in terms of the primary efficacy endpoint.

Inclusion and exclusion criteria for this study were, at least in part, motivated by the results of the phase 2 LBSL02 study in which efficacy was not demonstrated in the entire group of subjects but benefit from Belimumab appeared most promising in the subgroup of subjects who were autoantibody positive. To be eligible for studies 1056 and/or 1057 subjects were required to have a clinical diagnosis of systemic lupus erythematosus (SLE) according to the ACR criteria and clinically active SLE disease, defined as a SELENA SLEDAI disease activity score of at least 6 at screening. Subjects had to have an unequivocally positive ANA test result, from two independent time points within the study screening period or one positive historical test result and one positive test result during the screening period. ANA test results obtained in the screening period were only considered positive if the ANA titer was  $\geq 1:80$  and/or anti-dsDNA serum antibody was  $\geq 30$  IU/mL. In addition, subjects were required to be on a stable SLE treatment regimen for a period of at least 30 days prior to enrollment consisting of the following alone or in combination: prednisone or equivalent (from 0 to 40 mg/day when used in combination with other SLE treatment or from 7.5 to 40 mg/day alone), anti-malarials, NSAIDs, or any immunosuppressive therapy (i.e., methotrexate, azathioprine, leflunomide, mycophenolate, calcineurin inhibitors, sirolimus, oral cyclophosphamide, 6-mercaptopurine, or thalidomide). In total, the protocol specified ten inclusion and 19 exclusion criteria for enrollment in these studies.

Eligible subjects were randomized to one the following treatment groups (in a 1:1:1 ratio) to be received for the entire treatment period.

- Belimumab 1 mg/kg
- Belimumab 10 mg/kg
- placebo

Randomization was stratified by subjects' screening SELENA SLEDAI score (6-9 vs  $\geq 10$ ), screening proteinuria level ( $< 2$  g/24 hour vs  $\geq 2$  g/24 hour equivalent) and race (African descent or indigenous-American descent vs other). Subjects were to be dosed with study medication on days 0, 14, and 28, and then every 28 days through 76 weeks for study 1056 and 52 weeks for study 1057.

The primary efficacy endpoint, referred to as the SRI, was a composite endpoint that required success on all of the following criteria at week 52 to be considered a success overall.

- $\geq 4$  point reduction from baseline in SELENA SLEDAI score, and
- No worsening (increase of  $<0.30$  points from baseline) in physician's global assessment (PGA), and
- No new BILAG A organ domain score or 2 new BILAG B organ domain scores compared with baseline at the time of assessment

Subjects who dropped out of the study before the week 52 visit were imputed as failures for the primary efficacy analysis. In addition, the protocol specified that once a subject was randomized and received the first dose of study medication, adjustment to concurrent medications (add, eliminate, change dose level/frequency at certain times) was allowed as clinically required; however, certain changes required that the subject be coded as being a failure for purposes of the primary efficacy analysis and be discontinued from the study. These subjects are subsequently referred to as "medication failures". These medication restrictions included certain changes in anti-malarials, steroids, other immunosuppressive/immunomodulatory agents, HMG CoA3-hydroxy-3methyl-glutaryl co-enzyme A reductase inhibitors, angiotensin pathway antihypertensives, NSAIDs and aspirin. Prohibited medications included other investigational agents, anti-TNF therapy, other biologics, intravenous immunoglobulin, IV cyclophosphamide, and plasmapheresis.

The primary efficacy analysis was designed to demonstrate that for each Belimumab group a higher proportion of subjects achieved success on the SRI at week 52 compared to that in the placebo group. A logistic regression model with the following independent variables in the model: treatment group, baseline SELENA SLEDAI score ( $\leq 9$  versus  $\geq 10$ ), baseline proteinuria level ( $< 2$  g/24 hour versus  $\geq 2$  g/24 hour equivalent) and race (AIA versus other) was protocol-specified for these comparisons. A step-down sequential testing procedure was used to control for multiplicity in doses. The Belimumab 10 mg/kg treatment group was to be compared with the placebo group (two-sided  $\alpha=0.05$ ) first and if statistically significant, the Belimumab 1 mg/kg treatment group was to be compared with the placebo group (two-sided  $\alpha=0.05$ ). The primary efficacy analysis was to be conducted in the modified intent to treat group (mITT) defined as all subjects randomized who received at least one dose of study medication.

Numerous secondary efficacy endpoints were also examined as part of these studies. No multiplicity correction was planned for in the protocol for the secondary endpoints. Among the secondary endpoints, prednisone use and lupus flares were selected by the FDA medical team as being of particular interest and thus are examined in this review. In addition, the SRI response at week 76 was at the recommendation of the FDA pre-specified by the sponsor as a major secondary endpoint for study 1056 and is therefore examined in this review.

The protocol required the use of an independent Data Monitoring Committee (DMC) for these studies. The DMC was to review subject safety after the first 100 subjects had been treated through day 56 in study 1056 and 1057 combined or within approximately six months of the treatment of the first subject, whichever came first. After the initial review, the committee was to review the safety data approximately every four months. No efficacy data was reviewed by the DMC and thus no adjustment to the significance level in the primary efficacy analysis was made.

For the statistical review of safety, mortality was highlighted by the FDA medical team as important for this application and is thus commented upon in this review

