

CENTER FOR DRUG EVALUATION AND RESEARCH

APPROVAL PACKAGE FOR:

APPLICATION NUMBER

NDA 21-615

Statistical Review(s)



U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH
OFFICE OF PHARMACOEPIDEMIOLOGY AND STATISTICAL SCIENCE
OFFICE OF BIOSTATISTICS

STATISTICAL REVIEW AND EVALUATION

Clinical Studies

NDA/Serial Number: 21-615/000
Drug Name: Reminyl®(galantamine) Extended Release Capsules
Indication: Alzheimer
Applicant: Johnson & Johnson
Date: 5/27/2004
Review Priority: 2 months

Biometrics Division: I (HFD 710)
Statistical Reviewer: Kun He
Concurring Reviewers: Kun Jin, , Ph.D., Team Leader
Jim Hung, Ph.D., Acting Deputy Director

Medical Division: Neuropharmacological Drug Products (HFD 120)
Clinical Team: Ranjit Mani, M.D., Clinical Reviewer
Russell Katz, M.D., Director
Project Manager: Merrill Mille, R. Ph.

Keywords: Alzheimer, Reminyl, ANCOVA, CMH

TABLE OF CONTENTS

1. Executive Summary3

1.1 Conclusions and Recommendations3

1.2 Brief Overview of Clinical Studies.....3

1.3 Statistical Issues and Findings4

2. Introduction5

2.1 Overview.....5

2.2 Data Sources.....5

3. Statistical Evaluation5

3.1 Evaluation of Efficacy.....5

 3.1.1 Analysis Stratified by Study Site.....5

 3.1.2 Analysis Stratified by Screening MMSE and Country.....8

 3.1.3 Analysis Stratified by Baseline ADAS-cog/11 and Country.....10

 3.1.4 Analysis Stratified by Prior Cholinomimetic Use and Country.....12

 3.1.5 Analysis in US Population14

3.2 Evaluation of Safety15

 4. Findings in Special/Subgroup Populations.....15

5. Summary and Conclusions.....15

5.1 Statistical Issues and Collective Evidence15

5.2 Conclusions and Recommendations15

Appears This Way
On Original

Statistical Review and Evaluation

1. Executive Summary

1.1 Conclusions and Recommendations

The post-hoc re-analyses of CIBIC-plus at Week 26 doesn't provide convincing evidence that CIBIC-plus at Week 26 is significantly different between GAL-ER vs. Placebo, and GAL-IR vs. Placebo.

1.2 Brief Overview of Clinical Studies

The primary objective of this study was to evaluate the safety and efficacy (as measured by Alzheimer's Disease Assessment Scale: sum of 11 cognitive items [ADAS-cog/11] and Clinician's Interview Based Impression of Change – Plus Caregiver Input [CIBIC-plus]) of a flexible dosing regimen (16 or 24 mg/day) of galantamine controlled-release (CR) compared with placebo in subjects with mild to moderate Alzheimer's disease. Secondary analysis includes comparisons for efficacy measurements between galantamine IR and placebo and between galantamine CR and IR. This double-blind, parallel-group, placebo-controlled, flexible-dose study was conducted in the U.S., Australia, Canada, South Africa, and New Zealand. Following a 4-week, single-blind, placebo run-in period, subjects were randomized to receive placebo, galantamine IR, or galantamine CR in double-blind fashion for 26 weeks. Subjects in the galantamine treatment groups received 4 weeks of galantamine CR 8 mg q.d. or IR 4 mg b.i.d., followed by 4 weeks of galantamine CR 16 mg q.d. or IR 8 mg b.i.d. Based on safety and tolerability, the galantamine dose could be increased to CR 24 mg q.d. or IR 12 mg b.i.d. at Week 8, and could be reduced to CR 16 mg q.d. or IR 8 mg b.i.d. at Week 12. The dose chosen at the end of Week 12 was fixed for the remainder of the study. 971 subjects were randomized with 965 subjects analyzed for safety, and 925 subjects analyzed for efficacy (Intent-to-Treat [ITT] Analysis Set).

Efficacy was evaluated by the ADAS-cog/11, CIBIC-plus, ADCS-ADL, and NPI measurements at baseline and at Weeks 8, 12, and 26. The primary efficacy analyses were to compare galantamine CR with placebo with respect to change in ADAS-cog/11 scores from baseline to Week 26 and CIBIC-plus scores at Week 26. The changes in the ADAS-cog/11 score were analyzed using analysis of variance (ANOVA) models with treatment and country (U.S. vs. non-U.S.) factors. The Cochran-Mantel-Haenszel (CMH) test using modified ridit scores, derived from rank scores (the Van Elteren test) controlling for country (U.S. vs. non-U.S.) effect, was applied to compare the distributions between each pair of the treatments for the CIBIC-plus score. The primary efficacy analysis was based on the observed case (OC) data for the ITT analysis set. (*In the protocol stage, the Agency pointed out to the applicant that LOCF is commonly used for the primary analysis for this indication.*) Last observation carried forward (LOCF) analysis for the ITT analysis set was used for the primary analyses. Treatment with galantamine CR and IR led to statistically significant improvements in the primary efficacy endpoint (ADAS-cog/11) compared with placebo at Week 26.

Both galantamine CR and IR treatments were numerically better but not statistically different from that of the placebo group in maintaining global function assessed by CIBIC-plus scores at Week 26.

P-values for ADAS-cog/11 are .0001 using LOCF and OC. P-values for CIBIC-plus are .216 using LOCF and .0859 using OC, respectively. The decision rule is to have significance on both ADAS-cog/11 and CIBIC-plus simultaneously.

In this submission, the applicant re-analyzed CIBIC-plus at Week 26 LOCF data with CMH using rank scores. The applicant presented the results of CMH stratified by site, CMH stratified by country and screening MMSE, CMH stratified by country and ADAS-cog/11, CMH stratified by country and prior cholinomimetic use, and CMH stratified by site for USA population.

1.3 Statistical Issues and Findings

CMH stratified by region or by country is more commonly used method in analyzing CIBIC-plus when the study is an international study, and more appropriate to adjust for country effect. Unfortunately, none of them is significant. The significance result of CMH stratified by site is questionable since there is no pre-specified pooling plan. Furthermore, it is not clear whether the significance result of CMH stratified by site is driven by one extremely large pooled site.

Since none of CMH without stratification, CMH stratified by country, and CMH stratified by disease severity (MMSE, ADAS-cog/11, and prior cholinomimetic use) is significant for GAL-ER vs. Placebo, the significance result of CMH stratified by both disease severity and country might be due to interaction.

In USA population, CMH stratified by site shows significance and CMH without stratification also shows significant for GAL-ER vs. Placebo, but one should interpret the subgroup result with caution since it is a subgroup analysis.

Appears This Way
On Original

2. Introduction

2.1 Overview

In this resubmission, the applicant re-analyzed GAL-INT-10 CIBIC-plus data using the following approaches: a design-based analysis stratified by study site, 3 similar analyses stratified by country and baseline disease severity to adjust for prognostic factors, and an analysis of the U.S. population alone. All additional analyses presented were performed based on ITT (intent-to-treat) LOCF (last-observation-carried-forward) data using 2-sided statistical tests. All CMH tests were performed using rank scores.

2.2 Data Sources

The path to the CDER Electronic Document Room (EDR) is:
\\CDSESUB1\N21615\N_000\2003-05-27

3. Statistical Evaluation

3.1 Evaluation of Efficacy

3.1.1 Analysis Stratified by Study Site

The protocol-specified analysis of CIBIC-plus used a CMH with modified riddit scores and stratified by region (U.S. vs. non-U.S.). The modified riddit score is defined as the rank score further divided by the stratum sample size. This use of modified riddit scores and U.S. vs. non-U.S. stratification inappropriately gives equal weight to the non-U.S. region and the U.S. region, the latter of which comprised 69% of the study population. This analysis, therefore, effectively down-weights the contribution of the U.S. subjects. Since the randomization of subjects was stratified by study site, a design-based analysis, which weights each subject's contribution toward the overall analysis result equally, would be a CMH using rank scores and stratified by study site.

GAL-INT-10 CIBIC-plus results of analysis stratified by study site are summarized in Table 3.1.1.1. Small study sites with 3 or fewer subjects per treatment cell were pooled. When GAL-INT-10 CIBIC-plus data are analyzed using this design-based analysis, both GAL-ER treatment and GAL-IR treatment were statistically superior to placebo (Table 3.1.1.1). The sponsor believes that this analysis reflects more accurately the overall treatment effect of GAL-ER and GAL-IR on CIBIC-plus.

