Statistical Review and Evaluation

NDA 21-239
Name of Drug: GL701
Applicant: Genelabs Technologis
Documents Reviewed: Statistical Section (Vol.18-Vol.102 of NDA 21-239) Received by CDER on 06/01/2000
Medical Reviewer: Kent Johnson, M.D.
Statistical Reviewer: Laura Lu, Ph.D.
Date of Review:

I. Introduction

NDA21-239 has been submitted for approval of GL701 for treatment of systemic lupus erythematosus (SLE) in women. Two placebo controlled pivotal studies were conducted in US: Study 94-01 (phase II/III) and Study 95-02 (phase III). Reports for the following supportive clinical studies were also submitted: An open label uncontrolled safety study (95-01), a foreign study (Study 96-01) conducted in Taiwan, two small studies (<28 patients) conducted in Stanford University (one open label and one double blinded PK/clinical study). This review focuses on the efficacy evaluation of the two pivotal studies.

II. Study Protocols

II.1 Study 94-01

This study was designed as a double-blind, randomized, placebo-controlled, parallel group trial to evaluate GL701 100 and 200 mg/day versus placebo in female patients with mild to moderate prednisone-dependent systemic lupus erythematosus (SLE). The objective of this study was to determine whether GL701 100 or 200 mg/day would allow tapering of prednisone (a steroid) use in patients with steroid-dependent SLE while maintaining stable SLE disease activity.

This study included women with mild to moderate systemic lupus erythematosus requiring chronic treatment with prednisone dose of $\geq$10 and $\leq$30 mg/day and either a) in the last 12 months attempted to taper prednisone dose but failed and had a stable prednisone dose for at least 6 weeks preceding the study, or b) in whom there had been no attempt to taper in the last 12 months and had been receiving a stable prednisone dose for at least 3 months preceding the study. Patients returned at monthly visits for up to 7 to 9 months. To evaluate the efficacy of GL701, the trial was designed so that prednisone was tapered in the face of stable or improving manifestations of SLE. Prednisone dose was to be reduced if disease activity was stabilized or improved (i.e., if SLE Disease Activity Index (SLEDAI) score was the same or decreased) from the prior monthly visit. If the SLEDAI score worsened (increased) from the prior monthly visit, the daily dose of prednisone could be increased at the investigator's discretion.

This study included two primary efficacy variables. The first one was responder rate. A responder was defined as a patient with the achievement of a decrease in prednisone dose to 7.5 mg/day or less sustained for no less than three consecutive scheduled visits, including the termination visit (i.e., two consecutive months), on or after Visit 7. The second primary variable was percent decrease in prednisone dose determined by comparing the prescribed prednisone (or steroid equivalent) dose at Baseline (Qualifying Visit) and the last visit prednisone dose using the physician prescribed prednisone dose recorded on the Medication Record Form.

The secondary efficacy variables included 1) change from baseline in SLEDAI, 2) change from baseline in quality of life assessment by SF-36, 3) change from baseline in Krupp Fatigue Severity Score (KFSS), 4) change from baseline in global assessment of disease activity by physician, and 5) change from baseline in global assessment of disease activity by patient assessed percent reduction achieved in daily prednisone dose.

All analyses were performed as intent-to-treat analyses. For each efficacy variable, the intent-to-treat analysis only included patients randomized to treatment that had a baseline measurement, received at least one dose of study drug, and had at least one post-baseline measurement.

The **proportion of responders** was analyzed using logistic regression with treatment as a factor. Baseline variables which attain a 0.05 significance level for association with treatment assignment may be included as covariates. A subsidiary analysis was proposed (Amendment 5) adding baseline SLEDAI and treatment interaction to the model.
Percentage reduction of prednisone dosage (from baseline) was analyzed by one-way ANOVA (analysis of variance) with treatment as a factor. Baseline variables which attain a 0.05 significance level for association with treatment assignment may be included as covariates.

All secondary efficacy variables were to be analyzed by means of a one-way analysis of covariance model with treatment as a factor and baseline (Qualifying Visit) as a covariate. Treatment-by-baseline interaction was to be included in the model.
Bonferroni's method for adjustment for multiple comparisons was to be used for the comparisons of GL701 100 mg/day vs. placebo and GL701 200 mg/day vs. placebo.

A sample size of 190 was initially planned to allow for 168 patients to complete the study. An interim analysis was conducted to adjust the sample size to ensure adequate power to detect treatment effects upon the second primary efficacy variable. The statistical methodology of EM algorithm was used to determine sample size.
Based on the result of the interim analysis, the sample size was not changed. The interim analysis used no treatment code and relative efficacy information. Therefore, type I error rate was inflated minimally (in the order of 10^{-3}) and no adjustment was done.

II.2 Study 95-02

This was a Phase III, multi-center, randomized, parallel group, double-blind, placebo-controlled study in female patients with active SLE. Patients were randomized to receive 200 mg/day GL701 or placebo. The primary objective of this study was to demonstrate improvement in the disease and or its symptoms in women with active SLE.

Patients were treated for 52 weeks and remained on the same blinded treatment for the duration of the study. Patients were required to visit the clinic every 13 weeks. The primary efficacy variable was responder rate. A responder was defined as a patient who satisfies the following conditions: (1) improvement or stabilization in all disease activity (i.e., systemic lupus activity measure (SLAM), SLEDAI) and constitutional symptom assessments (i.e. KFSS, Patient VAS), i.e., post baseline weighted (by time interval) means of SLAM, SLEDAI, KFSS and Patient VAS scores were either the same or less than the baseline scores and (2) no clinical deterioration. In a later (than the original protocol) submitted 'Statistical Analysis Plan', the sponsor redefined 'improvement and stabilization' by the following window definition: (1) weighted average change from baseline for SLAM is less than 1; for SLEDAI less than 0.5; for KFSS less than 0.5; for Patient VAS less than 10; and (2) no clinical deterioration. The 'Statistical Analysis Plan' was submitted after 86% of all the randomized patients had finished study.

Secondary efficacy variables included change from baseline in SLAM, SLEDAI, Patient's VAS, KFSS, Physician's VAS, SF-36 Mental Component Summary (MCS), SF-36 Physical Component Summary (PCS), time to first clinical deterioration and DEXA scan summary. In the later submitted 'Statistical Analysis Plan', time to flare was added as a secondary endpoint.

In the original protocol, proportion of responders was to be analyzed by a logistic regression model with treatment and center as factors. Covariates attaining 0.05 significance level for association with treatment assignment was to be included in the model, with eight covariates specified in the protocol (race, cytotoxic use, prednisone use, menopausal status, and baseline SLEDAI/SLAM /KFSS /PG). In the later submitted 'Statistical Analysis Plan', treatment was the only factor included in the logistic regression model, and center was dropped from the model due to many small centers. An additional analysis for proportion of responders was also proposed in the 'Statistical Analysis Plan' with SLEDAI > 2 (yes/no), baseline prednisone dose > 0 mg (yes/no), menopausal status (pre/other) as factors in addition to treatment, and any imbalanced baseline variable which attains a 0.05 significance level for association with treatment assignment included in the logistic regression model.

In the original protocol, time to first clinical deterioration was to be analyzed by using Cox regression model with treatment, center and the baseline variables listed above as factors. In the later submitted 'Statistical Analysis Plan', treatment was the only factor included in the Cox model for time to deterioration and time to flare. In the original

protocol, weighted mean changes from baseline in SLAM, SLEDAI, patient's VAS, KFSS, physician's VAS, SF-36 MCS, SF-36 PCS, Systemic Lupus International Cooperating Clinics (SLICC) were to be analyzed in a two-way analysis of covariance model with treatment and trial center as factors, and baseline as a covariate and treatment-by-baseline and treatment-by-center interactions included in the model. In the later submitted 'Statistical Analysis Plan', only treatment, baseline and treatment-by-baseline interaction were to be included in the model.

In the original protocol, the ITT population including all randomized patients were the primary analysis population for all endpoints. In Protocol Amendment #1, a subgroup analysis was proposed for patients with baseline SLEDAI>2 based on the result observed in Study 94-01. In the later submitted 'Statistical Analysis Plan', the sponsor changed the primary analysis population to a per-protocol population: patients who had either clinical deterioration or who had baseline measurements and at least one post-baseline measurement at the on-treatment visits for at least one of the variables SLAM, SLEDAI, patient's VAS, and KFSS with at least 60 days study medication.

Since there was no prior information regarding the responder rates, the original sample size of 300 randomized patients was not based on statistical calculations, but was mainly based on feasibility considerations. In Protocol Amendment #1, additional 50 patients were proposed with an extra inclusion criteria that the baseline SLEDAI score should be larger than 2. This decision was made to increase the power of analysis in the subgroup with baseline SLEDAI>2.