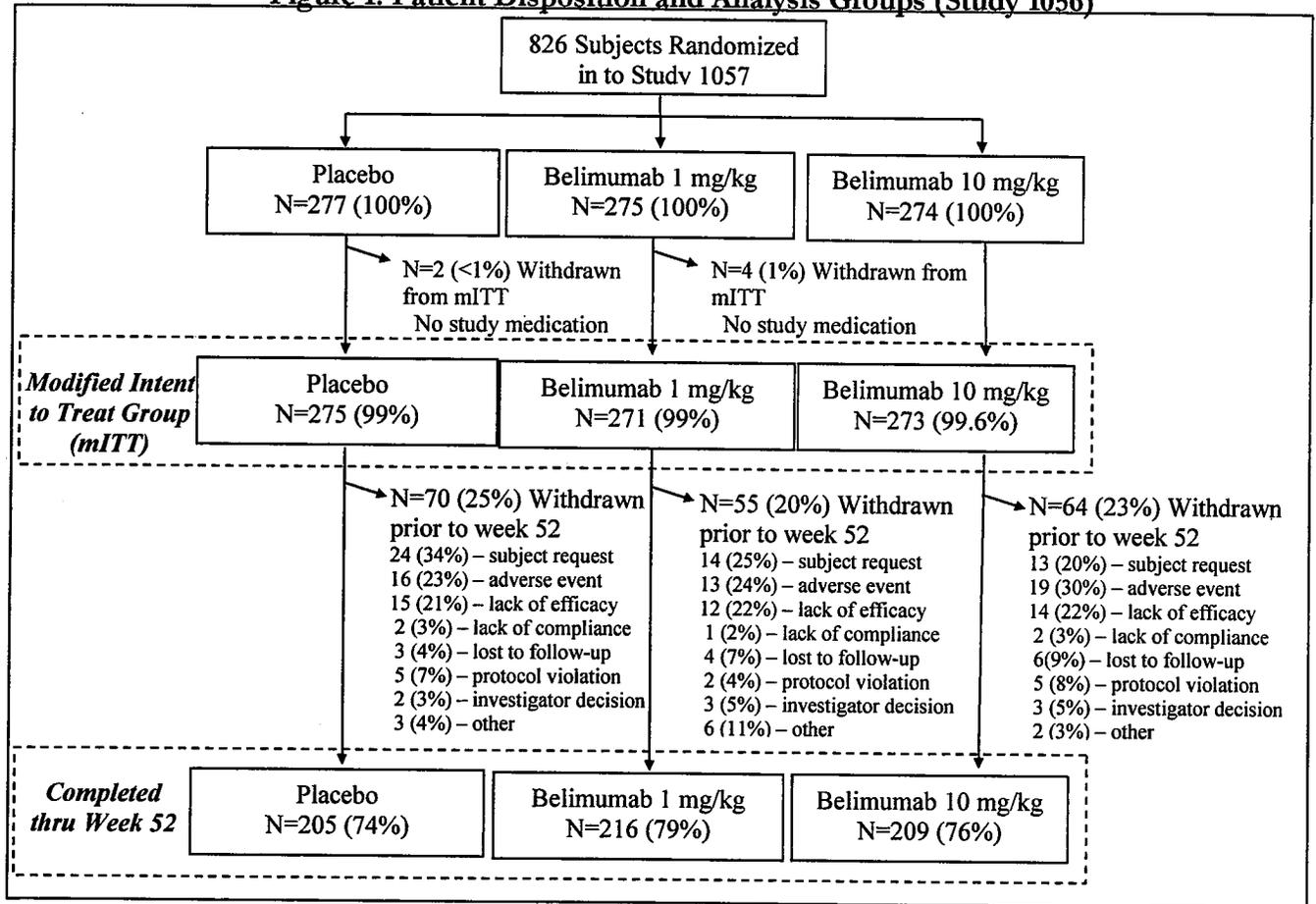
### **3.1.2 Results (Studies 1056 and 1057)**

Eight hundred twenty six subjects were randomized (1:1:1 stratified by subjects' screening SELENA SLEDAI score (6-9 vs  $\geq 10$ ), screening proteinuria level ( $< 2$  g/24 hour vs  $\geq 2$  g/24 hour equivalent) and race (African descent or indigenous-American descent vs other)) into study 1056 as follows: 277 to receive placebo, 275 to receive Belimumab 1 mg/kg and 274 to receive Belimumab 10 mg/kg. For study 1057, 867 subjects were randomized (1:1:1 stratified by subjects' screening SELENA SLEDAI score (6-9 vs  $\geq 10$ ), screening proteinuria level ( $< 2$  g/24 hour vs  $\geq 2$  g/24 hour equivalent) and race (African descent or indigenous-American descent vs other)) as follows: 288 to receive placebo, 289 to receive Belimumab 1 mg/kg and 290 to receive Belimumab 10 mg/kg. Seven subjects in study 1056 and 2 subjects in study 1057 did not receive study medication thus per protocol definition, there were 819 subjects in study 1056 and 865 subjects in study 1057 who were included in the mITT groups. Figures 1 and 2 describe the treatment randomizations, the inclusion or exclusion of subjects from the mITT analysis groups, and the rates of early treatment discontinuation for studies 1056 and 1057, respectively.

In both studies 1056 and 1057, exclusions from the mITT group were infrequent. In each study, approximately 20% of randomized subjects discontinued treatment before week 52. The most common reasons for early treatment discontinuation were subject request, adverse event, and lack of efficacy. The rates of early treatment discontinuation due to subject request were numerically lower in the Belimumab groups than placebo in each study while the rate of early treatment discontinuation due adverse event was slightly numerically higher in the Belimumab 10 mg/kg group than placebo in study 1056 and fairly consistent across treatment groups in study 1057. The rates of early treatment discontinuation due to lack of efficacy were fairly consistent across treatment groups in both studies. Many but not all of these subjects also discontinued the study. (Table 3 and the associated text further describe dropouts and "medication failures" in the context of the primary efficacy analyses.) Subjects

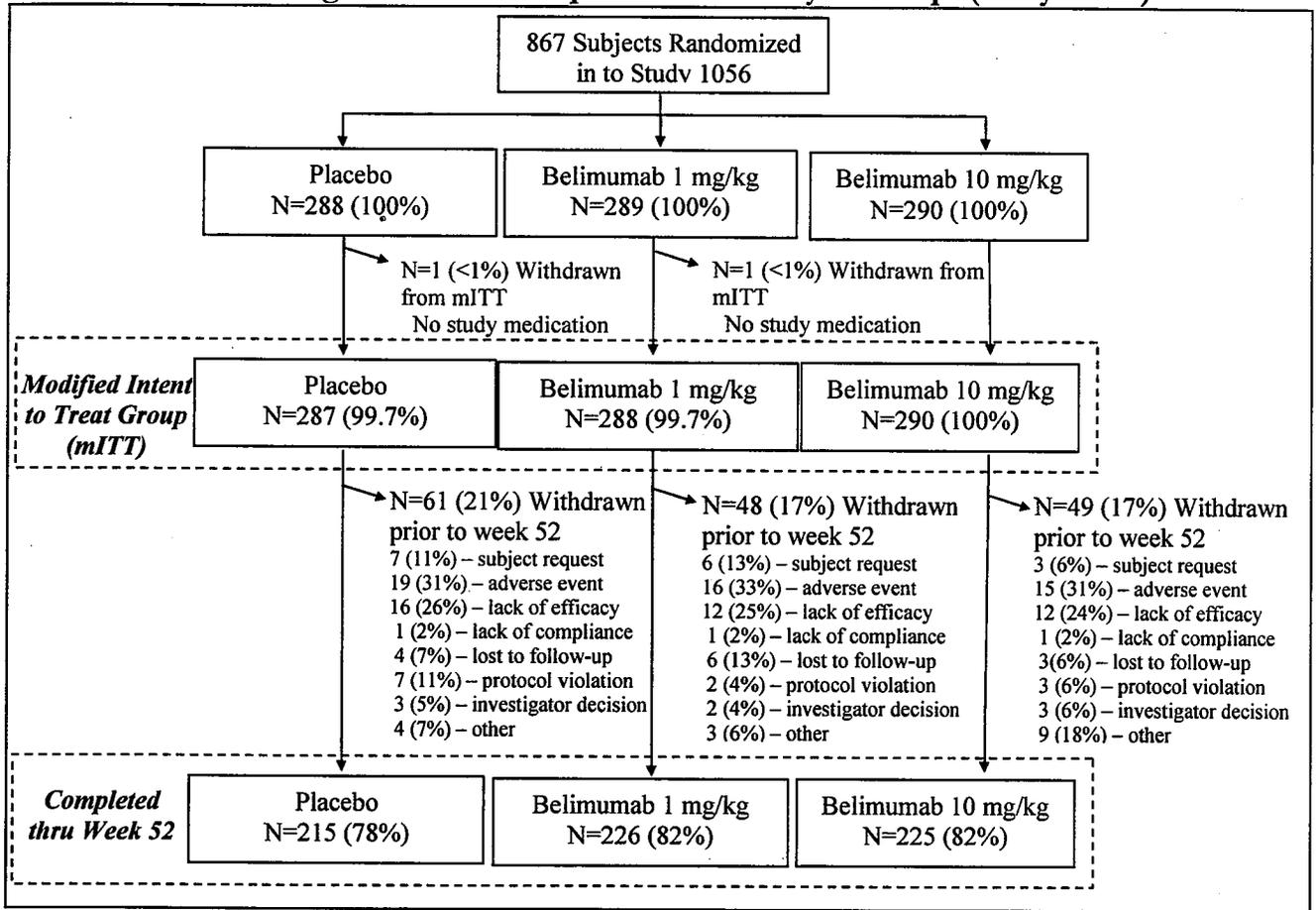
who dropped out of the study before the week 52 visit were, by protocol definition, imputed as failures for the primary efficacy analysis in the mITT group. This may be considered a fair representation of the efficacy in these subjects in that the subjects' reasons for withdrawal from treatment indicate the study treatment could not be tolerated in exchange for whatever efficacy may have been being achieved and thus for all intents and purposes, the study treatment failed for those subjects. Therefore, the primary efficacy results in the mITT group likely remain reliable despite the fairly high (but approximately balanced) early treatment discontinuation rate.