Table 3.1.1.1 GAL-INT-10 Week 26 (LOCF) CIBIC-plus Results

	PLACEBO	GAL ER	GAL IR
All ITT Subjects	N=301	N=296	N=302
CIBIC-plus			
Markedly improved	1.0%	1.0%	1.0%
Moderately improved	3.7%	4.7%	5.0%
Mildly improved	15.9%	16.6%	15.2%
No change	36.9%	38.5%	42.1%
Mildly worse	26.6%	27.4%	25.8%
Moderately worse	13.6%	9.8%	9.9%
Markedly worse	2.3%	2.0%	1.0%
p-value^a (vs Placebo)		0.030	0.027

^aCMH model stratified by study site.

Reviewer's Comment:

All CMH calculation in this review will use rank scores which are same as the applicant's method in this submission, despite the original protocol specified method using modified ridit scores.

CMH without any stratification give p-values .2770 for GAL-ER vs. Placebo, and .1512 for GAL-IR vs. Placebo, respectively, using rank scores.

CMH stratified by region (USA vs. non-USA) give p-values .0951 for GAL-ER vs. Placebo, and .0958 for GAL-IR vs. Placebo, respectively.

CMH stratified by country (Australia, Canada, New Zealand, South Africa, and USA) give p-values .0582 for GAL-ER vs. Placebo, and .0845 for GAL-IR vs. Placebo, respectively.

Since there are 91 sites in the study, after pooling sites with few than 3 patients to big sites within each country, there are still 32 sites as listed in Table 3.1.1.2. Number of subjects in pooled site in Australia is 991, in Canada 992, in New Zealand 993, in South Africa 994, and in USA 995, respectively.

Table 3.1.1.2 Number of Subjects and P-value by Site

Site	GAL-ER	Placebo	Total	P-value
2	4	4	8	.2248
3	8	8	16	.2312
7	6	8	14	.3650
15	7	6	13	.1002
16	4	4	8	.2248
21	5	5	10	.0647
24	11	10	21	.3184
30	6	6	12	.4344
35	8	8	16	.2695
36	5	4	9	.8927
38	9	10	19	.3012
41	7	7	14	.4170
48	5	6	11	.3976
49	5	4	9	.6547
50	6	5	11	.7798
51	5	6	11	.5627
54	4	4	8	.3428
56	9	10	19	.0824
101	4	4	8	.1342
102	5	5	10	.7346
106	5	5	10	.7364
107	4	4	8	.8770
109	9	8	17	.0505
213	5	6	11	.6889
302	5	5	10	.1175
304	6	6	12	.6045
305	7	8	15	.3645
307	8	7	15	.4551
991	21	22	43	.4726
992	5	6	11	.2189
993	2	3	5	1.000
994	8	8	16	1.000
995	88	89	177	.0293

P-values are calculated using CMH and listed. Since sample sizes in sites are small, p-values are not used to check difference within each site. Notice that the pooled site 995 is extremely large with p-value .0293, it is not clear whether the significance result of CMH stratified by site is driven by this particular pooled site. Without pre-specification of pooling plan, the result of CMH using the post-

hoc pooling site is difficult to interpret.

CMH stratified by region or by country is more commonly used method in analyzing CIBIC-plus when the study is an international study, and more appropriate to adjust for country effect. Unfortunately, none of them is significant. The significance result of CMH stratified by site is questionable since there is no pre-specified pooling plan. Furthermore, it is not clear whether the significance result of CMH stratified by site is driven by one extremely large pooled site.

3.1.2 Analysis Stratified by Screening MMSE and Country

A summary of CIBIC-plus data by screening MMSE categories ≤ 22 and > 22 is provided in Table 3.1.2.1. In the subgroup of subjects with screening MMSE score ≤ 22 , the CIBIC-plus responder rates were consistent with prior GAL-IR AD studies. In the subgroup of subjects with screening MMSE > 22 , the CIBIC-plus responder rates in all treatment groups are higher than for the remainder of the study population. The 77% CIBIC-plus responder rate in the placebo group was notably high and greatly reduced the CIBIC-plus sensitivity.

An analysis using a CMH stratified by country and screening MMSE category (≤ 22 versus > 22) was performed. Statistical significance in favor of GAL-ER is achieved for the comparison between GAL-ER and placebo. Sensitivity analyses using 20, 21, and 23 as the cutoff point confirmed the robustness of this analysis (Table 3.1.2.2). Statistical significance in favor of GAL-IR is achieved for the comparison between GAL-ER and placebo for the cutoff points of 21 and 23, but it just missed the 0.05 significance level for the cutoff points of 20 and 22 (Table 3.1.2.2).

Appears This Way
On Original

**Table 3.1.2.1 Summary of CIBIC-plus at Week 26 –
LOCF Data by Screening MMSE Category (≤ 22 and >22)**

	PLACEBO	GAL-ER	GAL-IR
Subjects with MMSE ≤ 22	N=248	N=255	N=256
CIBIC-plus			
Markedly improved	0.8 ^o _o	1.2 ^o _o	0.8 ^o _o
Moderately improved	2.0 ^o _o	4.7 ^o _o	4.3 ^o _o
Mildly improved	14.5 ^o _o	14.9 ^o _o	15.2 ^o _o
No change	35.9 ^o _o	39.6 ^o _o	40.6 ^o _o
Mildly worse	28.2 ^o _o	26.3 ^o _o	27.3 ^o _o
Moderately worse	15.7 ^o _o	11.0 ^o _o	10.6 ^o _o
Markedly worse	2.8 ^o _o	2.3 ^o _o	1.2 ^o _o
Subjects with MMSE >22	N=53	N=41	N=46
CIBIC-plus			
Markedly improved	1.9 ^o _o	0.0 ^o _o	2.2 ^o _o
Moderately improved	11.3 ^o _o	4.9 ^o _o	8.7 ^o _o
Mildly improved	22.6 ^o _o	26.8 ^o _o	15.2 ^o _o
No change	41.5 ^o _o	31.7 ^o _o	50.0 ^o _o
Mildly worse	18.9 ^o _o	34.2 ^o _o	17.4 ^o _o
Moderately worse	3.8 ^o _o	2.4 ^o _o	6.5 ^o _o
Markedly worse	0.0 ^o _o	0.0 ^o _o	0.0 ^o _o
p-value^a (vs Placebo)		0.019	0.053

^aCMH model stratified by screening MMSE and country

**Table 3.1.2.2 Summary of CIBIC-plus at Week 26 –
LOCF Data by Screening MMSE Category and Country**

Strata for CMH Model	GAL-ER vs. Placebo	GAL-IR vs. Placebo
MMSE (10-20, 20), country	0.022	0.052
MMSE (10-21, 21), country	0.013	0.026
MMSE (10-22, 22), country	0.019	0.053
MMSE (10-23, 23), country	0.022	0.043

Reviewer's Comment:

P-values for CMH stratified by MMSE are given in the following table. For GAL-ER vs. Placebo, none are significant.

**Table 3.1.2.3 Summary of CIBIC-plus at Week 26 –
LOCF Data by Screening MMSE Category**

Strata for CMH	GAL-ER vs. Placebo	GAL-IR vs. Placebo
MMSE (10-20, >20)	.1221	.0540
MMSE (10-21, >21)	.0681	.0229
MMSE (10-22, >22)	.0777	.0477
MMSE (10-23, >23)	.1101	.0496

Since none of CMH without stratification, CMH stratified by country, and CMH stratified by MMSE is significant for GAL-ER vs. Placebo, the significance result of CMH stratified by both MMSE and country might be due to interaction.

3.1.3 Analysis Stratified by Baseline ADAS-cog/11 and Country

CIBIC-plus data summarized by baseline ADAS-cog/11 categories of ≤ 18 and > 18 are provided in Table 3.1.3.1. In the subgroup of subjects with baseline ADAS-cog/11 > 18 , the CIBIC-plus responder rates were consistent with prior GAL-IR AD studies. In the subgroup of subjects with baseline ADAS-cog/11 ≤ 18 , representing more mild subjects, the CIBIC-plus responder rates in all treatment groups are considerably higher. The 78% CIBIC-plus responder rate in the placebo group was notably high and greatly reduced the CIBIC-plus sensitivity.

An analysis of CIBIC-plus data using a CMH stratified by country and baseline ADAS-cog/11 score (≤ 18 versus > 18) was performed. The cutoff point of ≤ 18 was used to create a comparable percentage to subjects with a screening MMSE > 22 . Sensitivity analyses using other cutoff points are also presented to show the robustness of this analysis. Statistical significance is achieved for the comparison between GAL-ER and placebo and between GAL-IR and placebo at all cutoff points (16, 17, 18, 19, and 20) (Tables 3.1.3.1 and 3.1.3.2).