Handling of Missing value was not discussed in the original protocol. In the later submitted 'Statistical Analysis Plan', the sponsor provided following methods in dealing with missing value for the primary responder analysis:

'For each missing item in SLAM or SLEDAI at the on-treatment visit, the measurement of that item at the previous visit (on-treatment visit or Qualifying Visit) will be carried forward for the measurement of this missing item. For any missing item in SLAM or SLEDAI at either Screening or Qualifying Visit, the other non-missing measurement of that item will be used for this missing item. For any missing item in SLAM at both Screening and Qualifying Visits, that item score at the on-treatment visits will be treated as missing.

SLAM and SLEDAI scores for each visit will be calculated by the sum of the item scores. KFSS Fatigue score for each visit will be calculated by the average of non-missing item scores.

For each missing measurement in SLAM, SLEDAI, KFSS, Patient VAS at the on-treatment visit, the average of the measurements at the two nearby (before and after) on-treatment visits will be used for that missing measurement.'

Methods used for handling missing value for all efficacy endpoints are described in Tables c.1-c.3 in Appendix C.

III. Sponsor's Reports

III.1 Study 94-01 (ITT)

III.1.i Patient Disposition

A total of 191 patients were randomized to receive study drug. The dropout rates were numerically higher in the two GL701 groups than that in the placebo group (23.4% in placebo, 27.0% in GL701 100 mg and 26.6% in GL701 200 mg). The highest dropout rate due to lack of efficacy was in placebo group (10.9%) and the highest dropout rate due to adverse event was in GL701 200 mg (7.8%). Detailed patient disposition is displayed in Table 1 below. The survival curves for withdrawal due to lack of efficacy and adverse events are presented in Figures b.1 and b.2 in Appendix B.

Table 1. Patient Disposition

	Placebo	GL701 100mg	GL701 200mg
Enrolled	64 (100.0%)	63 (100.0%)	64 (100.0%)
Completer	49 (76.6%)	46 (73.0%)	47 (73.4%)
Dropouts	15 (23.4%)	17 (27.0%)	17 (26.6%)
Reasons for Discontinuation			
Lack of Efficacy	7 (10.9%)	6 (9.5%)	5 (7.8%)
Adverse Event	1 (1.6%)	1 (1.6%)	5 (7.8%)
Other	7 (10.9%)	10 (15.9%)	7 (10.9%)

III.1.ii Demographics and Baseline Characteristics

The study population of 191 patients consisted of all women, primarily Caucasian (60%) and African-American (26%). Patient demographics and baseline characteristics were numerically comparable (see Tables a.1 and a.2 in Appendix A).

III.1.iii Efficacy Endpoints

Primary Efficacy Variables

One of the two primary efficacy variables was the achievement of a decrease in prednisone dose to 7.5 mg/day or less sustained for no less than three consecutive scheduled visits including the termination visit (i.e., two consecutive months). Patients who achieved this sustained prednisone dose reduction were defined as responders. The proportion of responders was analyzed by logistic regression analysis with treatment as a factor (no other baseline variables were included as covariates since none was different between treatments with p-value<0.05). The detailed results of responder analysis given in Table 2 show that no statistically significant advantage was observed for GL701 groups vs. placebo. The results when including baseline SLEDAI and treatment interaction to the model were consistent with those in Table 2.

Table 2. Percent of Responders by Treatment Group

Group	Placebo (N=64)	GL701 100 mg (N=63)	GL701 200 mg (N=64)
Responder	40.6% (26/64)	44.4% (28/63)	54.7% (35/64)
Non-Responder	59.4% (38/64)	55.6% (35/63)	45.3% (29/64)
P-value* vs. Placebo		0.66	0.11

*: P-value by logistic regression with treatment as a factor

Reviewer's brief comment: *After Bonferroni adjustment for multiple comparison, the p-values in Table 2 above should be 1 and 0.22.*

The second primary efficacy variable was percent decrease in prednisone use, comparing final dose with baseline. Although no statistically significant difference were detected for each of the pairwise comparisons (active treatment vs. placebo), placebo showed more percent deduction in prednisone use than the GL701 groups. Detailed results for percent decrease in prednisone use are presented in Table 3 below.

Table 3. Mean Percent Change from Baseline to Last Visit in Prescribed Prednisone Dose

Treatment Group	Mean % Change (SD)	P-values* (vs. Placebo)
Placebo (N=64)	-35.8 (50)	
GL701 100 mg (N=63)	-13.7 (91)	0.094
GL701 200 mg (N=664)	-30.3 (74)	0.672

*: P-value by one-way ANOVA with treatment as a factor

Secondary Efficacy Variables

The secondary efficacy variables included change from baseline in the measurements of the following: SLEDAI, each of the eight systems in SF36, Krupp fatigue score, physician's global assessment, and patient's global assessment. Since the main purpose of the trial design was to reduce prednisone dose and maintain a consistent SLEDAI score, it would not be expected that the secondary variables show treatment differences. Each secondary efficacy variable was analyzed by means of a one-way analysis of covariance model with treatment as a factor and baseline as a covariate. Treatment-by-baseline interaction was included in the model. There were no statistically significant or clinically meaningful differences between treatment groups for changes in any of these variables from baseline (See Table a.3 in Appendix A).

Exploratory Subgroup Analyses

in the subgroup of patients with baseline SLEDAI > 2, a larger treatment effect was observed in terms of responder rate, but not in the mean percent change from baseline to last visit in prescribed prednisone (the second primary endpoint. The responder rate and percent change from baseline in prednisone dose in each treatment group within the subgroup is presented in Table 4 and Table 5 below. Note that the p-values in these tables

are nominal and can not be interpreted as level of significance since the exploratory nature of the subgroup analysis.

Table 4. Percent of Responders by Treatment Group in Patients with SLEDAI>2

Group	Placebo (N=45)	GL701 100 mg (N=47)	GL701 200 mg (N=45)
Responder	28.9% (13/45)	38.3% (18/47)	51.1% (23/45)
Non-Responder	71.1% (17/45)	61.7% (21/47)	48.9% (15/45)
P-value* (vs. Placebo)		0.339	0.031

*: P-value by logistic regression with treatment as a factor

Table 5. Mean Percent Change from Baseline to Last Visit in Prescribed Prednisone Dose in Patients with SLEDAI>2

Treatment Group	Mean % Change (SD)	P-values* (vs. Placebo)
Placebo (N=45)	-25.74 (54)	
GL701 100 mg (N=47)	-0.01 (101)	0.129
GL701 200 mg (N=45)	-21.96 (77)	0.788

*: P-value by one-way ANOVA with treatment as a factor

III.2 Study 95-02

III.2.i Patient Disposition

A total of 381 patients were randomized to receive study drugs. The dropout rates were 26.0% in placebo group and 34.4% in GL701 200 mg group. The percentages of dropout due to lack of efficacy and treatment related adverse events were both higher in the GL701 group than that in the placebo group (5.8% vs. 4.7%, 14.3% vs. 5.7%). Detailed patient disposition for the ITT population is displayed in Table 6 below. The survival curves for withdrawal due to lack of efficacy and adverse events are presented in Figures b.3 and b.4 in Appendix B.

Table 6. Patient Disposition

	Placebo	GL701
No. of Patients Randomized	192	189
No. of Patients Completed Study Drug	142 (73.9%)	124 (65.6%)
No of Early Terminations from Study Drug	50 (26.0%)	65 (34.4%)
Lack of Efficacy or Required Immunosuppression	9 (4.7 %)	11 (5.8%)
Possible treatment-related adverse event	11 (5.7%)	27 (14.3%)
Terminated for Reasons Related to Neither Safety Nor Efficacy	30 (15.6%)	27 (14.3%)

III.2.ii Demographics

The study population consisted of all women, primarily Caucasian (74%) and African-American (14%). Patient demographics and baseline characteristics were numerically comparable (see Tables a.4 and a.5 in Appendix A for detailed demographics for the ITT population).

III.2.iii Efficacy Endpoints

Results for ITT population are reported below. Results for per-protocol population (patients who were on the study drug for more than 60 days and had measurements of SLE scores or other data beyond 60 days) are reported in Tables a.6-a.11 in Appendix A.

Reviewer's brief comment: *In general, GL701 showed more numerical advantage than placebo in per-protocol analysis compared with ITT analysis. Please see reviewer's further comment about ITT vs. per-protocol population in Section IV.2.i.*

Primary Efficacy Variable

When a responder was defined without a window (the definition in the original protocol), the responder rates for the GL701 group and the placebo group were 30.7% and 27.1%. When a responder was defined with a window (the definition in the 'Statistical Analysis Plan'), the responder rates for the GL701 group and the placebo group were 51.3% and 42.2%, respectively. The p-values listed in Table 7 below are obtained by logistic regression with only treatment as a factor. The results from analysis with covariates (SLEDAI > 2 (yes/no), baseline prednisone dose > 0 mg (yes/no), menopausal status (pre/other)) included were consistent with that in Table 7 in terms of level of statistical significance. Since the baseline variables are comparable among treatment groups, none of them were included in the logistic model. The p-values from analyses without/with windows were larger than .05 (0.4378 and 0.07 respectively).