**Figure 1: Patient Disposition and Analysis Groups (Study 1056)**



Source: Sponsor analyses (Figure 6-1 clinical study report) and reviewer analyses

**Figure 2: Patient Disposition and Analysis Groups (Study C1057)**



Source: Sponsor analyses (Figure 6-1 clinical study report) and reviewer analyses

Selected demographic and baseline characteristics for the mITT groups provided by the sponsor in the clinical study reports for studies 1056 and 1057 are summarized in Table 1. No differences between treatment groups with associated p-values less than 0.05 were noted in the demographic and baseline characteristics in the mITT groups for studies 1056 or 1057. As would be expected due to the random treatment assignment, balance among the treatment groups in demographic and baseline characteristics appears adequate to allow by-treatment group differences in post-randomization outcomes to be attributed to treatment effects and not an artifact of an imbalance in pre-randomization characteristics.

**Table 1: Demographic and Baseline Characteristics (mITT)**

Demographic/Baseline Characteristic		Study 1056			p-value <sup>1</sup>	Study 1057			p-value <sup>1</sup>
		Placebo N=275	Belimumab			Placebo N=287	Belimumab		
			1 mg/kg N=271	10 mg/kg N=273			1 mg/kg N=288	10 mg/kg N=290	
Gender	Female	252 (92%)	253 (93%)	259 (95%)	0.3	270 (94%)	271 (94%)	280 (97%)	0.3
	Male	23 (8%)	18 (7%)	14 (5%)		17 (6%)	17 (6%)	10 (3%)	
Race <sup>2</sup>	White/Caucasian	188 (68%)	192 (71%)	189 (69%)	0.9	82 (29%)	76 (26%)	71 (25%)	0.9
	Asian	11 (4%)	6 (2%)	11 (4%)		105 (37%)	106 (37%)	116 (40%)	
	Black/African American	39 (14%)	40 (15%)	39 (14%)		11 (4%)	8 (3%)	11 (4%)	
	Alaska Native or American Indian from North/Central/South America	36 (13%)	33 (12%)	34 (12%)		89 (31%)	98 (34%)	92 (32%)	
	Native Hawaiian or Other Pacific Islander	1 (<1%)	0 (0%)	0 (0%)		0 (0%)	0 (0%)	0 (0%)	
	Multiracial	2 (1%)	3 (1%)	3 (1%)		1 (<1%)	3 (1%)	1 (<1%)	
	Hispanic or Latino origin	Yes	55 (20%)	62 (23%)		56 (21%)	0.7	143 (50%)	
No	220 (80%)	209 (77%)	217 (79%)	144 (50%)	147 (51%)	154 (53%)			
Region and Country	USA/Canada	145 (53%)	155 (57%)	136 (50%)	0.6	NA			
	Western Europe/Israel	64 (23%)	63 (23%)	75 (28%)					
	Eastern Europe	36 (13%)	27 (10%)	30 (11%)					
	Americas excluding USA/Canada	30 (11%)	26 (10%)	32 (12%)					
Region and Country	East Europe	NA			33 (12%)	34 (12%)	31 (11%)	0.97	
	Latin America				145 (51%)	143 (50%)	140 (48%)		
	Asia				103 (36%)	106 (37%)	115 (40%)		
	Australia				6 (2%)	5 (2%)	4 (1%)		
Age (years)	Mean ± SD	40 ± 12	40 ± 11	41 ± 11	0.8	36 ± 12	35 ± 11	35 ± 11	0.4
	Min, Max	(18, 73)	(18, 70)	(18, 71)		(18, 69)	(18, 67)	(18, 71)	
Weight (kg)	Mean ± SD	72 ± 18	73 ± 18	74 ± 21	0.4	62 ± 12	61 ± 13	62 ± 13	0.5
	Min, Max	(43, 170)	(43, 135)	(45, 165)		(35, 128)	(36, 120)	(36, 129)	
Bilag organ domain involvement	At least 1A or 2B	187 (68%)	173 (64%)	160 (59%)	0.07	166 (58%)	166 (58%)	172 (59%)	0.9
	At least 1A	37 (14%)	38 (14%)	24 (9%)	0.1	52 (18%)	58 (20%)	54 (19%)	0.8
	At least 1A or 1B	258 (94%)	245 (90%)	251 (92%)	0.3	259 (90%)	255 (89%)	258 (89%)	0.8
	No A or B	17 (6%)	26 (10%)	22 (8%)	0.3	0 (0%)	0 (0%)	0 (0%)	
SELENA SLEDAI category	0 to 3	3 (1%)	5 (2%)	8 (3%)	0.7	1 (<1%)	4 (1%)	3 (1%)	0.2
	4 to 9	131 (48%)	122 (45%)	129 (47%)		128 (45%)	145 (50%)	127 (44%)	
	10 to 11	62 (23%)	72 (27%)	65 (24%)		75 (26%)	53 (18%)	72 (25%)	
	≥12	79 (29%)	72 (27%)	71 (26%)		83 (29%)	86 (30%)	88 (30%)	
PGA category	0-1	33 (12%)	39 (14%)	51 (19%)	0.3	43 (15%)	38 (13%)	32 (11%)	0.3
	>1-2.5	239 (87%)	230 (85%)	219 (80%)		243 (85%)	247 (86%)	256 (88%)	
	>2.5-3	3 (1%)	2 (0.7%)	3 (1%)		1 (<1%)	3 (1%)	2 (1%)	
SLICC Damage Index Score	Mean ± SD	0.99 ± 1.5	1.0 ± 1.4	0.94 ± 1.4	0.7	0.6 ± 0.9	0.6 ± 1.1	0.6 ± 1.0	0.8
SELENA SLEDAI score (stratification factor)	≤9	136 (50%)	128 (47%)	137 (50%)	0.8	128 (45%)	136 (47%)	134 (46%)	0.8
	≥10	139 (51%)	143 (53%)	136 (50%)		159 (55%)	152 (53%)	156 (54%)	
Proteinuria level (stratification factor)	<2 g/24hour	264 (96%)	261 (96%)	258 (95%)	0.6	264 (92%)	266 (92%)	269 (93%)	0.9
	≥2 g/24 hour	11 (4%)	10 (4%)	15 (6%)		23 (8%)	22 (8%)	21 (7%)	
Race (stratification factor)	AlA	74 (27%)	74 (27%)	72 (26%)	0.97	100 (35%)	106 (37%)	103 (36%)	0.9
	Other	201 (73%)	197 (73%)	201 (74%)		187 (65%)	182 (63%)	187 (65%)	