**Appears This Way
On Original**

Table 3.1.3.1 Summary of CIBIC-plus at Week 26 - LOCF Data by Baseline ADAS-cog/11 and Country

	PLACEBO	GAL-ER	GAL-IR
Subjects with Baseline ADAS-cog/11 >18	N=223	N=234	N=246
CIBIC-plus			
Markedly improved	0.5%	0.8%	0.8%
Moderately improved	2.2%	4.3%	4.1%
Mildly improved	15.3%	15.8%	14.6%
No change	32.7%	35.9%	42.3%
Mildly worse	29.2%	29.1%	25.2%
Moderately worse	17.0%	12.0%	11.8%
Markedly worse	3.1%	2.1%	1.2%
Subjects with Baseline ADAS-cog/11 ≤18	N=74	N=57	N=51
CIBIC-plus			
Markedly improved	2.7%	1.8%	2.0%
Moderately improved	8.1%	5.3%	9.8%
Mildly improved	18.9%	19.3%	17.6%
No change	48.6%	50.9%	45.1%
Mildly worse	17.6%	21.0%	23.5%
Moderately worse	4.1%	1.8%	2.0%
Markedly worse	0.0%	0.0%	0.0%
p-value^a (vs Placebo)		0.021	0.015

^aCMH model stratified by baseline ADAS-cog/11 score and country.

Table 3.1.3.2 GAL-INT-10 Week 26 CIBIC-plus Data Results (p-values) from CMH Model Stratified by Baseline ADAS-cog/11 Score and Country

Strata for CMH Model	GAL-ER vs Placebo	GAL-IR vs Placebo
ADAS-cog/11 (≤16, 16), Country	0.022	0.019
ADAS-cog/11 (≤17, 17), Country	0.019	0.016
ADAS-cog/11 (≤18, 18), Country	0.021	0.015
ADAS-cog/11 (≤19, 19), Country	0.025	0.020
ADAS-cog/11 (≤20, 20), Country	0.030	0.029

Reviewer's Comment:

P-values for CMH stratified by ADAS-cog/11 are given in the following table. For GAL-ER vs. Placebo, none is significant.

Table 3.1.3.3 GAL-INT-10 Week 26 CIBIC-plus Data Results (p-values) from CMH Model Stratified by Baseline ADAS-cog/11 Score

Strata for CMH	GAL-ER vs. Placebo	GAL-IR vs. Placebo
ADAS-cog/11 ($\leq 16, > 16$)	.1198	.0337
ADAS-cog/11 ($\leq 17, > 17$)	.1289	.0361
ADAS-cog/11 ($\leq 18, > 18$)	.1107	.0251
ADAS-cog/11 ($\leq 19, > 19$)	.0857	.0190
ADAS-cog/11 ($\leq 20, > 20$)	.1282	.0394

Since none of CMH without stratification, CMH stratified by country, and CMH stratified by ADAS-cog/11 is significant for GAL-ER vs. Placebo, the significance result of CMH stratified by both ADAS-cog/11 and country might be due to interaction.

3.1.4 Analysis Stratified by Prior Cholinomimetic Use and Country

A summary of CIBIC-plus data categorized by prior cholinomimetic use is provided in Table 3.1.4.1. In the subgroup of subjects with prior cholinomimetic use, the CIBIC-plus responder rates were consistent with prior GAL-IR AD studies. In the subgroup of subjects without prior cholinomimetic use, the CIBIC-plus responder rates in all treatment groups are higher than for the rest of the study population.

An analysis of CIBIC-plus data using a CMH stratified by country and prior cholinomimetic use status was performed. Statistical significance in favor of GAL-ER is achieved for the comparison between GAL-ER and placebo (Table 3.1.4.1). It just missed the 0.05 significance level for GAL-IR.

Appears This Way
On Original

**Table 3.1.4.1 Summary of CIBIC-plus at Week 26 –
LOCF Data by Prior Cholinomimetic Use**

	PLACEBO	GAL ER	GAL IR
Subjects with Prior Cholinomimetics Use	N=137	N=142	N=145
CIBIC-plus			
Markedly improved	2.2%	0.7%	1.4%
Moderately improved	2.9%	4.9%	3.4%
Mildly improved	10.9%	10.6%	11.7%
No change	31.4%	40.1%	40.7%
Mildly worse	29.9%	27.5%	29.0%
Moderately worse	19.0%	12.7%	11.7%
Markedly worse	3.6%	3.5%	2.1%
Subjects without Prior Cholinomimetics Use	N=164	N=153	N=157
CIBIC-plus			
Markedly improved	0.0%	1.3%	0.6%
Moderately improved	4.3%	4.6%	6.4%
Mildly improved	20.1%	22.2%	18.5%
No change	41.5%	37.2%	43.3%
Mildly worse	23.8%	26.8%	22.9%
Moderately worse	9.1%	7.2%	8.3%
Markedly worse	1.2%	0.7%	0.0%
p-value^a (vs Placebo)		0.030	0.052
^a CMH model stratified by cholinomimetics status and country.			

Reviewer's Comment:

P-values for CMH stratified by prior cholinomimetic use are .2140 for GAL-ER vs. Placebo, and .1192 for GAL-IR vs. Placebo, respectively.

Since CMH without stratification, CMH stratified by country, and CMH stratified by prior cholinomimetic use are not significant, it is not clear how to interpret the results of CMH stratified by both prior cholinomimetic use and country.

Since none of CMH without stratification, CMH stratified by country, and CMH stratified by prior cholinomimetic use is significant for GAL-ER vs. Placebo, the significance result of CMH stratified by both prior cholinomimetic use and country might be due to interaction.

3.1.5 Analysis in US Population

There was a difference between the U.S. and non-U.S. population in the severity of the subjects' dementia (i.e., severity was greater in the U.S. population). Since the U.S. population comprised 69% of the total GAL-INT-10 study population, an analysis of CIBIC-plus in the U.S. population only, stratified by study sites, was performed. This analysis examined the treatment effect in the large and more homogeneous U.S. population.

The percentage of CIBIC-plus responders for both GAL-ER and GAL-IR are almost identical (65%) and greater than placebo (58%), indicating very similar performance between GAL-ER and GAL-IR within the GAL-INT-10 U.S. population (Table 3.1.5.1). The CIBIC-plus responder rate for GAL-ER is comparable to that observed in previous AD studies. However, the responder rate in the placebo group (58%) is higher than that observed in previous AD studies (approximately 50%).

Statistical significance is achieved in the U.S. population for the comparison between GAL-ER and placebo ($p=0.026$) as well as the comparison between GAL-IR and placebo ($p=0.029$) (Table 3.1.5.1). These are based on a CMH stratified by study site in the U.S. population. Small study sites with 3 or fewer subjects per treatment cell were pooled.

In summary, in the U.S. population that comprised 69% of the GAL-INT-10 population, CIBIC-plus performed as would be expected. A large proportion of more mildly demented subjects in the non-U.S. sites reduced the CIBIC-plus sensitivity by inflating the placebo response.

**Table 3.1.5.1 Summary of CIBIC-plus at Week 26 –
LOCF Data (Study GAL-INT-10: U.S. Population)**

	PLACEBO	GAL-ER	GAL-IR
U.S. Subjects	N 204	N 202	N 200
CIBIC-plus			
Markedly improved	1.5%	1.0%	0.5%
Moderately improved	2.0%	3.0%	5.5%
Mildly improved	11.8%	16.8%	12.5%
No change	42.6%	44.5%	47.0%
Mildly worse	26.0%	24.3%	23.0%
Moderately worse	13.2%	8.9%	10.0%
Markedly worse	2.9%	1.5%	1.5%
p-value^a (vs Placebo)		0.026	0.029
^a CMH model stratified by study site			

Reviewer's Comment:

P-values for CMH without any stratification are .0478 for GAL-ER vs. Placebo, and .0823 for GAL-IR vs. Placebo, respectively. There are 19 sites after pooling small sites.

Since USA is a subgroup, interpretation of the significant results should be careful.

3.2 Evaluation of Safety

None.

4. Findings in Special/Subgroup Populations

None.

5. Summary and Conclusions

5.1 Statistical Issues and Collective Evidence

CMH stratified by region or by country is more commonly used method in analyzing CIBIC-plus when the study is an international study, and more appropriate to adjust for country effect. Unfortunately, none of them is significant. The significance result of CMH stratified by site is questionable since there is no pre-specified pooling plan. Furthermore, it is not clear whether the significance result of CMH stratified by site is driven by one extremely large pooled site.

Since none of CMH without stratification, CMH stratified by country, and CMH stratified by disease severity (MMSE, ADAS-cog/11, and prior cholinomimetic use) is significant for GAL-ER vs. Placebo, the significance result of CMH stratified by both disease severity and country might be due to interaction.

In USA population, CMH stratified by site shows significance and CMH without stratification also shows significant for GAL-ER vs. Placebo, but one should interpret the subgroup result with caution since it is a subgroup analysis.