Table 7. Percent of Responders by Treatment Group

	Placebo	GL701 200 mg	P-value*
	192	189	
Without Window			
Responder	52 (27.1%)	58 (30.7%)	0.4378
Non-Responder	140 (72.9%)	131 (69.3%)	
With Window			
Responder	81 (42.2%)	97 (51.3%)	0.0744
Non-Responder	111 (57.8%)	92 (48.7%)	

*: P-value by logistic regression with treatment as a factor

Reviewer's brief comment: *The result displayed in Table 7 above did not take into account treatment failures, i.e., if a patient dropped out due to lack of efficacy or adverse event, the patient is still counted as a responder as long as the patient's SLAM, SLEDAI, KFSS and Patient VAS satisfied the responder criteria. Please see reviewer's further comment in Section IV.2.iii. Since the window definition was not included in the original protocol (proposed at a time when 86% of the ITT patients had finished the study) and there was no clear rationale for this definition, robustness of the window definition will be examined. Please see reviewer's further comment in Section IV.2.ii.*

Secondary Efficacy Variables

Time to first definite flare were analyzed by log-rank test. No statistically significant difference was found between the treatment groups. Detailed results are presented in Table 8 below. The survival curve for first definite flare is presented in Figure b.5 in Appendix B.

Table 8. Survival Analysis for First Definite Flare

	Placebo (N=192)	GL701 200 mg (N=189)
Number of Patients Experiencing Definite Flare	57 (29.7%)	45 (23.8%)
P-value*		0.2657

*: P-value by log-rank test.

Reviewer's brief comment: *The pre-specified analysis for time to flare was Cox regression model instead of log-rank test. However, the result from Cox regression is consistent with that from log-rank test with p=0.2417 based on the reviewer's analysis.*

Time to clinical deterioration was analyzed by log-rank test. The percent of patients experience clinical deterioration were similar and no statistically significant difference was found between the treatment groups. Detailed results are presented in Table 9 below. The survival curve for Clinical Deterioration is presented in Figure b.6 in Appendix B.

Table 9. Survival Analysis for Clinical Deterioration

	Placebo (N=192)	GL701 200 mg (N=189)
Number of Patients Experiencing Clinical Deterioration	16 (8.3%)	16(8.5%)
P-value (GL701 vs. plavebo)		0.869

*: P-value by log-rank test.

Reviewer's brief comment: *The pre-specified analysis for time to clinical deterioration was Cox regression model instead of log-rank test. However, the result from Cox regression is consistent with that from log-rank test with p=0.8555 based on the reviewer's analysis.*

Means of change from baseline for scoring instruments were summarized in Table 10 below. Although GL701 did show numerical advantage in most of these secondary efficacy endpoints, no statistical significance was demonstrated in any of them.

Table 10. Change in Scoring Instruments from Baseline

Variable	Placebo	GL701 200 mg
SLEDAI	(N=178)	(N=178)

Mean Change from Baseline	-1.7	-2.2
Mean at Baseline (SD)	5.8 (4.3)	6.5 (4.3)
Patient VAS	(N=178)	(N=169)
Mean Change from Baseline	-4.5	-6.2
Mean at Baseline (SD)	55.4 (18.5)	55.2 (18.8)
Physician VAS	(N=178)	(N=169)
Mean Change from Baseline	-5.1	-5.6
Mean at Baseline (SD)	30.3 (13.5)	30.2 (13.8)
KFSS	(N=178)	(N=169)
Mean Change from Baseline	-0.4	-0.3
Mean at Baseline (SD)	5.6 (1.2)	5.5 (1.2)
SLAM	(N=178)	(N=170)
Mean Change from Baseline	-2.7	-3.1
Mean at Baseline (SD)	12.0 (3.0)	12.2 (2.8)
SLICC	(N=140)	(N=128)
Mean Change from Baseline	-0.1	-0.1
Mean at Baseline (SD)	1.3 (1.4)	1.3 (1.4)
SF36 – MCS	(N=175)	(N=166)
Mean Change from Baseline	1.8	2.6
Mean at Baseline (SD)	41.7 (11.8)	42.5 (10.2)
SF36 – PCS	(N=175)	(N=166)
Mean Change from Baseline	1.7	1.8
Mean at Baseline (SD)	30.3 (13.5)	30.2 (13.8)

Bone density loss was measured only at 8 out of 23 centers, and only on patients who had been on prednisone for at least 6 months. Thirty-seven (37) patients were included, 18 on DHEA and 19 on PLC. Summary results are presented in Table 11 below.

Table 11. Bone Density (gm/cm^2)

Treatment Group	Baseline Mean (SD)*	Last Visit Mean (SD)**	Percent Change (SD) From Baseline
Location:Hip			
Placebo (N = 19)	0.8735 (0.1194)	0.8721 (0.1206)	-0.16% (2.43%)
GL701 200 mg (N = 18)	0.8528 (0.1268)	0.8664 (0.1153)	2.08% (4.82%)
Location:Spine			
Placebo (N = 19)	0.9695 (0.1368)	0.9529 (0.1422)	-1.78% (3.04%)
GL701 200 mg (N = 18)	0.9447 (0.1422)	0.9595 (0.1374)	1.83% (4.10%)

* Baseline refers to the qualifying visit or within three days following the qualifying visit.
** Last visit refers to the last post-baseline measurement of an on-treatment visit.

Subgroup Analyses

In protocol amendment #1, the sponsor proposed subgroup analysis in patients with baseline SLEDAI>2. Patient disposition, results for primary and secondary efficacy endpoints are presented in Tables 12 to 16 below.

Reviewer's brief comment: *Consistent with the ITT population, GL701 group had more patients' withdrawal due to lack of efficacy and adverse events in this subgroup. Similar to the ITT population, GL701 showed numerical advantage over placebo in this subgroup in terms of responder rate, number of patients with definite flare, but not in terms of number of patients with clinical deterioration. P-values for all efficacy endpoints are larger than 0.05 except for the responder analysis with a window definition (p=0.017).*

Table 12. Patient Disposition in Patients with Baseline SLEDAI>2

	Placebo	GL701
No. of Patients Randomized	146	147
No. of Patients Completed Study Drug	105 (71.9%)	93 (63.3%)
No of Early Terminations from Study Drug	50 (28.1%)	65 (36.7%)
Lack of Efficacy or Required Immunosuppression	9 (6.2 %)	11 (7.5%)
Possible treatment-related adverse event	9 (6.2%)	20 (13.6%)
Terminated for Reasons Related to Neither Safety Nor Efficacy	23 (15.8%)	23 (15.7%)

Table 13. Percent of Responders by Treatment Group in Patients with Baseline SLEDAI>2

	Placebo	GL701 200 mg	P-value*
	146	147	
Without Window			
Responder	42 (28.8%)	55 (37.4%)	0.1166
Non-Responder	104 (71.2%)	92 (62.6%)	
With Window			
Responder	65 (44.5%)	86 (58.5%)	0.0170
Non-Responder	81 (55.5%)	61 (41.5%)	

*: P-value by logistic regression with treatment as a factor

Table 14. Survival Analysis for First Definite Flare in Patients with Baseline SLEDAI>2

	Placebo (N=146)	GL701 200 mg (N=147)
Number of Patients Experiencing Definite Flare	50 (34.2%)	36 (24.5%)
P-value (vs. placebo)		0.0967

*: P-value by log-rank test.

Table 15. Survival Analysis for Clinical Deterioration in Patients with Baseline SLEDAI>2

	Placebo (N=146)	GL701 200 mg (N=147)
Number of Patients Experiencing Clinical Deterioration	13 (8.9%)	15 (10.2%)
P-value (GL701 vs. placebo)		0.639

*: P-value by log-rank test.