1. P-value for comparison across 3 treatment groups obtained from likelihood ratio or Fisher's exact test for categorical data or 1-way ANOVA for continuous data.

All primary efficacy analyses were conducted using the statistical procedures specified in the protocol. The primary efficacy endpoint, referred to as the SRI, is a composite endpoint that requires success on all of the following criteria at week 52 to be considered a success overall.

- $\geq 4$  point reduction from baseline in SELENA SLEDAI score, and
- No worsening (increase of  $<0.30$  points from baseline) in physician's global assessment (PGA), and
- No new BILAG A organ domain score or 2 new BILAG B organ domain scores compared with baseline at the time of assessment

Subjects who dropped out of the study before the week 52 visit were imputed as failures for the primary efficacy endpoint. In addition, the protocol specified that once a subject was randomized and received the first dose of study medication, adjustment to concurrent medications (add, eliminate, change dose level/frequency at certain times) was allowed as clinically required; however, certain changes required that the subject be coded as being a failure for purposes of the primary efficacy analysis and be discontinued from the study. These subjects are referred to as "medication failures".

The primary efficacy analysis for comparing each Belimumab dose to placebo was a logistic regression model with a term for treatment and adjusted for baseline stratification factors. A step-down sequential testing procedure was used to control for multiplicity in doses. The Belimumab 10 mg/kg treatment group was compared with the placebo group (two-sided  $\alpha=0.05$ ) first and if statistically significant, the belimumab 1 mg/kg treatment group was compared with the placebo group (two-sided  $\alpha=0.05$ ). The primary efficacy analysis was conducted in the modified intent to treat group (mITT) defined as all subjects randomized who received at least one dose of study medication.

The primary efficacy results for studies 1056 and 1057 are given in Table 2.

<b>Table 2: Primary Efficacy Analysis – Proportion of Subjects with Successful SRI Response (mITT)</b>						
	<b>Study 1056</b>			<b>Study 1057</b>		
	<b>Placebo N=275</b>	<b>Belimumab 1 mg/kg N=271</b>	<b>Belimumab 10 mg/kg N=273</b>	<b>Placebo N=287</b>	<b>Belimumab 1 mg/kg N=288</b>	<b>Belimumab 10 mg/kg N=290</b>
<b>Number SRI Responders (%)</b>	93 (34%)	110 (41%)	118 (43%)	125 (44%)	148 (51%)	167 (58%)
<b>Observed Diff. vs. placebo</b>		7%	9%		8%	14%
<b>OR (95% CI) vs. placebo<sup>1</sup></b>		1.3 (0.9, 1.9)	1.5 (1.1, 2.1)		1.6 (1.1, 2.2)	1.8 (1.3, 2.6)
<b>p-value for comparison to placebo<sup>1</sup></b>		0.1	0.02		0.01	0.0006
<b>Subcomponents</b>						
	<b>Study 1056</b>			<b>Study 1057</b>		
<b>Number with 4-Point Reduction in SELENA SLEDAI (%)</b>	98 (36%)	116 (43%)	128 (47%)	132 (46%)	153 (53%)	169 (58%)
<b>OR (95% CI) vs. placebo<sup>1</sup></b>		1.4 (0.96, 2)	1.6 (1.1, 2.3)		1.5 (1.1, 2.1)	1.7 (1.2, 2.4)
<b>p-value for comparison to placebo<sup>1</sup></b>		0.09	0.006		0.01	0.002
<b>Number with No Worsening in PGA (%)</b>	173 (63%)	197 (73%)	189 (69%)	199 (69%)	227 (79%)	231 (80%)
<b>OR (95% CI) vs. placebo<sup>2</sup></b>		1.6 (1.1, 2.3)	1.3 (0.9, 1.9)		1.7 (1.2, 2.5)	1.7 (1.2, 2.6)
<b>p-value for comparison to placebo<sup>2</sup></b>		0.01	0.1		0.008	0.005
<b>Number with No New 1A/2B BILAG Domain Score (%)</b>	179 (65%)	203 (75%)	189 (69%)	210 (73%)	226 (79%)	236 (81%)
<b>OR (95% CI) vs. placebo<sup>3</sup></b>		1.6 (1.1, 2.4)	1.2 (0.8, 1.7)		1.4 (0.9, 2.0)	1.6 (1.1, 2.4)
<b>p-value for comparison to placebo<sup>3</sup></b>		0.01	0.3		0.1	0.02

1. OR (95% CI) and p-value were from logistic regression for the comparison between each Belimumab dose and placebo with covariates including baseline SELENA SLEDAI ( $\leq 9$  vs  $\geq 10$ ), baseline proteinuria level ( $<2$  g/24 hour equivalent) and race (AIA vs other)

2. OR (95% CI) and p-value were from logistic regression for the comparison between each Belimumab dose and placebo with covariates as in footnote 1 and baseline PGA

3. OR (95% CI) and p-value were from logistic regression for the comparison between each Belimumab dose and placebo with covariates as in footnote 1 and baseline BILAG domain involvement (at least 1A/2B vs at most 1B)

Source: Sponsor analyses (Table 7-1 clinical study reports)

While the magnitude of the difference between each Belimumab group and placebo were fairly small, in each study, the Belimumab 10 mg/kg had a statistically significantly higher SRI success rate than the placebo group ( $p=0.02$  and  $p=0.0006$  for studies 1056 and 1057, respectively). A statistically higher rate of SRI response for the Belimumab 1 mg/kg group as compared to placebo was demonstrated for only study 1057 ( $p=0.02$ ). Reviewer analyses indicate that these conclusions are consistent even in the face of slight variations in the logistic regression models employed.

The results from the analyses of the subcomponents of the SRI were generally consistent with those of the primary analysis. The proportions of subjects achieving success for each of the subcomponents of the SRI were numerically higher in the Belimumab groups than the placebo group in each study.