5.2 Conclusions and Recommendations

The post-hoc re-analyses of CIBIC-plus at Week 26 doesn't provide convincing evidence that CIBIC-plus at Week 26 is significantly different between GAL-ER vs. Placebo, and GAL-IR vs. Placebo.

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

/s/

Kun He
7/12/04 09:43:19 AM
BIOMETRICS

Kun Jin
7/12/04 09:50:01 AM
BIOMETRICS

James Hung
7/12/04 11:35:33 AM
BIOMETRICS



U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH
OFFICE OF PHARMACOEPIDEMIOLOGY AND STATISTICAL SCIENCE
OFFICE OF BIOSTATISTICS

STATISTICAL REVIEW AND EVALUATION

Clinical Studies

NDA/Serial Number: 21-615/000
Drug Name: Reminyl®(galantamine) Extended Release
Capsules
Indication: Alzheimer
Applicant: Johnson & Johnson
Date: 2/24/2002
Review Priority: Standard

Biometrics Division: I (HFD 710)
Statistical Reviewer: Kun He
Concurring Reviewers: Kun Jin, , Ph.D., Team Leader
Kooros Mahjoob, Ph.D., Deputy Director

Medical Division: Neuropharmacological Drug Products (HFD 120)
Clinical Team: Ranjit Mani, M.D., Clinical Reviewer
Russell Katz, M.D., Director
Project Manager: Merrill Mille, R. Ph.

Keywords: Alzheimer, Reminyl, ANCOVA

TABLE OF CONTENTS

1. Executive Summary3

1.1 Conclusions and Recommendations3

1.2 Brief Overview of Clinical Studies3

1.3 Statistical Issues and Findings4

2. Introduction5

2.1 Overview5

2.2 Data Sources6

3. Statistical Evaluation6

3.1 Evaluation of Efficacy6

 3.1.1 Objective of Study GAL-INT-106

 3.1.2 Study Design6

 3.1.3 Efficacy Measures8

 3.1.4 Statistical Analysis Plan8

 3.1.5 Study Population9

 3.1.6 Applicant’s Efficacy Results15

 3.1.6.1 ADAS-cog/11 Change from Baseline to Week 2616

 3.1.6.2 CIBIC-Plus Score at Week 2617

 3.1.6.3 ADCS-ADL Change from Baseline to Week 2618

 3.1.7 Reviewer’s Analysis19

3.2 Evaluation of Safety19

4. Findings in Special/Subgroup Populations19

4.1 Gender, Race, and Age19

4.2 Other Special/Subgroup Populations20

 4.2.1 Regional analysis20

5. Summary and Conclusions21

5.1 Statistical Issues and Collective Evidence21

5.2 Conclusions and Recommendations21

Statistical Review and Evaluation

1. Executive Summary

1.1 Conclusions and Recommendations

The data and analyses from the current submission doesn't support the applicant's claim, because CIBIC-plus, one of the two primary endpoints ADAS-cog/11 and CIBIC-plus, didn't achieve statistical significance.

1.2 Brief Overview of Clinical Studies

The primary objective of this study was to evaluate the safety and efficacy (as measured by Alzheimer's Disease Assessment Scale: sum of 11 cognitive items [ADAS-cog/11] and Clinician's Interview Based Impression of Change – Plus Caregiver Input [CIBIC-plus]) of a flexible dosing regimen (16 or 24 mg/day) of galantamine controlled-release (CR) compared with placebo in subjects with mild to moderate Alzheimer's disease. Secondary analysis includes comparisons for efficacy measurements between galantamine IR and placebo and between galantamine CR and IR. This double-blind, parallel-group, placebo-controlled, flexible-dose study was conducted in the U.S., Australia, Canada, South Africa, and New Zealand. Following a 4-week, single-blind, placebo run-in period, subjects were randomized to receive placebo, galantamine IR, or galantamine CR in double-blind fashion for 26 weeks. Subjects in the galantamine treatment groups received 4 weeks of galantamine CR 8 mg q.d. or IR 4 mg b.i.d., followed by 4 weeks of galantamine CR 16 mg q.d. or IR 8 mg b.i.d. Based on safety and tolerability, the galantamine dose could be increased to CR 24 mg q.d. or IR 12 mg b.i.d. at Week 8, and could be reduced to CR 16 mg q.d. or IR 8 mg b.i.d. at Week 12. The dose chosen at the end of Week 12 was fixed for the remainder of the study. 971 subjects were randomized with 965 subjects analyzed for safety, and 925 subjects analyzed for efficacy (Intent-to-Treat [ITT] Analysis Set).

Efficacy was evaluated by the ADAS-cog/11, CIBIC-plus, ADCS-ADL, and NPI measurements at baseline and at Weeks 8, 12, and 26. The primary efficacy analyses were to compare galantamine CR with placebo with respect to change in ADAS-cog/11 scores from baseline to Week 26 and CIBIC-plus scores at Week 26. The changes in the ADAS-cog/11 score were analyzed using analysis of variance (ANOVA) models with treatment and country (U.S. vs. non-U.S.) factors. The Cochran-Mantel-Haenszel (CMH) test using modified ridit scores, derived from rank scores (the Van Elteren test) controlling for country (U.S. vs. non-U.S.) effect, was applied to compare the distributions between each pair of the treatments for the CIBIC-plus score. The primary efficacy analysis was based on the observed case (OC) data for the ITT analysis set. *(In the protocol stage, the Agency pointed out to the applicant that LOCF is commonly used for the primary analysis for this indication.)* Last observation carried forward (LOCF) analysis for the ITT analysis set was used for the primary analyses. Treatment with galantamine CR and IR led to statistically significant improvements in the primary efficacy endpoint (ADAS-cog/11) compared with placebo at Week 26.

Both galantamine CR and IR treatments were numerically better but not statistically different from that of the placebo group in maintaining global function assessed by CIBIC-plus scores at Week 26.

P-values for ADAS-cog/11 are .0001 using LOCF and OC. P-values for CIBIC-plus are .216 using LOCF and .0859 using OC, respectively. The decision rule is to have significance on both ADAS-cog/11 and CIBIC-plus simultaneously.

1.3 Statistical Issues and Findings

The primary analyses include ANOVA on change from baseline to Week 26 in ADAS-cog/11 and CMH on CIBIC-plus at Week 26. P-values for ADAS-cog/11 are .0001 using LOCF and OC. However, p-values for CIBIC-plus are .216 using LOCF and .086 using OC, respectively. The decision rule is to have significance on both ADAS-cog/11 and CIBIC-plus simultaneously, so the result of the current trial didn't achieve statistical significance.

Appears This Way
On Original

2. Introduction

Texts, tables, and figures presented in Section 3.1.1 to 3.1.6 are mainly from the applicant's submission.

2.1 Overview

Natural and synthetic galantamine has been filed and approved in all major markets, except Japan (where Phase 3 development is ongoing), in IR tablet and oral liquid formulations for prescribed twice-daily use. The galantamine-IR development program was a full program for a new chemical entity. In that program, safety and efficacy in patients with mild to moderate AD were established by data from 5 placebo-controlled studies of 3- to 6-month duration.

The current submission presents data to support the approval of a CR capsule form of the synthetic galantamine product. The presentation of data from a single Phase 3 study was agreed upon by the Medicinal Products Agency (MPA) and the Food and Drug Administration (FDA). There were no Phase 2 studies since dose-finding studies were not required. Dose regimens used in the galantamine-CR program were based on findings in galantamine-IR studies.

In the galantamine-CR program, 7 Phase 1 studies were conducted. Two studies, GAL-BEL-17 and GAL-BEL-18, were conducted using pilot formulations of galantamine CR that were not selected for further development. Pharmacokinetic findings are summarized for 5 Phase 1 studies of galantamine CR conducted in a total of 109 healthy subjects (GAL-BEL-19, GAL-BEL-20, GAL-NED-8, GAL-NED-9, and GAL-NED-12).

The pivotal Phase 3 program consisted of GAL-INT-10, a placebo and active-controlled, double-blind study of galantamine CR in 971 randomized subjects with mild to moderate AD. In this review, GAL-INT-10 will be discussed.