Table 16. Mean Change in Scoring Instruments from Baseline in Patients with Baseline SLEDAI>2

Variable	Intent to Treat - Baseline SLEDAI > 2	
	Placebo (N=146)	GL701 200 mg (N=147)
SLEDAI	-2.5	-3.2
	(N=134)	(N=132)
Patient VAS	-3.0	-7.2
	(N=134)	(N=131)
Physician VAS	-4.3	-5.4
	(N=134)	(N=131)
KFSS	-0.3	-0.3
	(N=134)	(N=131)
SLAM	-2.7	-3.2
	(N=134)	(N=132)
SLICC	-0.1	-0.1
	(N=104)	(N=97)
SF36 – MCS	1.6	2.3
	(N=132)	(N=129)
SF36 – PCS	0.9	1.9
	(N=132)	(N=129)

IV. Reviewer's Comments

IV.1 Comments on Efficacy Results of Study 94-01

As presented in Tables 2 and 3, Study 94-01 did not demonstrate statistically significant advantage by the two primary endpoints: responder rate and percent of prednisone reduction. The responder rate for the GL701 groups were not significantly higher than that in the placebo group, and the mean percent of prednisone reduction in both GL701 groups were numerically lower than that in the placebo group.

In the post-hoc subgroup analysis for patients with baseline SLEDAI>2, the GL701 groups showed a larger numerical advantage over placebo than in the overall ITT population. Nonetheless, the mean percent of prednisone reduction in both GL701 groups were still numerically lower than that in the placebo group. Since this subgroup analysis was not pre-planned, the results was only used for hypothesis generating instead of efficacy confirmation. Due to the lack of advantage in mean percent of prednisone reduction, the result of the subgroup analysis did not provide a robust base for generating the hypothesis that GL701 is efficacious in patients with baseline SLEDAI>2.

IV.2 Comments on Efficacy Results in Study 95-02

IV.2.i ITT Population vs. Per-Protocol Population

The ITT population in Study 95-02 included all randomized patients, while the per-protocol population only included only the patients that had stayed in the trial for more than 60 days and with post-baseline measurements. From statistical point of view, ITT analysis preserves randomization which enables valid statistical inference.

The sponsor argued that, based on results from the Stanford University studies, the treatment effect required approximately a minimum of 2 months treatment, therefore, drop-outs within 60 days should not be treatment related. However, as presented in Table 17 below, among the 35 patients who withdrawn the study within the first 60 days, 12 (35%) of them were due to treatment related reasons. Among the 12 patients, 11 of them dropped out due to treatment related adverse events with 9 of them in the GL701 group. In ITT analysis, the 35 patients were treated as non-responders with the assumption that the dropouts were treatment related. While the ITT analysis may over-estimate the treatment risks, it counter-balances the per-protocol analysis which optimistically assessing treatment risks at early study stage.

Table 17. Patient Disposition in Patients Withdrawn within the First 60 Days of Treatment

	Placebo	GL701
No of Early Terminations within the first 60 Days	15	19
Lack of Efficacy or Required Immunosuppression	0	1
Possible treatment-related adverse event	3	8
Terminated for Reasons Related to Neither Safety Nor Efficacy	12	10

IV.2.ii Robustness of Responder Analysis with Window Definition

The window definition for a responder was proposed by the sponsor in a later submitted statistical analysis plan after 86% of the ITT population finished the study. The window for 'improvement and stabilization' for each score were: weighted average change from baseline for SLAM is less than 1; for KFSS less than 0.5; for Patient VAS less than 10; for SLEDAI less than 0.5. So 'worse in some extent' from baseline was allowed for a responder. This reviewer calculated the mean and range of the window margins over the baseline scores. As presented in Table 18 below, for SLAM, 1 unit can be 4.8%-25% (mean is 8.8%) of baseline for all ITT patients; for KFSS, 0.5 unit can be 7.1%-45% (mean is 9.6%) of baseline for all ITT patients; for Patient VAS, 10 unit can be 10.1%-500% (mean is 23.2%) of baseline for all ITT patients; for SLEDAI, 1 unit can be 2.1%-50% (mean is 10.9%) of baseline for all ITT patients when thirty eight zero (0) baseline scores are excluded. This window definition seems inadequate since a patient could be classified as a responder even the Patient VAS becomes 4 times worse than that at baseline as long as other scores are within the window margin.

Table 18. Window Margin in Terms of Percent of Baseline

Scores	Window Margin	Range of % of Baseline in ITT Patients	Mean of % of Baseline in ITT Patients	% of SD of Baseline in ITT Patients
SLAM	1	4.8%-25%	8.8%	34.7%
KFSS	0.5	7.1%-45%	9.6%	43.1%
Patient VAS	10	10.1%-500%	23.2%	53.7%
SLEDAI	0.5	2.1%-50%	10.9%	11.8%

This reviewer believes that a by-patient window is more appropriate for a by-patient endpoint (responder rate). This reviewer assessed the robustness of the responder analysis by defining by-patient windows according to percent change from baseline. For example, a -5% window definition for a responder is (1) weighted averages of SLAM, SLEDAI, KFSS and Patient VAS are not worse for more than 5% from baseline and (2) no clinical deterioration. A positive percent means improvement from baseline. The range of percentage explored is –70% to 50%.

Figure 1 displayed the responder rates of placebo and GL701 for ITT population versus the percent for window definition. The specific responder rates are given in Table a.12 in Appendix A. As shown in Figure 1 and Table a.12, GL701 had numerical advantage over placebo in term of responder rate at <1% windows, and the p-values were less than 0.05 at -15%, -10% and -3% windows. Therefore, if a responder is defined by >-3% windows (i.e., no worse than placebo than baseline by 3%), GL701 would not be significantly better than placebo at 0.05 level; If a responder is defined by >1% windows (i.e., improve from baseline by at least 1%), GL701 would lose numerical advantage over placebo. In general, GL701 would show larger numerical advantages over placebo if worsening from baseline is allowed for a responder definition than when the strict definition is used.

Figure 1. Responder Rate in ITT Patients

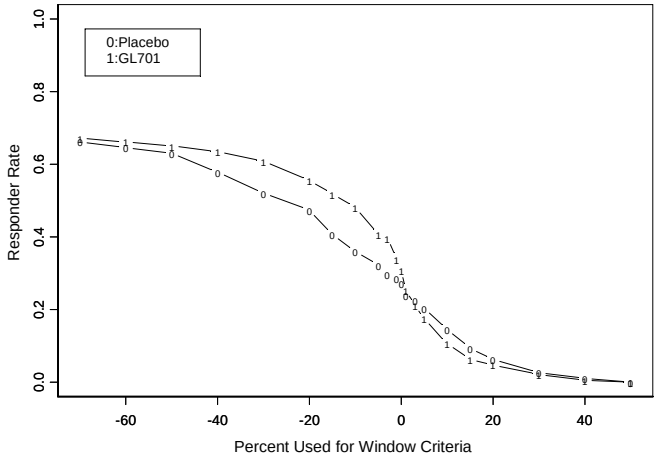
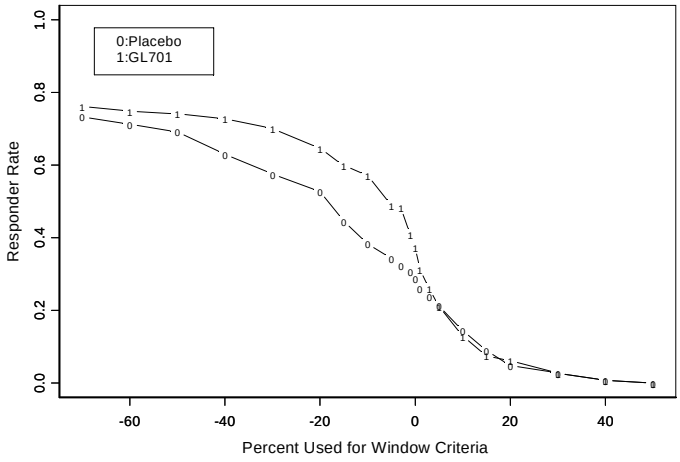


Figure 2 displayed the responder rates of placebo and GL701 for patients with baseline SLEDAI>2 along the percent for window definition. The specific responder rates were given in Table a.13 in Appendix A. In general, Figure 2 presented similar pattern to that in Figure 1, although the results were more favorable to GL701 over placebo in this subgroup than that in the overall ITT population (GL701 had numerical advantage over placebo in term of responder rate at <5% windows, and the p-values were less than 0.05 at -30% to -3% windows).

Figure 2. Responder Rate in ITT Patients with SLEDAI>2



IV.2.iii Influence of Patient Disposition to Responder Rates

As specified in the protocol, a patient's response was evaluated by weighted averages of change of SLAM, SLEDAI, KFSS from baseline while the patient was on treatment. A patient could be classified as a responder even if the patient terminated the study early. When a patient terminated study due to lack of efficacy (LOE) or adverse event (AE), the corresponding treatment should not be considered successful for the patient. Figures 3 and 4 below show that the drop-out rates due to LOE and AE were both higher in the GL701 than that in the placebo group, so the result of responder rate may bias against placebo when the early drop-outs due to treatment failure (LOE and AE) were not properly taken into account. This reviewer conducted a sensitivity analysis by treating early dropouts due to treatment failure as non-responders even when SLAM, SLEDAI, KFSS SLAM, SLEDAI, KFSS scores were stabilized or improved from baseline while the patient was on treatment. The results of this sensitivity analysis are compared with the sponsor's original results without window in Table 18 for all ITT patients and Table 19 for the subgroup with baseline SLEDAI>2. Table 18 and Table 19 show that the numerical advantages of GL701 over placebo in responder rates are less with the sensitivity analysis than with the original analysis.