Table 3 provides the reasons subjects failed to achieve a positive SRI response in studies 1056 and 1057. Note that the categories provided are mutually exclusive and mutually exhaustive. The proportions of subjects who dropped out are approximately 16% in study 1056 and 12% in study 1057 and are fairly balanced across treatment groups within each study thus the impact of imputing dropouts as failures on the treatment effect in the primary analysis should be small. However, unlike dropouts, “medication failures” are not balanced across treatment groups (17%, 9%, and 10% for placebo, 1 mg/kg Belimumab, and 10 mg/kg Belimumab respectively in study 1056 and 11%, 7%, and 6% for the same in study 1057). Since medication failures are more frequent in the placebo groups than the Belimumab groups, imputing medication failures as efficacy failures could bias the treatment effect in the primary efficacy endpoint in favor of Belimumab (unless these subjects would truly have been unable to achieve success on the primary endpoint had they not taken the prohibited medication).

<b>Table 3: Reasons for Not Achieving Successful SRI Response (mITT)</b>						
<b>Reason for Failure</b>	<b>Study 1056</b>			<b>Study 1057</b>		
	<b>Placebo N=275</b>	<b>Belimumab 1 mg/kg N=271</b>	<b>Belimumab 10 mg/kg N=273</b>	<b>Placebo N=287</b>	<b>Belimumab 1 mg/kg N=288</b>	<b>Belimumab 10 mg/kg N=290</b>
<b>Medication Failure</b>	47 (17%)	24 (9%)	27 (10%)	30 (10%)	21 (7%)	18 (6%)
<b>Drop out (and not medication failure)</b>	43 (16%)	40 (15%)	46 (17%)	38 (13%)	34 (12%)	31 (11%)
<b>Failed to satisfy <math>\geq 1</math> component of primary endpoint (and not medication failure or drop out)</b>	92 (33%)	97 (36%)	82 (30%)	94 (33%)	85 (30%)	74 (26%)

Source: Sponsor analyses, with modification (Table 7-3 clinical study reports)

Analyses of two of the numerous secondary efficacy endpoints examined in these studies (i.e., prednisone reduction by  $\geq 25\%$  from baseline to  $\leq 7.5$  mg/day during weeks 40 through 52 and severe flares) are provided in Tables 4 and 5. These two endpoints were selected by the FDA medical team as being of particular interest for evaluation of the efficacy of Belimumab. In considering these analyses, the reader should be cautioned that no multiplicity correction was planned for in the protocol for the secondary endpoints or applied here and therefore these hypothesis tests should be interpreted with caution as the

probability of at least one type I error occurring is increased beyond the usual 0.05 due to the examination of a large number of secondary endpoints.

Analyses of the proportion of subjects who reduced their prednisone by  $\geq 25\%$  from baseline to  $\leq 7.5$  mg/day during weeks 40 through 52 are provided in Table 4. Note that this analysis excludes subjects who were not receiving at least 7.5 mg/day of prednisone at baseline. Approximately half of the subjects in study 1056 and approximately 70% of the subjects in study 1057 were receiving at least 7.5 mg/day of prednisone and thus are included in these analyses. Of these subjects, the proportion of subjects who reduced their average prednisone dose by at least 25% to  $\leq 7.5$  mg/day during Weeks 40 through 52 were not consistently significantly different for either Belimumab dose relative to placebo in both studies. In study 1056, neither Belimumab dose group to placebo comparison was associated with a nominal p-value less than 0.05. In study 1057, only the Belimumab 1 mg/kg to placebo comparison resulted in a nominal p-value less than 0.05.

<b>Table 4: Secondary Efficacy Analysis – Prednisone Reduction by <math>\geq 25\%</math> from Baseline to <math>\leq 7.5</math> mg/day During Weeks 40 through 52 (subset of mITT)</b>						
	<b>Study 1056</b>			<b>Study 1057</b>		
	<b>Placebo N=275</b>	<b>Belimumab 1 mg/kg N=271</b>	<b>Belimumab 10 mg/kg N=273</b>	<b>Placebo N=287</b>	<b>Belimumab 1 mg/kg N=288</b>	<b>Belimumab 10 mg/kg N=290</b>
<b>Number of Subjects with baseline prednisone &gt; 7.5 mg/day (%)</b>	126 (46%)	130 (48%)	120 (44%)	192 (67%)	204 (71%)	204 (70%)
<b>Number of Subjects with Prednisone Reduction <math>\geq 25\%</math> from baseline to <math>\leq 7.5</math> mg/day</b>	16 (13%)	25 (19%)	20 (17%)	23 (12%)	42 (21%)	38 (19%)
<b>Odds Ratio (95% CI) vs. placebo<sup>1</sup></b>		1.6 (0.8, 3.1)	1.3 (0.6, 2.6)		1.9 (1.1, 3.3)	1.8 (1.0, 3.1)
<b>p-value<sup>1</sup></b>		0.2	0.5		0.03	0.053

1. From logistic regression for the comparison between each Belimumab dose and placebo with covariates, including baseline prednisone level and the stratification factors.

Source: Sponsor analyses (Table 7-24 for study 1056 and 7-15 for study 1057 clinical study reports)

Analyses of the time to first severe SLE flare over 52 weeks are provided in Table 5. The risk of experiencing a severe flare was not consistently significantly reduced for either Belimumab dose relative to placebo for both studies 1056 and 1057. In study 1056, comparison of the Belimumab 1 mg/kg dose to placebo for the risk of experiencing a severe flare was associated with a nominal p-value smaller than 0.05 and the comparison of the Belimumab 10 mg/kg dose to placebo for the same was not ( $p=0.02$  and  $p=0.09$  for the 1 mg/kg and 10 mg/kg, respectively). In study 1057, the comparison of the Belimumab 10 mg/kg dose to placebo was associated with a nominal p-value smaller than 0.05 while the comparison of the Belimumab 1 mg/kg dose to placebo was not ( $p=0.1$  and  $p=0.0006$  for the 1 mg/kg and 10 mg/kg, respectively).

<b>Table 5: Secondary Efficacy Analysis –Severe Flares (mITT)</b>						
	<b>Study 1056</b>			<b>Study 1057</b>		
	<b>Placebo N=275</b>	<b>Belimumab 1 mg/kg N=271</b>	<b>Belimumab 10 mg/kg N=273</b>	<b>Placebo N=287</b>	<b>Belimumab 1 mg/kg N=288</b>	<b>Belimumab 10 mg/kg N=290</b>
<b>Number of Subjects with at least one flare over 52 weeks (%)</b>	67 (24%)	44 (16%)	48 (18%)	66 (23%)	51 (18%)	40 (14%)
<b>Hazard Ratio (95% CI) vs. placebo<sup>1</sup></b>		0.6 (0.4, 0.9)	0.7 (0.5, 1.1)		0.8 (0.5, 1.1)	0.6 (0.4, 0.8)
<b>p-value<sup>1</sup></b>		0.02	0.09		0.1	0.006

1. From Cox proportional hazards model for the comparison between each Belimumab dose and placebo, adjusted for baseline stratification factors.