The primary objective of this study was to evaluate the safety and efficacy (as measured by Alzheimer's Disease Assessment Scale: sum of 11 cognitive items [ADAS-cog/11] and Clinician's Interview Based Impression of Change – Plus Caregiver Input [CIBIC-plus]) of a flexible dosing regimen (16 or 24 mg/day) of galantamine controlled-release (CR) compared with placebo in subjects with mild to moderate Alzheimer's disease. Secondary analysis includes comparisons for efficacy measurements between galantamine IR and placebo and between galantamine CR and IR. This double-blind, parallel-group, placebo-controlled, flexible-dose study was conducted in the U.S., Australia, Canada, South Africa, and New Zealand. Following a 4-week, single-blind, placebo run-in period, subjects were randomized to receive placebo, galantamine IR, or galantamine CR in double-blind fashion for 26 weeks. Subjects in the galantamine treatment groups received 4 weeks of galantamine CR 8 mg q.d. or IR 4 mg b.i.d., followed by 4 weeks of galantamine CR 16 mg q.d. or IR 8 mg b.i.d. Based on safety and tolerability, the galantamine dose could be increased to CR 24 mg q.d. or IR 12 mg b.i.d. at Week 8, and could be reduced to CR 16 mg q.d. or IR 8 mg b.i.d. at Week

12. The dose chosen at the end of Week 12 was fixed for the remainder of the study. 971 subjects were randomized with 965 subjects analyzed for safety, and 925 subjects analyzed for efficacy (Intent-to-Treat [ITT] Analysis Set).

2.2 Data Sources

The path to the CDER Electronic Document Room (EDR) is:

\\CDSESUB1\N21615\N_000\2003-02-24

3. Statistical Evaluation

3.1 Evaluation of Efficacy

3.1.1 Objective of Study GAL-INT-10

The primary objective of this study was to evaluate the safety and efficacy (as measured by ADAS-cog/11 and CIBIC-plus) of a flexible dosing regimen (16 or 24 mg/day) of galantamine CR compared with placebo in subjects with mild to moderate AD.

Secondary objectives were to evaluate the effects of galantamine CR and IR treatment on those subjects with regard to the activities of daily living and behavior, using the ADCS-ADL and NPI scores, as well as ADAS-cog/13, ADAS-cog/10, and ADAS-cog/mem scores, and to estimate the difference in effect between the galantamine CR and IR treatment groups.

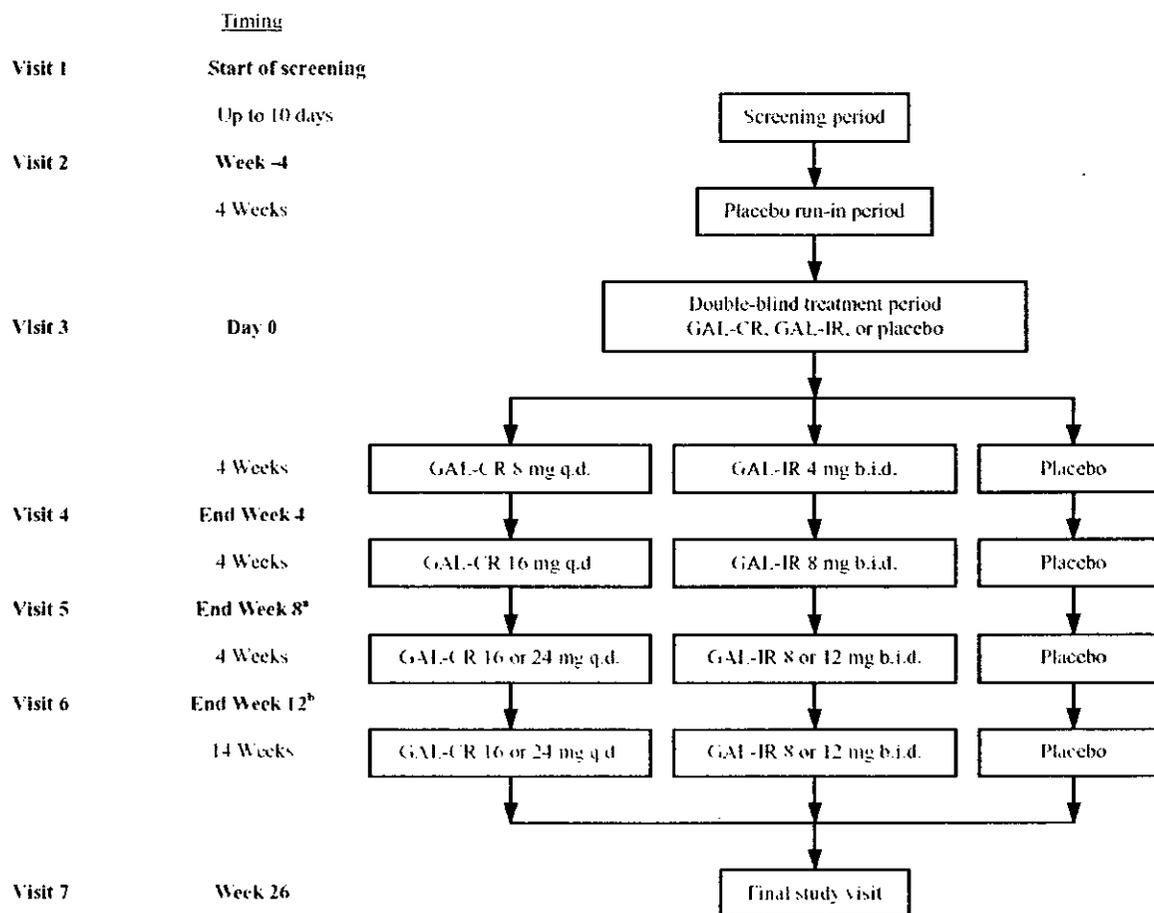
3.1.2 Study Design

This was a 26-week, randomized, double-blind, parallel-group, placebo-controlled, flexible-dose study conducted in Australia, Canada, New Zealand, South Africa, and the U.S., comparing the safety, tolerability, and efficacy of galantamine CR with that of placebo in subjects with mild to moderate AD. Following a 4-week, single-blind, placebo run-in period, subjects were randomized to receive galantamine CR, galantamine IR, or a matching placebo in a double-blind fashion for 26 weeks.

Subjects in the galantamine-IR group received 4 weeks of 8 mg/day galantamine (4 mg b.i.d.) followed by 4 weeks of 16 mg/day (8 mg b.i.d.). After Week 8, at the investigator's discretion based on safety and tolerability, the dose could be increased to 24 mg/day (12 mg b.i.d.). At Week 12, subjects receiving 24 mg/day (12 mg b.i.d.) were re-evaluated, and the dose could have been reduced to 16 mg/day (8 mg b.i.d.) based on safety and tolerability. The dose chosen at the end of Week 12 was fixed for the remainder of the study.

A diagrammatic representation of the study design and the treatment received is provided in Figure 3.1.2.1. Office visits were scheduled at screening of subjects for entry into the study (Visit 1), at the beginning of the single-blind treatment phase (Visit 2), at baseline (Visit 3), and at Weeks 4, 8, 12, and 26 (Visits 4, 5, 6, and 7). Office visits consisted of physical examinations including neurological examinations (Visits 1, 3, 6, and 7), dose adjustment evaluations (Visits 4, 5, and 6), and the recording of the efficacy endpoint scores from the following tests: ADAS-cog (Visits 1, 3, 5, 6, and 7), CIBIC-plus (Visits 3, 5, 6, and 7), and ADCS-ADL and NPI (Visits 3, 5, 6, and 7). Safety evaluations included recording adverse event reports, conducting physical examinations, monitoring vital signs and ECG, and evaluating laboratory evaluations. Figure 3.1.2.1 is adapted from Study Report Figure 1.

Figure 3.1.2.1 Overview of Study Design



^aDose increase based on safety and tolerability only permitted at end of Week 8.

^bDose reduction based on safety and tolerability only permitted at end of Week 12, and the dose at Week 12 was fixed thereafter.

There were 2 amendments to the final GAL-INT-10 protocol.

The first amendment dated 10 May 2001 resulted in clarifications of cardiovascular and concomitant medication exclusion criteria, withdrawal criteria, separation of Visits 1 and 2, and DNA sampling procedures, elimination of CIBIC-plus rater access to subject source documentation, change from digital to manual measurement of ECG intervals, updating of the Pharmacogenomics Supplement, and a change in the reporting pregnancy as an immediately reportable adverse event. At the time of Amendment implementation, 636 subjects had been screened (signed informed consent form [ICF] at Visit 1) in the study.

The second amendment dated 18 October 2001 resulted in a reduction in sample size from approximately 1020 randomized subjects (340/group) to approximately 885 randomized subjects (295/group) as the power of the study to differentiate between galantamine CR and placebo (15% expected difference) in terms of CIBIC-plus was reduced from 95% to 90%, while the overall (global) statistical power was approximately 90% for the dual endpoints (ADAS-cog and CIBIC-plus). The amendment was implemented due to increasing use of the commercially available REMINYL formulation. At the time of amendment implementation, 1243 subjects had been screened (signed ICF at Visit 1) in the study. Subsequent to these amendments, an additional 272 subjects were screened.