Robustness of window definition is also assessed for the above sensitivity analysis. Figure 3 and Figure 4 present the responder rates in GL701 and placebo in the sensitivity analysis along the percent for window definition. Compared with the results displayed Figure 1 and Figure 2, the numerical advantage of GL701 mitigated with the sensitivity

analysis along all the percent for window definition, and there were no statistically significant advantage demonstrated for GL701 over placebo with any percentage window.

Table 18. Results Comparison between Sensitivity Analysis and Responder's Original Analysis in ITT Patients (Without Window)

	Placebo	GL701 200 mg	P-value**
	192	189	
Sponsor's Result			
Responder	52 (27.1%)	58 (30.7%)	0.438
Non-Responder	140 (72.9%)	131 (69.3%)	
Sensitivity Analysis*			
Responder	51 (26.6%)	53 (28.0%)	0.746
Non-Responder	141 (73.4%)	136 (72.0%)	

*: sensitivity analysis refers to the analysis with dropouts due to LOE and AE considered as non-responders

**: P-values are from Mantel-Haenszel Tests

Table 19. Results Comparison between Sensitivity Analysis and Responder's Original Analysis in Patients with Baseline SLEDAI>2 (Without Window)

	Placebo	GL701 200 mg	P-value**
	146	147	
Sponsor's Result			
Responder	42 (28.8%)	55 (37.4%)	0.117
Non-Responder	104 (71.2%)	92 (62.6%)	
Sensitivity Analysis*			
Responder	41 (28.1%)	50 (34.0%)	0.273
Non-Responder	105 (71.9%)	97 (66.0%)	

*: sensitivity analysis refers to the analysis with dropouts due to LOE and AE considered as non-responders

**: P-values are from Mantel-Haenszel Tests

Figure 3. Responder Rate in ITT Patients with Dropouts Due to LOE and AE Considered as Nonrresponders

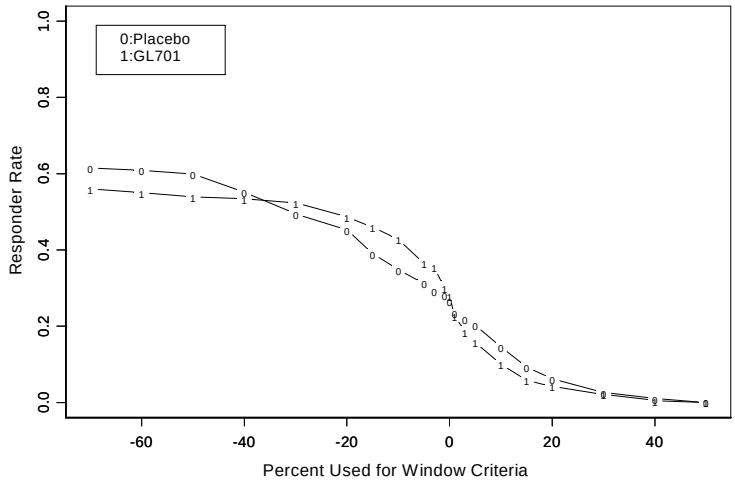
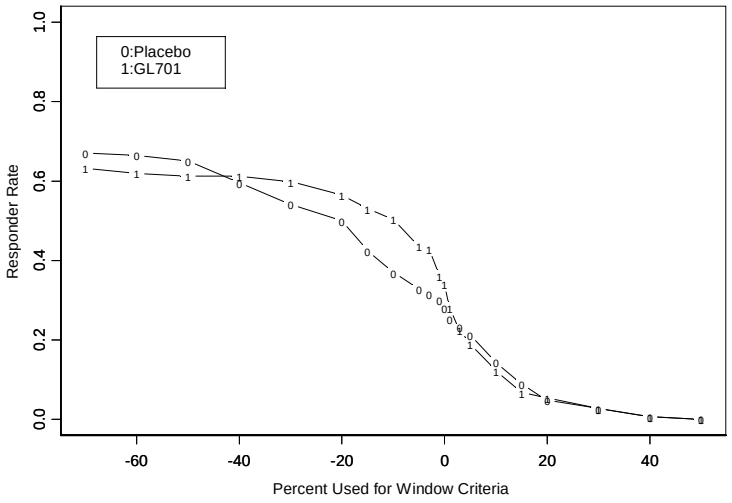


Figure 4. Responder Rate in ITT Patients with Baseline SLEDAI>2 with Dropouts Due to LOE and AE Considered as Nonrresponders



V. Final Conclusion

1. In Study 94-01, efficacy of GL701 was not demonstrated over placebo by any of the two primary endpoints (responder rate and percent decrease in predisone dose). For responder rate, although GL701 100 mg and 200 mg groups showed numerical advantage over placebo, no statistical significance was found. For percent decrease in predisone dose, placebo showed numerical advantage over the GL701 groups by mean.
2. In Study 95-02, although GL701 200 mg showed numerical advantage over placebo in responder rate, but no statistically significance was demonstrated. The dropout rates due to adverse events and lack of efficacy were both higher in the GL701 group. Therefore, when dropouts due to treatment failures were treated as non-responders, the numerical advantage of GL701 200 mg was mitigated (see reviewer's comment on Section IV.2.iii).

3. As discussed in Section IV.1, the result of the post-hoc subgroup (baseline SLEDAI>2) analysis in Study 94-01 did not provide a robust base for generating the hypothesis that GL701 is efficacious in patients with baseline SLEDAI>2. Further, in Study 95-02, although GL701 200 mg showed larger numerical advantages over placebo in responder rate in the subgroup with baseline SLEDAI>2 than in the overall ITT population, the advantages were not statistically significant. Therefore, additional data is needed in supporting the efficacy of GL701 200 mg in the subgroup with SLEDAI>2.

Laura Lu, Ph.D.

Mathematical Statistician

Concur:

Stan Lin, Ph.D.
Team Leader

CC:
NDA21239
HFD-550/MO/Johnson/Goldkind/Midthun
HFD-550/PM/Cook
HFD-550/Div. File
HFD-725/Lu/Lin ST./Huque
HFD-725/Div. File

Appendix A. Tables

Table a1. Demographics (Study 94-01)

Summary Statistics of Demographic/Medical History Characteristics by Treatment Group				
		Placebo (N = 64)	GL701 100mg (N = 63)	GL701 200mg (N = 64)
Age	Mean	40.6	40.0	40.2
	Median	39.0	39.0	41.0
	SD	10.96	12.17	9.84
	Range	22-70	18-75	21-66
Race n (%)	Asian	2 (3.1)	2 (3.2)	1 (1.6)
	African-American	17 (26.6)	16 (25.4)	17 (26.6)
	Caucasian	44 (68.8)	36 (57.1)	35 (54.7)
	Hispanic	0 (0.0)	8 (12.7)	9 (14.1)
	Other	1 (1.6)	1 (1.6)	2 (3.1)
Menopausal Status (%)	Pre- menopausal	38 (59.4)	37 (58.7)	48 (75.0)
	Post- menopausal	16 (25.0)	17 (27.0)	7 (10.9)
	Other	10 (15.6)	9 (14.3)	9 (14.1)
Smoke Now?	No (%)	47 (73.4)	45 (71.4)	49 (76.6)
	Yes (%)	17 (26.6)	18 (28.6)	15 (23.4)

Table a2. Baseline Characteristics (Study 94-01)