Source: Sponsor analyses (Table 7-18 for study 1056 and 7-9 for study 1057 clinical study reports)

The SRI response at week 76 was at the recommendation of the FDA pre-specified by the sponsor as a major secondary endpoint for study 1056. The SRI response at week 76 is provided in Table 6. Although there are no statistically significant differences between either Belimumab dose and placebo at week 76, the week 76 results are actually very similar to the week 52 results with the possible exception that the dropout rate in the 10 mg/kg Belimumab group increases more than it does in the other treatment groups. This small change in the pattern of dropouts could explain the move from a statistically significant difference between Belimumab 10 mg/kg and placebo at week 52 to a nonsignificant difference of the same at week 76.

<b>Table 6: Secondary Efficacy Analysis – SRI Response at Week 76 (mITT)</b>			
	<b>Study 1056</b>		
	<b>Placebo N=275</b>	<b>Belimumab 1 mg/kg N=271</b>	<b>Belimumab 10 mg/kg N=273</b>
<b>Number SRI Responders (%)</b>	89 (32%)	106 (39%)	105 (39%)
<b>Observed Diff. vs. placebo</b>		7%	6%
<b>OR (95% CI) vs. placebo<sup>1</sup></b>		1.3 (0.9, 1.9)	1.3 (0.9, 1.9)
<b>p-value for comparison to placebo<sup>1</sup></b>		0.10	0.13
<b>Reasons for Not Achieving Successful SRI Response</b>			
<b>Medication Failure</b>	53 (19%)	30 (11%)	33 (12%)
<b>Drop out (and not medication failure)</b>	53 (19%)	50 (19%)	59 (22%)
<b>Failed to satisfy ≥1 component of primary endpoint (and not medication failure or drop out)</b>	80 (29%)	85 (31%)	76 (28%)

1. OR (95% CI) and p-value were from logistic regression for the comparison between each Belimumab dose and placebo with covariates including baseline SELENA SLEDAI ( $\leq 9$  vs  $\geq 10$ ), baseline proteinuria level ( $<2$  g/24 hour equivalent) and race (AIA vs other)

### 3.2 Evaluation of Safety

For the statistical review of safety, mortality was highlighted by the FDA medical team as important for this application and is thus commented upon in this review. This analysis pools data from studies 1056, 1057, and a controlled phase 2 study, study LBSL02. As shown in Table 7, the exposure-adjusted incidence rates for mortality are 0.4%, 0.7%, and 0.9% for placebo, Belimumab 1 mg/kg, and Belimumab 10 mg/kg, respectively. Pooling the Belimumab groups, this difference translates to a number needed to harm of 342 patient years with a wide 95% confidence interval from 167 to infinity (which corresponds to no increased risk in mortality for Belimumab).

<b>Table 7: Safety Analysis – Mortality (Double Blind Periods of Studies 1056, 1057, and LBSL02)</b>			
	<b>Placebo</b>	<b>Belimumab 1 mg/kg</b>	<b>Belimumab 10 mg/kg</b>
<b>Number of Deaths (exposure-adjusted incidence)</b>	3 (0.4%)	5 (0.7%)	6 (0.9%)

Source: Sponsor analyses (Table 10-13 Advisory Committee Briefing package) and reviewer analyses

#### 4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

The sponsor evaluated the consistency of the treatment effect on the primary efficacy endpoint across subgroups using logistic regression with main effects for treatment, subgroup, and treatment-by-subgroup interactions. The statistical significance of the interaction term indicates whether the treatment effect is different among the subgroups.

The prespecified subgroup analyses that were considered included the following.

- County region (USA/Canada, Americas excluding USA/Canada, Western Europe, Eastern Europe)
- Baseline C4 levels (normal/high vs. low)
- Baseline C3 levels (normal/high vs. low)
- Baseline average dose of steroids ( $\leq 7.5$  mg/day vs.  $> 7.5$  mg/day)
- Baseline anti-dsDNA ( $\geq 30$  IU/mL vs.  $< 30$  IU/mL)
- Baseline proteinuria level ( $< 2$  g/24 hour vs.  $\geq 2$  g/24 hour equivalent, stratification factor)
- Race (AIA vs. other, stratification factor)
- Baseline SELENA SLEDAI score ( $\leq 9$  vs.  $\geq 10$ , stratification factor)

In study 1056, the only pre-specified subgroup analysis with a significant treatment-by-subgroup interaction for both Belimumab groups versus placebo was the race stratification factor ( $p=0.03$  and  $p=0.009$  for 1 mg/kg and 10 mg/kg, respectively). The nature of this interaction suggests that there may be a reversal of the treatment effect (i.e., a qualitative interaction) in the AIA race category versus other race category. Response rates of subjects of AIA races were highest in the placebo group while response rates of subjects of other races were highest in the Belimumab groups. (Note that due to the locations where studies 1056 and 1057 took place, subgroup analyses by race in study 1057 cannot be used to either confirm or refute these results. AIA subjects in study 1056 were primarily blacks and American Indians from the US and Canada while the AIA subjects in study 1057 were primarily Latin Americans.) To further investigate the significant treatment-by-race interaction, the sponsor undertook exploratory subgroup analyses by race with a different categorization (black vs. white vs. Alaska Native or American Indian vs. other). This subgroup analysis also seemed to indicate that there may be a reversal of treatment effect in black subjects. A marginally significant treatment-by-subgroup interaction was also observed in study 1056 for country/region for the comparison of 10 mg/kg group vs. placebo ( $p=0.07$ ). However, this result may have been influenced by the interaction observed for race in that 94% of the population in the Americas excluding USA/Canada fell into the AIA stratum.

In study 1057, the only pre-specified subgroup analysis with a significant treatment-by-subgroup interaction for both Belimumab groups versus placebo was baseline SELENA SLEDAI score ( $p=0.04$  and  $p=0.03$  for 1 mg/kg and 10 mg/kg, respectively). The nature of the interaction suggests that response to treatment with Belimumab may be greater in subjects with more active disease at baseline (i.e., a quantitative interaction).

Table 8 provides analyses of the primary efficacy endpoint, SRI responder, for the four subgroups previously discussed, race (AIA vs. other), race (black vs. white vs. Alaska Native or American Indian vs. other), country/region, and baseline SELENA SLEDAI score. Analyses by age ( $\leq 45$  years vs.  $>45$  years to  $<65$  years, age  $\geq 65$  excluded from inferential analyses due to sparse data) and gender (male vs. female) are also provided in Table 8.