3.1.3 Efficacy Measures

The primary efficacy endpoints were change in Alzheimer's Disease Assessment Scale (ADAS-cog/11) score from baseline to Week 26, and Clinician's Interview Based Impression of Change-Plus (CIBIC-plus) score at Week 26. The ADAS-cog was performed at Visits 1, 3, 5, 6, and 7 (screening, baseline, Weeks 8, 12, and 26, or upon premature discontinuation of study). The CIBIC-plus was performed at Visits 5, 6, and 7 (Weeks 8, 12, and 26, or upon premature discontinuation of the study). Secondary endpoints include Alzheimer's Disease Cooperative Study (ADCS-ADL) Scale, Neuropsychiatric Inventory (NPI), ADAS-cog/13, ADAS-cog/10, and ADAS-cog/mem scores from baseline to week 26.

3.1.4 Statistical Analysis Plan

The primary efficacy analyses, based on the observed case data (ITT population), were to compare galantamine CR with placebo with respect to change in ADAS-cog/11 score from baseline to Week 26 and CIBIC-plus score at Week 26. The primary comparison is the GAL CR versus placebo for both primary efficacy variables.

Change from baseline to Week 26 in ADAS-cog/11 score: an analysis of variance (ANOVA) model with treatment and pooled country factors (U.S. vs non-U.S.) was used for comparison of the least-squares (LS) means between the treatments. The country factor used had 2 levels, i.e., U.S. and non-

U.S..

CIBIC-plus score at Week 26: the Cochran-Mantel-Haenszel (CMH) test statistic using modified ridit scores, derived from rank scores (the Van Elteren test) controlling for country effect was applied to compare the distributions between each pair of the treatments. The country factor used had 2 levels, i.e., U.S. and non-US.

3.1.5 Study Population

Nine-hundred seventy-one subjects were randomized to receive galantamine CR (n=320), IR (n=327), or placebo (n=324) at 93 study sites in 5 countries: the U.S., Australia, Canada, South Africa, and New Zealand. The study was conducted from 08 February 2001 to 15 July 2002. Of the 971 subjects randomly assigned to study treatment, 965 subjects received at least 1 dose of double-blind study medication and comprise the all randomized and treated population. The majority (69%) of all randomized and treated subjects were enrolled at study sites in the U.S. Table 3.1.5.1 is adapted from Study Report Table 9.

**Table 3.1.5.1 Subjects Distribution by Country
(All Randomized and Treated Subjects)**

Country	PLACEBO (N=320) n (%)	GAL-IR (N=326) n (%)	GAL-CR (N=319) n (%)	Total (N=965) n (%)
U.S.	221 (69)	221 (68)	221 (69)	663 (69)
South Africa	35 (11)	39 (12)	34 (11)	108 (11)
Canada	33 (10)	32 (10)	34 (11)	99 (10)
Australia	28 (9)	30 (9)	27 (8)	85 (9)
New Zealand	3 (1)	4 (1)	3 (1)	10 (1)

Of the 971 subjects randomized, 768 (79%) completed the 26-week study. The most common reasons for discontinuations during the double-blind treatment phase were adverse events (7%) and withdrawal of consent (6%), which occurred at similar rates for the 3 treatment groups (Table 3.1.5.2, adapted from Study Report Table 10). Figure 3.1.5.1 is adapted from Study Report Figure 2.

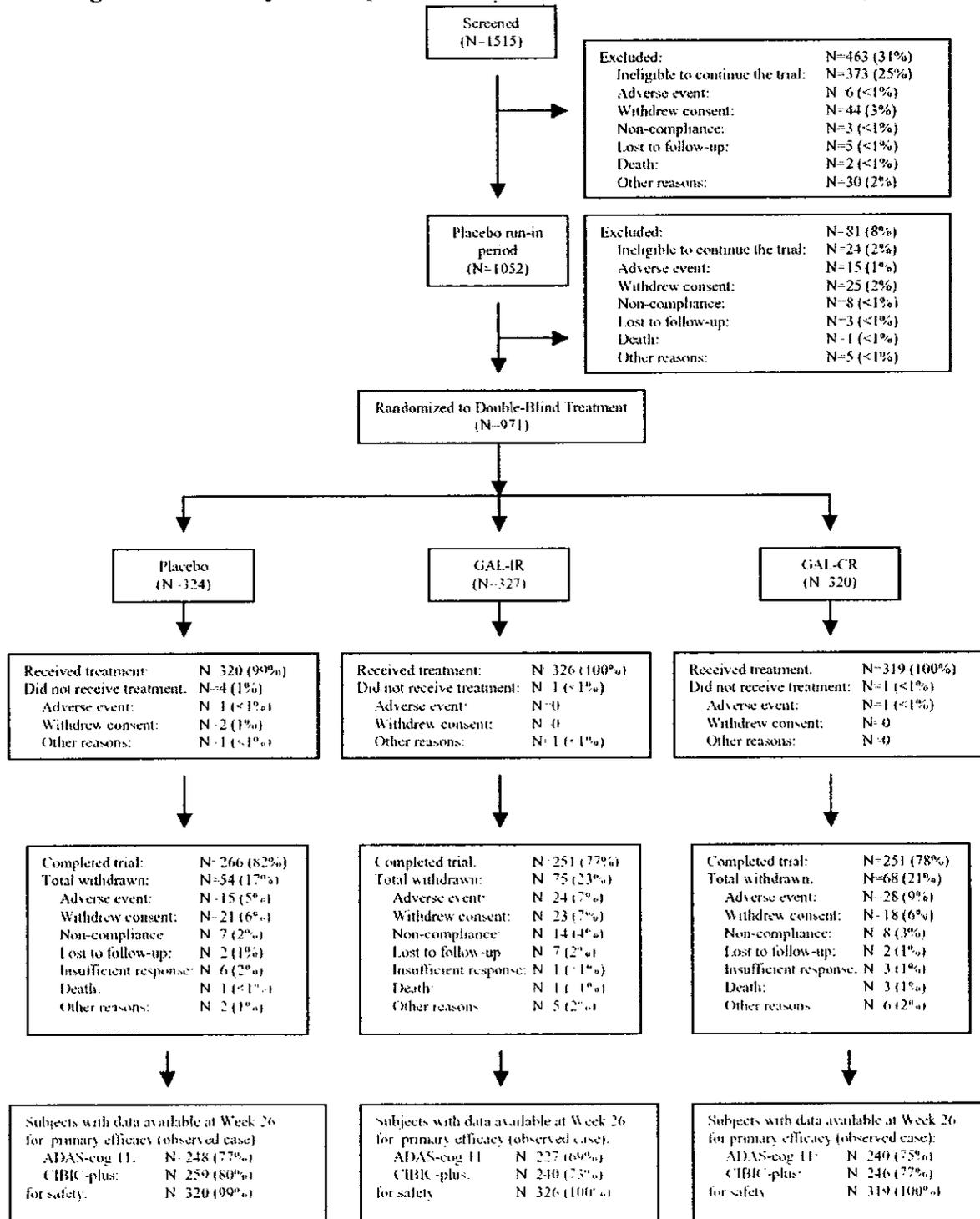
Table 3.1.5.2 Study Termination Reasons (All Randomized Subjects)

All Randomized Subjects Status	PLACEBO (N=324)	GAL-IR (N=327)	GAL-CR (N=320)	Total (N=971)
Termination Reasons	n (%)	n (%)	n (%)	n (%)
Randomized and treated	320 (99)	326 (100)	319 (100)	965 (99)
<u>Completed</u>	266 (82)	251 (77)	251 (78)	768 (79)
<u>Discontinued</u>	54 (17)	75 (23)	68 (21)	197 (20)
Adverse event	15 (5)	24 (7)	28 (9)	67 (7)
Subject withdrew consent	21 (6)	23 (7)	18 (6)	62 (6)
Subject non-compliant	7 (2)	14 (4)	8 (3)	29 (3)
Subject lost to follow-up	2 (1)	7 (2)	2 (1)	11 (1)
Insufficient response	6 (2)	1 (<1)	3 (1)	10 (1)
Death	1 (<1)	1 (<1)	3 (1)	5 (1)
Subject ineligible to continue the study	0	0	4 (1)	4 (<1)
Other	2 (1)	5 (2)	2 (1)	9 (1)
Randomized and not treated	4 (1)	1 (<1)	1 (<1)	6 (1)
<u>Discontinued</u>	4 (1)	1 (<1)	1 (<1)	6 (1)
Adverse event	1 (<1)	0	1 (<1)	2 (<1)
Subject withdrew consent	2 (1)	0	0	2 (<1)
Subject ineligible to continue the study	1 (<1)	0	0	1 (<1)
Subject non-compliant	0	1 (<1)	0	1 (<1)

Percentage for each category in a group was calculated based on all randomized subjects for that group as denominator.