Efficacy Variable		Placebo	GL701 100 mg	GL701 200 mg
Prescribed Prednisone Dose (mg)*	Number	64	63	64
	Mean (SD)	15.2 (5.69)	13.7 (5.09)	13.7 (4.94)
	Median	15.0	12.5	10.0
	Range	10-30	10-30	10-30
SLEDAI Score	Number	64	63	64
	Mean (SD)	6.4 (5.58)	5.5 (3.93)	5.9 (5.00)
	Median	4.0	4.0	6.0
	Range	0-22	0-16	0-22
Patient's VAS	Number	64	63	64
	Mean (SD)	49.1 (25.04)	46.4 (22.38)	46.8 (22.02)
	Median	48.5	47.0	47.5
	Range	5-100	0-100	9-91
Physician's VAS	Number	64	63	64
	Mean (SD)	28.0 (19.94)	26.0 (17.02)	23.3 (15.10)
	Median	23.0	24.0	21.5
	Range	0-76	1-80	2-65
SF-36 Mental Component Summary (MCS)	Number	62	63	63
	Mean (SD)	42.8 (11.01)	45.4 (10.43)	45.1 (10.75)
	Median	43.0	47.5	48.4
	Range	17.19 - 61.00	20.80 - 62.94	18.21 - 62.33
SF-36 Physical Component Summary (PCS)	Number	62	63	63
	Mean (SD)	33.1 (11.43)	34.6 (10.36)	31.9 (9.27)
	Median	32.0	34.5	29.3
	Range	12.68 - 60.00	8.12 - 55.52	15.77 - 54.54
Krupp Fatigue Severity Score	Number	64	63	64
	Mean (SD)	5.3 (1.46)	5.1 (1.50)	5.4 (1.26)
	Median	5.7	4.9	5.7
	Range	1.9 - 7.0	1.1 - 7.0	1.0 - 7.0
SLICC Damage Index†	Number	64	63	64
	Mean (SD)	2.1 (2.02)	2.5 (2.66)	2.3 (2.53)
	Median	2.0	2.0	1.0
	Range	0-9	0-13	0-9

Table a.3 Mean Change from Baseline in Secondary Efficacy Variables (Study 94-01)

Secondary Efficacy Variables		Last Visit	P-value Between-Treatment (vs. Placebo)
SLEDAI Score	Placebo	-0.5	
	GL701 100	0.5	0.384
	GL701 200	0.0	0.753
SF-36 Physical Functioning Score	Placebo	-1.9	
	GL701 100	1.5	0.185
	GL701 200	-0.0	0.563
SF-36 Role-Physical Score	Placebo	0.0	
	GL701 100	0.0	0.853
	GL701 200	1.1	0.887
SF-36 Body-Pain Score	Placebo	2.1	
	GL701 100	0.8	0.965
	GL701 200	-5.6	0.062
SF-36 General Health Score	Placebo	1.8	
	GL701 100	0.2	0.81
	GL701 200	-0.5	0.477
SF-36 Vitality Score	Placebo	3.4	
	GL701 100	0.8	0.851
	GL701 200	1.3	0.325
SF-36 Social Functioning Score	Placebo	-1.4	
	GL701 100	-1.2	0.464
	GL701 200	-1.8	0.954
SF-36 Role Emotional Score	Placebo	-2.2	
	GL701 100	-0.8	0.489
	GL701 200	-11.1	0.498
SF-36 Mental Health Score	Placebo	2.5	
	GL701 100	2.6	0.622
	GL701 200	-0.7	0.428

Table a.3 Results on Secondary Variables (Study 94-01) (cont.)

Secondary Efficacy Variables		Last Visit	P-value Between-Treatment (vs. Placebo)
SF-36 Physical Component (PCS) Score	Placebo	-0.1	
	GL701 100	0.2	0.623
	GL701 200	0.0	0.998
SF-36 Mental Component (MCS) Score	Placebo	0.9	
	GL701 100	0.4	0.758
	GL701 200	-1.5	0.345
Krupp Fatigue Score	Placebo	-0.0	
	GL701 100	0.1	0.775
	GL701 200	-0.0	0.963
Physician Global Assessment	Placebo	-1.0	
	GL701 100	0.4	0.929
	GL701 200	3.2	0.655
Patient Global Assessment	Placebo	-0.9	
	GL701 100	-2.7	0.284
	GL701 200	4.1	0.367

Table a4. Demographic Summary By Treatment Group (Study 95-02)

	Placebo (N=192)	GL701 200 mg (N=189)
Age (yrs)		
Mean (SD)	43.8 (10.6)	44.4 (11.2)
Median	43.4	44.7
Range	18.0- 67.8	18.6- 69.1
Race		
Caucasian	137 (71.4)	146 (77.2)
African-American	33 (17.2)	22 (11.6)
Asian	3 (1.6)	2 (1.1)
Hispanic	16 (8.3)	15 (7.9)
Other	3 (1.6)	4 (2.1)

Table a.5 Baseline Values of Principal Efficacy Variables by Treatment Group (Study 95-02)

	Placebo	GL701 200 mg
SLAM Score (Range 0-60)		
N	192	189
Mean (SD)	12.0 (3.0)	12.2 (2.8)
Median	12.0	12.0
Range	4.0- 21.0	6.5- 21.0
SLEDAI Score (Range 0-105)		
N	192	189
Mean (SD)	5.8 (4.3)	6.5 (4.3)
Median	5.0	6.0
Range	0.0- 24.0	0.0- 18.0
Krupp Fatigue Score (Range 0-7)		
N	192	189
Mean (SD)	5.6 (1.2)	5.5 (1.2)
Median	5.7	5.9
Range	2.1- 7.0	1.1- 7.0
Patient Self Assessment (Range 0-100)		
N	192	189
Mean (SD)	55.4 (18.5)	55.2 (18.8)
Median	57.0	57.0
Range	8.5- 99.0	2.0- 91.5
Physician Global Assessment (Range 0-100)		
N	192	189
Mean (SD)	30.3 (13.5)	30.2 (13.8)
Median	28.5	27.0
Range	6.0- 77.0	2.5- 78.0
Mental Component Summary (MCS) (Range 0-100)		
N	190	187
Mean (SD)	41.7 (11.8)	42.5 (10.2)
Median	40.9	42.7
Range	12.8- 65.4	17.2- 65.4

Table a.5 Baseline Values of Principal Efficacy Variables by Treatment Group (cont.)
(Study 95-02)

	Placebo	GL701 200 mg
Physical Component Summary (PCS) (Range 0-100)		
N	190	187
Mean (SD)	31.6 (8.9)	31.1 (8.4)
Median	30.3	30.6
Range	12.2- 57.2	14.0- 57.2
SLICC Damage Index (Range 0-47)		
N	192	188
Mean (SD)	1.3 (1.4)	1.3 (1.4)
Median	1.0	1.0
Range	0.0- 9.0	0.0- 7.0

Table a.6 Percent Of Responders* by Treatment Group
(Per Protocol Population)

	Placebo N=176	**GL701 200 mg N=170**	**Percentage of Improvement GL701 over Placebo**	**P-value*** **
With Window**				
Responder	80 (45.5)	99 (58.2)	27.9%	0.0177
Non-Responder	96 (54.5)	71 (41.8)		
Without Window				
Responder	52 (29.5)	60 (35.3)	19.7%	0.2537
Non-Responder	124 (70.5)	110 (64.7)		

* A responder is defined as a patient who satisfies the following conditions: (1) Improvement or stabilization in all disease activity (i.e., SLAM, SLEDAI) and constitutional symptom assessments (i.e. KFSS, Patient VAS) and (2) no clinical deterioration.

** A responder with window is defined as a patient who satisfies the following conditions: (1) Weighted average change from baseline for SLAM is less than 1; for SLEDAI less than 0.5; for KFSS less than 0.5; for Patient VAS less than 10; and (2) no clinical deterioration.

*** P-value is from a logistic regression analysis with treatment as a factor.

Table a.7 Percent Of Responders* by Treatment Group (Per-Protocol Patients with Baseline SLEDAI >2)

	Placebo (N=133)	GL701 200 mg (N=132)	Percentage of Improvement GL701 over Placebo	P-value***
With Window**				
Responder	65 (48.9)	87 (65.9)	34.8%	0.0053
Non-Responder	68 (51.1)	45 (34.1)		
Without Window				
Responder	42 (31.6)	56 (42.4)	34.2%	0.0682
Non-Responder	91 (68.4)	76 (57.6)		

* A responder is defined as a patient who satisfies the following conditions: (1) Improvement or stabilization in all disease activity (i.e., SLAM, SLEDAI) and constitutional symptom assessments (i.e. KFSS, Patient VAS) and (2) no clinical deterioration.

** A responder with window is defined as a patient who satisfies the following conditions: (1) Weighted average change from baseline for SLAM is less than 1; for SLEDAI less than 0.5; for KFSS less than 0.5; for Patient VAS less than 10; and (2) no clinical deterioration.

*** P-value is from a logistic regression analysis with treatment as a factor.

Table a.8 First Definite Flares (Per-Protocol Population)

	Per Protocol Population*		Per Protocol Population Baseline SLEDAI >2**	
	Placebo (N=176)	GL701 200 mg (N=170)	Placebo (N=133)	GL701 200 mg (N=132)
Number of Patients Experiencing At Least One Definite Flare While on Study Drug	47 (26.7%)	37 (21.8%)	41 (30.8%)	31 (23.5%)

* P-value (p=0.3353) is from a log-rank test for time to first definite flare.

** P-value (p=0.2013) is from a log-rank test for time to first definite flare.

First 60 days were excluded from analysis; patients were followed up for 7 days after their last medication date.