<b>Table 8: Subgroup Analyses – Proportion of Subjects with Successful SRI Response by Subgroup (mITT)</b>						
	<b>Study 1056</b>			<b>Study 1057</b>		
<b>Number of Responders (%) by Subgroup</b>	<b>Placebo N=275</b>	<b>Belimumab 1 mg/kg N=271</b>	<b>Belimumab 10 mg/kg N=273</b>	<b>Placebo N=287</b>	<b>Belimumab 1 mg/kg N=288</b>	<b>Belimumab 10 mg/kg N=290</b>
<b>Race (stratification factor)</b>						
African descent or indigenous-American descent (AIA)	36/74 (49%)	30/74 (41%)	29/72 (40%)	47/100 (47%)	59/106 (56%)	64/103 (62%)
Other	57/201 (28%)	80/197 (41%)	89/201 (44%)	78/187 (42%)	89/182 (49%)	103/187 (55%)
<b>Interaction p-value<sup>1</sup></b>		0.03	0.009		0.9	0.8
<b>Race (post-hoc defn)</b>						
White – Caucasian	56/188 (30%)	78/192 (41%)	86/189 (46%)	38/82 (46%)	47/76 (62%)	47/71 (66%)
Black – African American or African Heritage	15/39 (39%)	12/40 (30%)	13/39 (33%)	7/11 (64%)	3/8 (38%)	5/11 (46%)
Alaska Native or American Indian	21/36 (58%)	18/33 (55%)	16/34 (47%)	40/105 (38%)	42/106 (40%)	56/116 (48%)
Other	1/12 (8%)	2/6 (33%)	3/11 (27%)	40/89 (45%)	56/98 (57%)	59/92 (64%)
<b>Interaction p-value<sup>1</sup></b>		0.2	0.07		0.2	0.3
<b>Region</b>						
USA/Canada	46/145 (32%)	59/155 (38%)	47/136 (35%)	12/33 (36%)	21/34 (62%)	23/31 (74%)
Western Europe/Israel	15/64 (23%)	25/63 (40%)	38/75 (51%)	71/145 (49%)	85/143 (59%)	85/140 (61%)
Western Europe	15/36 (42%)	11/27 (41%)	16/30 (53%)	40/103 (39%)	42/106 (40%)	56/115 (49%)
Americas excluding USA/Canada	17/30 (57%)	15/26 (58%)	17/32 (53%)	2/6 (33%)	0/5 (0%)	3/4 (75%)
<b>Interaction p-value<sup>1</sup></b>		0.6	0.07		0.4	0.2
<b>Baseline SELENA SLEDAI score (stratification factor)</b>						
$\leq 9$ points	39/134 (29%)	39/127 (31%)	45/137 (33%)	47/129 (36%)	55/149 (37%)	53/130 (41%)
$\geq 10$ points	54/141 (38%)	71/144 (49%)	73/136 (54%)	78/158 (49%)	93/139 (67%)	114/160 (71%)
<b>Interaction p-value<sup>1</sup></b>		0.3	0.2		0.04	0.03
<b>Age</b>						
$\leq 45$ years	65/189 (34%)	76/184 (41%)	80/178 (45%)	99/225 (44%)	117/236 (50%)	139/236 (59%)
$>45$ to $<65$ years	25/77 (33%)	33/83 (40%)	36/92 (39%)	25/57 (44%)	31/48 (65%)	27/52 (52%)
$\geq 65$ years <sup>2</sup>	3/9 (33%)	1/4 (25%)	2/3 (67%)	1/5 (20%)	0/4 (0%)	1/2 (50%)
<b>Interaction p-value<sup>1</sup></b>		0.9	0.7		0.4	0.7
<b>Gender</b>						
Male	8/23 (35%)	7/18 (39%)	6/14 (43%)	7/17 (41%)	8/17 (47%)	7/10 (70%)
Female	85/252 (34%)	103/253 (41%)	112/259 (43%)	118/270 (44%)	140/271 (52%)	160/280 (57%)
<b>Interaction p-value<sup>1</sup></b>		0.9	0.9		0.9	0.4

1. For treatment-by-subgroup interaction term from logistic regression.

2. Category excluded from logistic regression analysis due to sparse data.

Source: Sponsor analyses (Tables 7-5 and 7-6 clinical study reports)

To further explore the statistically significant qualitative interaction for race observed for each Belimumab doses versus placebo in study 1056, Table 9 contains the primary efficacy analysis, including the reasons for failure, within the AIA stratum. The patterns in the rates of medication

failure within the AIA stratum are similar to that of the overall group. However, the patterns in the dropout rates are not. Within the AIA stratum, there is a higher rate of dropouts in the 10 mg/kg Belimumab group than in the other groups. This disproportionately high rate of dropouts in the 10 mg/kg group could largely explain the reversal in the treatment effect in this stratum. This pattern of dropouts (i.e., an increasing rate of dropouts with increasing dose) is consistent with that of a drug with an undesirable or toxic effect at higher doses.

<b>Table 9: Primary Efficacy Analysis – Proportion of Subjects with Successful SRI Response within the AIA Stratum (mITT in AIA Stratum)</b>						
	<b>Study 1056</b>			<b>Study 1057</b>		
	<b>Placebo N=74</b>	<b>Belimumab 1 mg/kg N=74</b>	<b>Belimumab 10 mg/kg N=72</b>	<b>Placebo N=100</b>	<b>Belimumab 1 mg/kg N=106</b>	<b>Belimumab 10 mg/kg N=103</b>
<b>Number SRI Responders (%)</b>	36 (49%)	30 (41%)	29 (40%)	47 (47%)	59 (56%)	64 (62%)
<b>Observed Diff. vs. placebo</b>		-8%	-8%		9%	15%
<b>OR (95% CI) vs. placebo<sup>1</sup></b>		0.7 (0.4, 1.4)	0.7 (0.4, 1.4)		1.4 (0.8, 2.5)	1.9 (1.1, 3.2)
<b>p-value for comparison to placebo<sup>1</sup></b>		0.3	0.3		0.2	0.03
<b>Reasons for Not Achieving Successful SRI Response (mITT in AIA Stratum)</b>						
<b>Medication Failure</b>	13 (18%)	14 (19%)	11 (15%)	13 (13%)	9 (8%)	2 (2%)
<b>Drop out (and not medication failure)</b>	12 (16%)	14 (19%)	19 (26%)	15 (15%)	15 (14%)	15 (15%)
<b>Failed to satisfy <math>\geq 1</math> component of primary endpoint (and not medication failure or drop out)</b>	13 (18%)	16 (22%)	13 (18%)	25 (25%)	23 (22%)	22 (21%)

Source: Reviewer Analyses

The reasons for dropout within the AIA stratum are shown in Table 10. The rates of dropouts due to subject request, AE, and lost to follow-up are numerically higher for the 10 mg/kg Belimumab group than the other groups.

<b>Table 10: Reasons for Dropout in AIA Stratum (mITT within AIA Stratum)</b>			
	<b>Study 1056</b>		
	<b>Placebo N=74</b>	<b>Belimumab 1 mg/kg N=74</b>	<b>Belimumab 10 mg/kg N=72</b>
<b>Dropouts (and not a medication failure)</b>			
<b>Subject Request</b>	3 (4%)	4 (5%)	6 (8%)
<b>Adverse Event</b>	3 (4%)	3 (4%)	7 (10%)
<b>Lack of Efficacy</b>	4 (5%)	3 (4%)	3 (4%)
<b>Lack of Compliance</b>	1 (1%)	1 (1%)	0 (0%)
<b>Lost to Follow-up</b>	1 (1%)	1 (1%)	3 (4%)
<b>Other</b>	0 (0%)	2 (3%)	0 (0%)

Source: Reviewer Analyses

## 5. SUMMARY AND CONCLUSIONS

### 5.1 Statistical Issues and Collective Evidence

The following statistical issues and their impact have been described in the context of the review. Please refer to the specified section for details. Issues that are of particular consequence are shown below in boldface type.