Appears This Way
On Original

Figure 3.1.5.1 Subject Completion and Withdrawal Information (All Subjects)



Demographic and baseline characteristics of subjects included in the all randomized and treated population are summarized by treatment in Table 3.1.5.4. Most subjects were white (91%) and female (64%). Subjects were between 48 and 93 years of age and had a mean age of 76.5 years. Of the 965 subjects in the all randomized and treated population, 886 (92%) were 65 years of age or older, including 99 subjects who were older than 85 years of age. The subjects had a mean weight of 68.23 kg, and a mean height of 163.2 cm. Most (93%) of the subjects were nonsmokers. The treatment groups were generally well matched with regard to sex, race, history of smoking, age, weight, and height. Subjects with AD and MMSE scores ranging from 10 to 24 and ADAS-cog/11 scores ≥ 18 at screening were to be entered in this study. The subjects had a median MMSE score at screening of 18 and a median ADAS-cog/11 score at screening of 25. Median MMSE and ADAS-cog/11 scores were comparable across the treatment groups. Subjects were categorized according to the screening MMSE scores (10 to 22) used in the GAL-USA-10 study. The number of subjects with screening MMSE scores >22 , (indicating a milder form of AD) were higher for the placebo (17%) than for the galantamine-IR (15%), and the galantamine-CR groups (13%).

The percentage of subjects with ADAS-cog/11 score <18 at baseline was higher than those at screening in all 3 treatment groups. This difference was more pronounced in the placebo group (18% vs. 1%) than in the galantamine-IR (14% vs. 1%) and galantamine-CR (17% vs. 3%) treatment groups. Table 3.1.5.3 is adapted from Study Report Table 11.

Appears This Way
On Original

Table 3.1.5.3 Demographic, Baseline Characteristics, and MMSE and ADAS-cog/11 Data at Screening (All Randomized and Treated Subjects)

	PLACEBO (N=320)	GAL-IR (N=326)	GAL-CR (N=319)	Total (N=965)
Sex, n (%)				
N	320	326	319	965
Male	115 (36)	118 (36)	114 (36)	347 (36)
Female	205 (64)	208 (64)	205 (64)	618 (64)
Race, n (%)				
N	320	326	319	965
Black	12 (4)	12 (4)	9 (3)	33 (3)
Caucasian	289 (90)	293 (90)	297 (93)	879 (91)
Hispanic	6 (2)	6 (2)	2 (1)	14 (1)
Oriental	7 (2)	5 (2)	9 (3)	21 (2)
Other	6 (2)	10 (3)	2 (1)	18 (2)
Smoking, n (%)				
N	319	325	318	962
Yes	24 (8)	22 (7)	19 (6)	65 (7)
No	295 (92)	303 (93)	299 (94)	897 (93)
Age, years				
N	320	326	319	965
Category, n (%)				
<65	27 (8)	30 (9)	22 (7)	79 (8)
65-85	254 (79)	268 (82)	265 (83)	787 (82)
>85	39 (12)	28 (9)	32 (10)	99 (10)
Mean (SD)	76.3 (8.03)	76.5 (7.77)	76.6 (7.64)	76.5 (7.81)
Median	77.0	78.0	77.0	77.0
Range	48 - 92	49 - 92	55 - 93	48 - 93
Weight at baseline, kg				
N	319	325	319	963
Mean (SD)	67.78 (14.591)	68.29 (15.857)	68.60 (14.159)	68.23 (14.881)
Median	66.00	67.30	67.30	67.00
Range	37.4 - 119.5	37.0 - 136.4	35.8 - 120.9	35.8 - 136.4
Height, cm				
N	320	326	318	964
Mean (SD)	162.7 (10.03)	162.7 (10.85)	164.1 (10.23)	163.2 (10.39)
Median	162.6	161.0	162.6	162.0
Range	132 - 189	122 - 207	142 - 191	122 - 207
Sum of MMSE at screening				
N	320	326	319	965
Category, n (%)				
10-22	265 (83)	276 (85)	276 (87)	817 (85)
>22	55 (17)	50 (15)	43 (13)	148 (15)
Mean (SD)	18.08 (4.082)	17.80 (4.138)	17.96 (3.966)	17.95 (4.061)
Median	19.00	18.00	18.00	18.00
Range	10 - 24	10 - 24	10 - 24	10 - 24
ADAS-cog/11 score at screening				
N	317	323	315	955
Category, n (%)				
<18	3 (1)	3 (1)	8 (3)	14 (1)
≥18	314 (99)	320 (99)	307 (97)	941 (99)
Mean (SD)	26.97 (8.300)	27.80 (8.486)	26.73 (8.036)	27.17 (8.282)
Median	25.00	26.00	25.00	25.00
Range	13 - 57	13 - 55	12 - 52	12 - 57

Note: Percentages calculated with the number of subjects by parameter as denominator.

(continued)

	PLACEBO (N=320)	GAL-IR (N=326)	GAL-CR (N=319)	Total (N=965)
ADAS-cog/11 score at baseline				
N	316	320	314	950
Category, n (%)				
<18	58 (18)	44 (14)	52 (17)	154 (16)
≥18	258 (82)	276 (86)	262 (83)	796 (84)
Mean (SD)	26.23 (9.588)	27.47 (9.935)	26.43 (9.303)	26.71 (9.620)
Median	25.00	26.00	25.00	25.00
Range	9 - 62	9 - 58	7 - 55	7 - 62

Note: Percentages calculated with the number of subjects by parameter as denominator.

The AD history for subjects included in the all randomized and treated population is summarized by treatment group in Table 3.1.5.4. (Adapted from Study Report Table 12). The treatment groups were generally well matched with regard to age at onset and duration since diagnosis of cognitive problems. The age at diagnosis and duration since diagnosis of AD as well as number of subjects with first-degree relatives with AD were also similar across the treatments groups. The number of subjects who had taken cholinomimetics before study enrollment was similar across treatments.

Appears This Way
On Original

Table 3.1.5.4 Alzheimer's Disease History (All Randomized and Treated Subjects)

	PLACEBO (N=320)	GAL IR (N=326)	GAL CR (N=319)	Total (N=965)
Age at onset of cognitive problems				
N	319	326	319	964
Mean (SD)	72.3 (8.30)	72.0 (8.30)	72.6 (8.12)	72.3 (8.24)
Median	73.0	73.0	74.0	73.0
Range	44 - 89	45 - 89	49 - 91	44 - 91
Duration (yrs) since diagnosis of cognitive problems				
N	319	326	319	964
Mean (SD)	4.00 (2.439)	4.48 (2.674)	4.07 (2.501)	4.18 (2.547)
Median	3.40	3.90	3.40	3.50
Range	0.5 - 15	0.7 - 16	0.5 - 15	0.5 - 16
Age at diagnosis of probable Alzheimer's disease				
N	320	326	319	965
Mean (SD)	75.0 (8.32)	75.2 (7.97)	75.4 (7.85)	75.2 (8.04)
Median	76.0	76.0	76.0	76.0
Range	48 - 92	48 - 91	51 - 93	48 - 93
Duration (yrs) since diagnosis of probable Alzheimer's disease				
N	320	326	319	965
Mean (SD)	1.28 (1.584)	1.24 (1.474)	1.23 (1.573)	1.25 (1.542)
Median	0.60	0.65	0.60	0.60
Range	0.0 - 7.5	0.0 - 6.3	0.0 - 9.2	0.0 - 9.2
First-degree relatives with Alzheimer's disease, n (%)				
N	318	325	317	960
Yes	96 (30)	86 (26)	90 (28)	272 (28)
No	222 (70)	239 (74)	227 (72)	688 (72)
Subject taken cholinomimetics, n (%)				
N	320	326	318	964
Yes	150 (47)	157 (48)	153 (48)	460 (48)
No	170 (53)	169 (52)	165 (52)	504 (52)

3.1.6 Applicant's Efficacy Results

Summaries of efficacy data were based on the ITT population, defined as all randomized and treated subjects who received at least 1 dose of study medication and provided at least 1 postbaseline

primary efficacy measurement. Observed case data was used for the primary efficacy analysis. Two imputation methods were used for missing data at visits including Week 26. For the missing data at postbaseline visits in the ITT subjects, the last observation carried forward (LOCF) method was used; for the all randomized subjects set, the classical intent-to-treat (CITT) method was used. For LOCF and CITT data, the endpoint was defined as the last available observation up to 14 days after the last dose of study medication.