Table a.9 Survival Analysis for Clinical Deterioration (Per Protocol Population)

	Placebo (N=176)	GL701 200 mg (N=170)
Number of Patients Experiencing Clinical Deterioration	15 (8.5%)	13 (7.6%)
P-value (GL701 vs. placebo)		0.639

*p-value by log-rank test.

Table a.10 Survival Analysis for Clinical Deterioration (Per-Protocol Patients with Baseline SLEDAI >2)

	Placebo (N=133)	GL701 200 mg (N=132)
Number of Patients Experiencing Clinical Deterioration	12 (9.0%)	13 (9.8%)
P-value (GL701 vs. placebo)		0.788

*p-value by log-rank test.

Table a.11 Mean Change in Scoring Instruments From Baseline

Variable	Per-Protocol		Per-Protocol Baseline SLEDAI > 2	
	Placebo (N=176)	**GL701 (N=170)**	**Placebo (N=133)**	**GL701 (N=132)**
SLEDAI	-1.72	-2.24	-2.57	-3.17
	(N=175)	(N=170)	(N=132)	(N=132)
Patient VAS	-4.35	-6.24	-2.85	-7.22
	(N=175)	(N=169)	(N=132)	(N=131)
Physician VAS	-5.19	-5.64	-4.52	-5.38
	(N=175)	(N=169)	(N=132)	(N=131)
KFSS	-0.39	-0.33	-0.27	-0.32
	(N=175)	(N=169)	(N=132)	(N=131)
SLAM	-2.65	-3.10	-2.63	-3.16
	(N=175)	(N=170)	(N=132)	(N=132)
SLICC	-0.05	-0.08	-0.06	-0.09
	(N=138)	(N=128)	(N=102)	(N=97)
SF36 - MCS	1.80	2.64	1.64	2.33
	(N=172)	(N=166)	(N=130)	(N=129)
SF36 - PCS	1.71	1.76	0.90	1.87
	(N=172)	(N=166)	(N=130)	(N=129)

Table a.12 Responder Rates along Percentage Used in Window Definition (ITT)

Percentage Used in Window Definition	Responder Rate in Placebo	Responder Rate in GL701 200 mg	P-value* (GL701 200 mg vs. Placebo)
-70	66.1%	67.2%	0.828
-60	64.6%	66.1%	0.750
-50	63.0%	65.1%	0.675
-40	57.8%	63.5%	0.257
-30	52.1%	60.8%	0.085
-20	47.4%	55.6%	0.111
-15	40.6%	51.9%	0.028**
-10	35.9%	48.1%	0.016**
-5	32.3%	40.7%	0.087
-3	29.7%	39.7%	0.040**
-1	28.6%	33.9%	0.272
0	27.1%	30.7%	0.438
1	24.5%	25.0%	0.745
3	22.4%	21.2%	0.771
5	20.3%	17.5%	0.477
10	14.6%	10.6%	0.239
15	9.4%	6.3%	0.273
20	6.3%	4.8%	0.525
30	2.6%	2.1%	0.754
40	1.0%	0.5%	0.571
50	0.0%	0.0%	1.000

*: P-values from Mantel-Haenszel Tests
**: P-values Less Than 0.05

Table a.13 Responder Rates along Percentage Used in Window Definition in Subgroup with Baseline SLEDAI>2 (ITT)

Percentage Used in Window Definition	Responder Rate in Placebo (%)	Responder Rate in GL701 200 mg (%)	P-value* (GL701 200 mg vs. Placebo)
-70	73.3	76.2	0.557
-60	71.2	74.8	0.488
-50	69.2	74.1	0.345
-40	63.0	72.8	0.073
-30	57.5	70.1	0.026
-20	52.7	64.6	0.039
-15	44.5	59.9	0.009
-10	38.4	57.1	0.001
-5	34.2	49.0	0.011
-3	32.2	48.3	0.005
-1	30.8	40.8	0.074
0	28.8	37.4	0.116
1	26.0	31.3	0.319
3	24.0	25.9	0.710
5	21.2	21.1	0.976
10	14.4	12.9	0.716
15	8.9	7.5	0.657
20	4.8	6.1	0.617
30	2.7	2.7	0.992
40	0.7	0.7	0.996
50	0.0	0.0	1.000

*: P-values from Mantel-Haenszel Tests
**: P-values Less Than 0.05

Appendix B. Figures

Figure b.1 Survival Curves for Early Termination of Study Medication Due to Lack of Efficacy (Study 94-01)

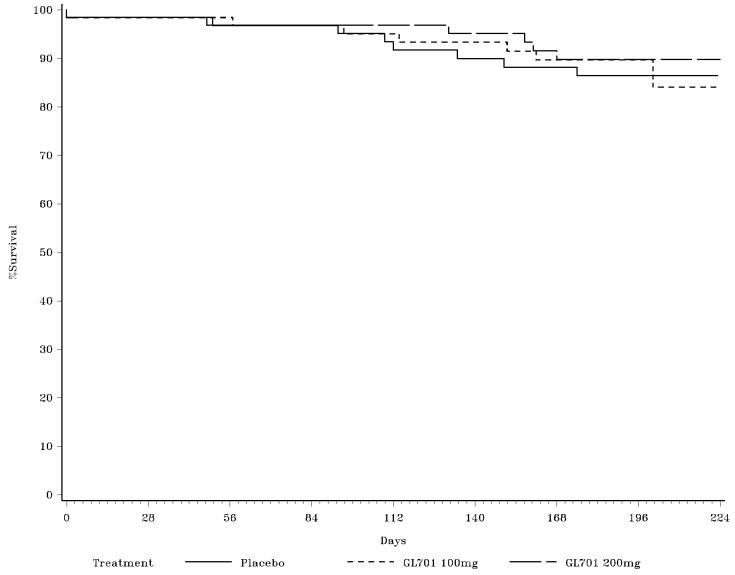


Figure b.2 Survival Curves for Early Termination of Study Medication Due to Adverse Events (Study 94-01)

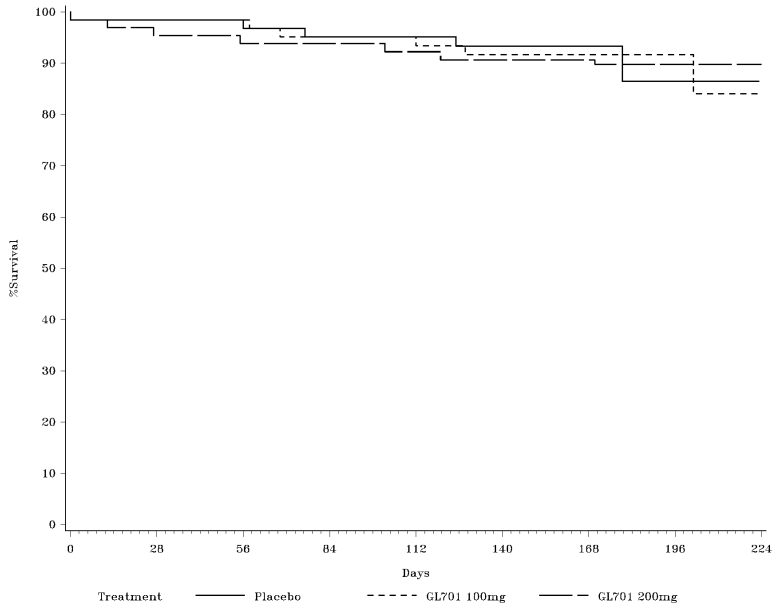


Figure b.3 Survival Curves for Early Termination of Study Medication Due to Lack of Efficacy (Study 95-02)

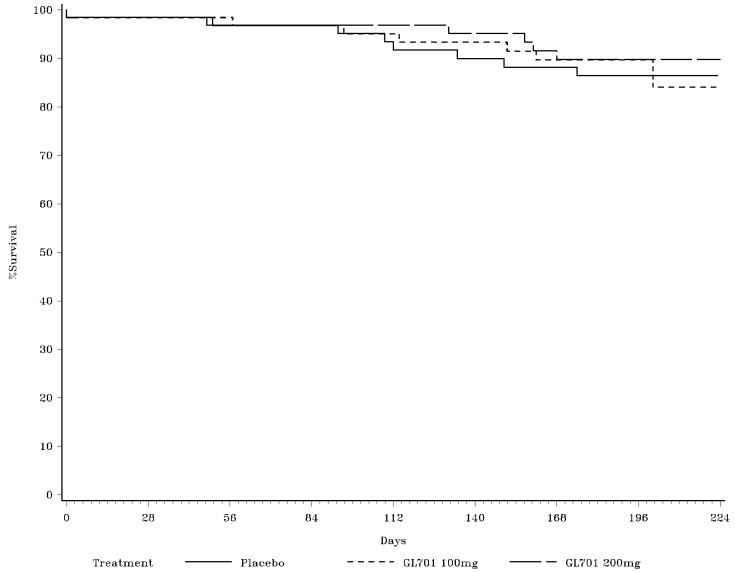


Figure b.4 Survival Curves for Early Termination of Study Medication Due to Adverse Events (Study 95-02)