- In each study, approximately 20% of randomized subjects discontinued treatment before week 52. The most common reasons for early treatment discontinuation were subject request, adverse event, and lack of efficacy. Many but not all of these subjects also discontinued the study. Considering subjects who discontinued the study as failures for the primary efficacy analysis is likely a fair representation of the efficacy in these subjects in that the subjects' reasons for withdrawal from treatment indicate the study treatment could not be tolerated in exchange for whatever efficacy may have been being achieved. Therefore, the primary efficacy results in the mITT group likely remain reliable despite the fairly high early treatment discontinuation rate.
- No significant differences between treatment groups in the demographic and baseline characteristics in the mITT groups for studies 1056 or 1057 were noted. As would be expected due to the random treatment assignment, balance among the treatment groups in demographic and baseline characteristics appears adequate to allow by-treatment group differences in post-randomization outcomes to be attributed to treatment effects and not an artifact of an imbalance in pre-randomization characteristics.
- **While the magnitude of the differences between treatment groups were fairly small (observed difference between Belimumab 10 mg/kg and placebo of 9% and 14% for studies 1056 and 1057, respectively), in each study, the Belimumab 10 mg/kg had a statistically significantly higher SRI success rate than the placebo group ( $p=0.02$  and  $p=0.0006$  for studies 1056 and 1057, respectively). A statistically higher rate of SRI response for the Belimumab 1 mg/kg group as compared to placebo was demonstrated for only study 1057 ( $p=0.02$ ). Reveiwer analyses indicate that these conclusions are consistent in the face of slight variations in the logistic regression models employed.**
- The results from the analyses of the subcomponents of the SRI were generally consistent with those of the primary analysis. The proportions of subjects achieving success for each of the subcomponents of the SRI were numerically higher in the Belimumab groups than the placebo group in each study.
- Subjects who dropped out of the study early were considered failures for the primary efficacy analysis. The proportions of subjects who dropped out are approximately 16% in study 1056 and 12% in study 1057 and are fairly balanced across treatment groups within each study thus the impact of imputing dropouts as failures on the treatment effect in the primary analysis should be small.
- **Subjects who were “medication failures” were considered failures for the primary efficacy analysis. The rates of “medication failures” are not balanced across treatment groups (17%, 9%, and 10% for placebo, 1 mg/kg Belimumab, and 10 mg/kg Belimumab respectively in study 1056 and 11%, 7%, and 6% for the same in study 1057). Since medication failures are more frequent in the placebo groups than**

**the Belimumab groups, imputing medication failures as efficacy failures could bias the treatment effect in the primary efficacy endpoint in favor of Belimumab (unless these subjects would truly have been unable to achieve success on the primary endpoint had they not taken the prohibited medication).**

- No multiplicity correction was planned for in the protocol for the secondary endpoints; however, two of the numerous secondary efficacy endpoints examined in these studies (i.e., prednisone reduction by  $\geq 25\%$  from baseline to  $\leq 7.5$  mg/day during weeks 40 through 52 and severe flares) were of particular interest to the FDA medical team for evaluation of the efficacy of Belimumab. The proportion of subjects who reduced their average prednisone dose by at least 25% to  $\leq 7.5$  mg/day during Weeks 40 through 52 were not consistently significantly different for either Belimumab dose relative to placebo in both studies. The risk of experiencing a severe flare was not consistently significantly reduced for either Belimumab dose relative to placebo for both studies 1056 and 1057.
- The SRI response at week 76 was at the recommendation of the FDA pre-specified by the sponsor as a major secondary endpoint for study 1056. There were no statistically significant differences between either Belimumab dose and placebo at week 76.
- **In study 1056, a significant treatment-by-race interaction for both Belimumab groups versus placebo suggest that there may be a reversal of the treatment effect (i.e., a qualitative interaction) in the AIA race category versus other race category ( $p=0.03$  and  $p=0.009$  for 1 mg/kg and 10 mg/kg, respectively). Response rates of subjects of AIA races were highest in the placebo group while response rates of subjects of other races were highest in the Belimumab groups. Exploratory subgroup analyses by race with a different categorization (black vs. white vs. Alaska Native or American Indian vs. other) also indicate that there may be a reversal of treatment effect in black subjects. This reversal of treatment effect in AIA subjects and/or black subjects could be due to the disproportionately high rate of dropouts in the 10 mg/kg group. This pattern of dropouts (i.e., an increasing rate of dropouts with increasing dose) is consistent with that of a drug with an undesirable or toxic effect at higher doses. Within the AIA stratum, the rates of dropouts due to subject request, AE, and lost to follow-up are numerically higher for the 10 mg/kg Belimumab group than the other groups.**

## 5.2 Conclusions and Recommendations

Studies 1056 and 1057 adequately demonstrate that the proportion of subjects achieving SRI response at week 52 is higher with Belimumab 10 mg/kg than placebo. However, the magnitudes of the treatment effects observed in these studies were small and the results were somewhat dependent on the handling of missing data due to “medication failure”. Clinical interpretation regarding the importance of these two factors is needed. Subgroup efficacy analyses suggest that Belimumab 10 mg/kg could be harmful in black subjects. Labeling clearly communicating this data and further investigation of the effect of Belimumab in black subjects is needed.

Numerical results for the Belimumab 1 mg/kg to placebo comparison were similar to the 10 mg/kg Belimumab to placebo comparison but did not reach statistical significance in both studies.

Analyses of secondary efficacy endpoints (i.e., prednisone reduction by  $\geq 25\%$  from baseline to  $\leq 7.5$  mg/day during weeks 40 through 52 and severe flares) and the SRI response at 76 weeks were not consistently significantly better for either Belimumab dose relative to placebo in both studies 1056 and 1057.

## STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

**BLA Number: STN125370**

**Applicant: HGS**

**Stamp Date: 6/9/10**

**Drug Name: Belimumab**

**NDA/BLA Type: priority**

On **initial** overview of the NDA/BLA application for RTF:

	<b>Content Parameter</b>	<b>Yes</b>	<b>No</b>	<b>NA</b>	<b>Comments</b>
1	Index is sufficient to locate necessary reports, tables, data, etc.	X			
2	ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)	X			
3	Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated (if applicable).	X			
4	Data sets in EDR are accessible and do they conform to applicable guidances (e.g., existence of define.pdf file for data sets).	X			

**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. **None**

<b>Content Parameter (possible review concerns for 74-day letter)</b>	<b>Yes</b>	<b>No</b>	<b>NA</b>	<b>Comment</b>
Designs utilized are appropriate for the indications requested.				
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.				
Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available.				
Appropriate references for novel statistical methodology (if present) are included.				
Safety data organized to permit analyses across clinical trials in the NDA/BLA.				
Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate.				

**STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA**

[Redacted] (b) (6)

7/21/10

Reviewing Statistician

Date

[Redacted] (b) (6)

7/21/10

Supervisor/Team Leader

Date