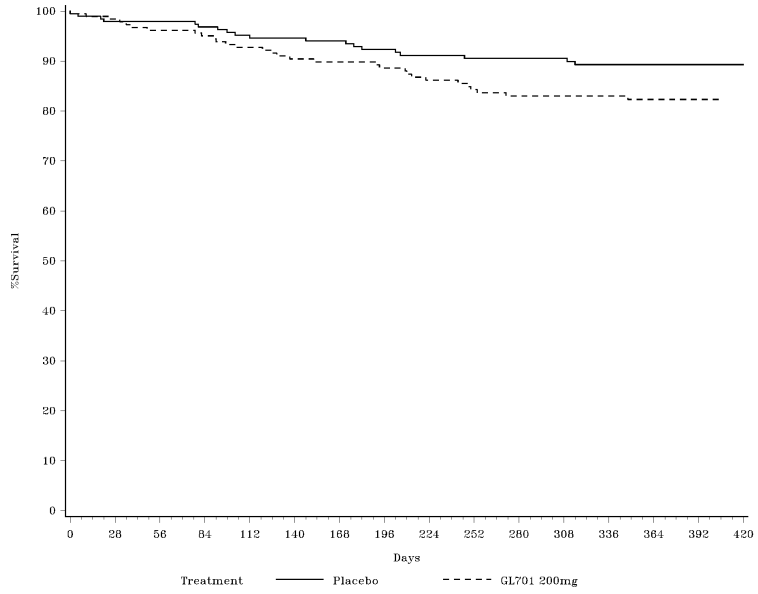


Figure b.5 Time to First Definite Flare Survival Curve in ITT Population
(Study 95-02)

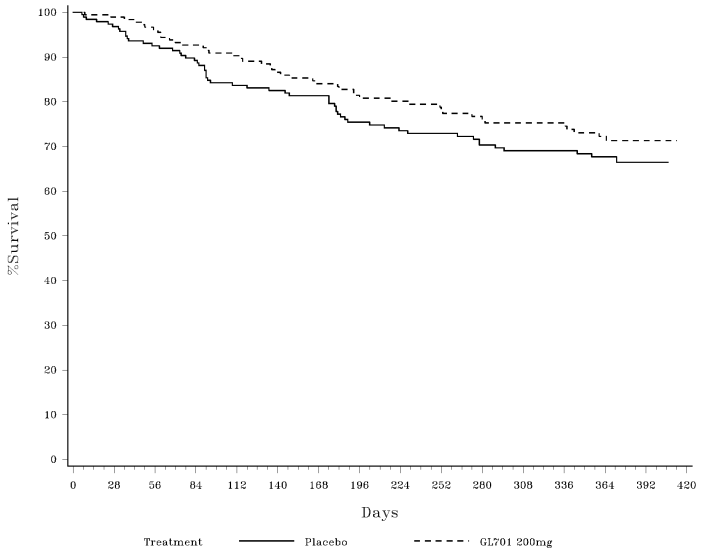
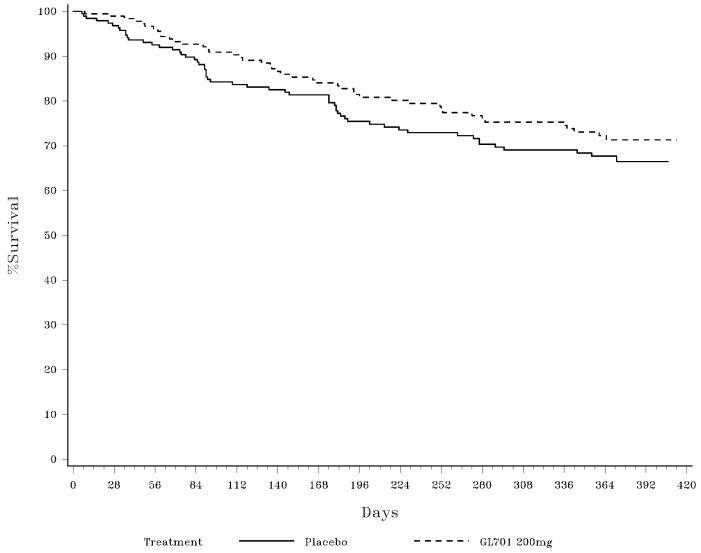


Figure b.6 Time to Clinical Deterioration Survival Curve in ITT Population
(Study 95-02)



Appendix C. Handling of Missing Value in Study 95-02

Table c.1 Handling of Patients' Missing VAS, KFSS Average, SLEDAI or SLAM Score for The Determination Of The Primary Variable (Response)

Missing status	**Per-protocol analysis +**	**Intention-to-treat analysis++**
Baseline #		
Individual Item* KFSS	Average of non-missing items	Same as the per-protocol analysis
Individual Item SLEDAI or SLAM	**Item(s) missing at baseline will be considered as missing for all visits.** The rest of the item total will be used to compare changes.	Same as the per-protocol analysis
Missing entire KFSS, Patient VAS, SLAM (unlikely), or SLEDAI (unlikely)	Missing score(s) will be considered as missing for all visits. The rest of scores will be used to determine responder status	**Same as the per-protocol analysis**
All 4 scores (unlikely)	Patient excluded	Failure
Post-baseline (when baseline is non-missing)		
Individual Item KFSS	Average of non-missing items	Same as the per-protocol analysis
Individual Item SLEDAI or SLAM	Last item carried forward	Same as the per-protocol analysis
Missing entire KFSS, Patient VAS, SLAM (unlikely), or SLEDAI (unlikely)	Weighted average (by duration) of remaining on-treatment scores	Same as the per-protocol analysis
All post-baseline measurements and all visits	Exclusion from analysis	Failure
Treatment duration less than 61 days	Exclusion from analysis	Include in the analysis.

+ Patients who were on study drug > 60 days and who have any measurements of SLE scores or other data (pertinent to determination of flare or clinical deterioration) beyond 60 days.
++ Any randomized patient.
* An individual item of a score is the item sub-score (e.g., "Weight loss" in SLAM).
# Individual item at baseline is the item average of the two pre-treatment visits.
Note: all items refer to "on-treatment" visit data.

Table c.2 Determination of The Variable "Number of Patients with at Least One Flare", "Time to First Flare", rr "Time to Clinical Deterioration"

Missing status	**Per-protocol analysis**	**Intention-to-treat analysis**
Every patient will be classified	If a patient's on-treatment period is shorter than 61 days, she will not be included in the per-protocol analysis	If a patient's on-treatment period is shorter than 61 days, she will be included in the analysis.
Treatment Duration	**Per-protocol analysis**	**Intention-to-treat analysis**
Treatment duration $\leq$ 60 days	Patient not included in either clinical deterioration or flare analyses	Included in both responder and flare analyses, with any missing data handled as described above. Responder and flare status analysed only while patient is on treatment
Treatment duration > 60 days	Patient included, but clinical deterioration and flare status analysed only while patient is on treatment, excluding first 60 days	Included in both responder and flare analyses, with any missing data handled as described above. Responder and flare analysed only while patient is on treatment

Table c.3 Handling of Missing Data for Analysis of Mean Changes in SLE Scores

Missing status	**Per-protocol analysis**	**Intention-to-treat analysis**
Baseline #		*No separate intention-to-treat analysis for analysis of mean changes*
Individual Item * KFSS	Average of non-missing items	
Individual Item SF-36	Missing items handled by the SF-36 system automatically (see Ware's SF-36 Health Survey Manual and Interpretation Guide for detail).	
Individual Item SLEDAI or SLAM	All post-baseline will be missing	
Any score	All post-baseline will be missing	
Post-baseline (when baseline is non-missing)		
Individual Item * KFSS	Average of non-missing items	
Individual Item SF-36	Missing items handled by the SF-36 system automatically (see Ware's SF-36 Health Survey Manual and Interpretation Guide for detail).	
Individual Item SLEDAI or SLAM	Last visit carried forward	
Any score @	Score will be missing for that visit's average. Weighted average for overall change from baseline.	

* An individual item of a score is the item sub-score.
# Individual item at baseline is the item average of the two pre-treatment visits.
@ A score is either Patient's VAS, physician's VAS, KFSS average, SLEDAI total, SLAM total or SF-36 (M & P).
Note: all measurements and items refer to "on-treatment" visit data.