Of the 971 subjects randomly assigned to 1 of the 3 treatment groups (all randomized analysis set), 965 (99.4%) received the study medication (all randomized and treated analysis set). The 6 subjects who did not receive study medication were distributed as follows: 4 subjects in the placebo group and 1 subject in each of the galantamine groups. These subjects were not included in the all randomized and treated analysis set. All subjects (n=8) from 1 site in the U.S. were excluded from all efficacy analyses before database lock as the site failed to adhere to good clinical practices (GCPs). Nine hundred twenty-five subjects were included in the ITT analysis as 32 subjects were excluded because they did not present with postbaseline primary efficacy data.

3.1.6.1 ADAS-cog/11 Change from Baseline to Week 26

Both galantamine treatments were statistically better than placebo in improving cognition based on the change from baseline in ADAS-cog/11 at Week 26 for the OC and LOCF data analyses. See Tables 3.1.6.1.1 (adapted from Study Report Table 16) and 3.1.6.1.2 (adapted from Study Report Table 15).

Table 3.1.6.1.1 ADAS-cog/11 Change from Baseline (ITT-LOCF)

Timepoint	PLACEBO			GAL-IR			GAL-CR			P value ^a
	N	Mean (SE)	Change (SE)	N	Mean (SE)	Change (SE)	N	Mean (SE)	Change (SE)	
Baseline										
LOCF	305	26.1 (0.54)	--	306	27.3 (0.55)	--	300	26.3 (0.54)	--	
CITT	316	26.1 (0.54)	--	319	27.5 (0.55)	--	313	26.4 (0.53)	--	
Week 8										
LOCF	293	25.9 (0.63)	0.0 (0.30)	294	25.4 (0.57)	-1.7 (0.29)	287	24.7 (0.57)	-1.5 (0.30)	<0.001
CITT	316	26.2 (0.61)	0.0 (0.30)	319	25.9 (0.57)	-1.7 (0.29)	313	25.0 (0.56)	-1.5 (0.30)	<0.001
Week 12										
LOCF	296	26.0 (0.64)	0.2 (0.31)	296	24.5 (0.56)	-2.5 (0.30)	290	24.2 (0.56)	-2.0 (0.31)	<0.001
CITT	316	26.3 (0.62)	0.2 (0.31)	319	25.2 (0.56)	-2.5 (0.30)	313	24.6 (0.55)	-2.0 (0.31)	<0.001
Endpoint ^b										
LOCF	296	27.0 (0.67)	1.2 (0.33)	296	25.4 (0.62)	-1.6 (0.36)	291	24.9 (0.62)	-1.3 (0.31)	<0.001
CITT	316	27.2 (0.65)	1.2 (0.33)	319	26.0 (0.62)	-1.6 (0.36)	313	25.3 (0.60)	-1.3 (0.31)	<0.001

^aPairwise comparison for no difference between GAL-CR and Placebo from ANOVA model with factors Treatment and Pooled Country (type III SS).

^bThe endpoint was defined as the last available observation up to 14 days after the last dose of study medication. GAL-IR vs. Placebo at endpoint: p=0.001 (LOCF-CITT)

Table 3.1.6.1.2 ADAS-cog/11 Change from Baseline (ITT-OC)

Timepoint	PLACEBO			GAL-IR			GAL-CR			P value ^a
	N	Mean (SE)	Change (SE)	N	Mean (SE)	Change (SE)	N	Mean (SE)	Change (SE)	
Baseline	305	26.1 (0.54)	--	306	27.3 (0.55)	--	300	26.3 (0.54)	--	
Week 8	289	25.8 (0.63)	0.0 (0.30)	286	25.4 (0.58)	-1.7 (0.30)	284	24.6 (0.58)	-1.5 (0.30)	<0.001
Week 12	275	25.9 (0.66)	0.0 (0.32)	268	24.0 (0.57)	-2.6 (0.31)	269	23.9 (0.57)	-2.2 (0.32)	<0.001
Week 26	248	26.4 (0.72)	1.3 (0.36)	227	24.7 (0.69)	-1.8 (0.42)	240	24.8 (0.69)	-1.4 (0.34)	<0.001

^aPairwise comparison for no difference between GAL-CR and Placebo from ANOVA model with factors Treatment and Pooled Country (type III SS).

GAL-IR vs. Placebo at Week 26: p<0.001.

3.1.6.2 CIBIC-Plus Score at Week 26

Both galantamine treatment groups were numerically better but not statistically different from placebo in improving global functioning based on CIBIC-plus scores at Week 26 for the OC and LOCF data. See Tables 3.1.6.2.1 (adapted from Study Report Table 18) and 3.1.6.2.2 (adapted from Study Report Table 17).

Table 3.1.6.2.1 CIBIC-Plus at Week 26 (ITT-LOCF)

7-Point Category	PLACEBO		GAL-IR		GAL-CR		P value ^d
	n (%) ^b	Cum %	n (%) ^b	Cum %	n (%) ^b	Cum %	
LOCF/CITT at Endpoint ^b	301		302		296		
Markedly improved	3 (1.0)	(1.0)	3 (1.0)	(1.0)	3 (1.0)	(1.0)	
Moderately improved	11 (3.7)	(4.7)	15 (5.0)	(6.0)	14 (4.7)	(5.7)	
Mildly improved	48 (15.9)	(20.6)	46 (15.2)	(21.2)	49 (16.6)	(22.3)	
No change	111 (36.9)	(57.5)	127 (42.1)	(63.2)	114 (38.5)	(60.8)	
Mildly worse	80 (26.6)	(84.1)	78 (25.8)	(89.1)	81 (27.4)	(88.2)	
Moderately worse	41 (13.6)	(97.7)	30 (9.9)	(99.0)	29 (9.8)	(98.0)	
Markedly worse	7 (2.3)	(100.0)	3 (1.0)	(100.0)	6 (2.0)	(100.0)	0.216

^bGAL-CR vs. Placebo comparison using the Van Elteren test controlling for Pooled Country.

^cThe endpoint was defined as the last available observation up to 14 days after the last dose of study medication

Cum % = cumulative percent.

Note: Percentages calculated with the number of subjects at Week 26 as denominator.

GAL-IR vs. Placebo at endpoint: p=0.144 (LOCF, CITT).

Table 3.1.6.2.2 CIBIC-Plus at Week 26 (ITT-OC)

7-Point Category	PLACEBO		GAL-IR		GAL-CR		P value ^a
	n (%)	(Cum %)	n (%)	(Cum %)	n (%)	(Cum %)	
N at Week 26	259		240		246		
Markedly improved	3 (1.2)	(1.2)	3 (1.3)	(1.3)	3 (1.2)	(1.2)	
Moderately improved	9 (3.5)	(4.6)	14 (5.8)	(7.1)	14 (5.7)	(6.9)	
Mildly improved	41 (15.8)	(20.5)	36 (15.0)	(22.1)	43 (17.5)	(24.4)	
No change	94 (36.3)	(56.8)	93 (38.8)	(60.8)	90 (36.6)	(61.0)	
Mildly worse	70 (27.0)	(83.8)	67 (27.9)	(88.8)	69 (28.0)	(89.0)	
Moderately worse	36 (13.9)	(97.7)	25 (10.4)	(99.2)	23 (9.3)	(98.4)	
Markedly worse	6 (2.3)	(100.0)	2 (0.8)	(100.0)	4 (1.6)	(100.0)	0.086

^aGAL-CR vs. Placebo comparison using the Van Elteren test controlling for Pooled Country.

Cum % = cumulative percent

Note: Percentages calculated with the number of subjects at Week 26 as denominator.

GAL-IR vs. Placebo at Week 26; p=0.223.

3.1.6.3 ADCS-ADL Change from Baseline to Week 26

GAL-CR was statistically better than placebo in improving cognition based on the change from baseline in ADCS-ADL to Week 26 for the OC, and LOCF. See Tables 3.1.6.3.1 (adapted from Study Report Table 18) and 3.1.6.3.2 (adapted from Study Report Table 19).

Table 3.1.6.3.1 ADCS-ADL Change from Baseline (ITT-LOCF)

Timepoint	PLACEBO			GAL-IR			GAL-CR			P value ^a
	N	Mean (SE)	Change (SE)	N	Mean (SE)	Change (SE)	N	Mean (SE)	Change (SE)	
Baseline										
LOCF	308	54.5 (0.87)	--	310	52.0 (0.90)	--	303	53.5 (0.88)	--	
CITT	319	54.3 (0.87)	--	323	51.9 (0.88)	--	316	53.3 (0.86)	--	
Week 8										
LOCF	299	53.8 (0.96)	-0.8 (0.45)	300	52.7 (0.91)	0.8 (0.41)	295	54.2 (0.94)	0.8 (0.40)	0.007
CITT	319	53.5 (0.93)	-0.8 (0.45)	323	52.6 (0.87)	0.8 (0.41)	316	54.0 (0.91)	0.8 (0.40)	0.007
Week 12										
LOCF	301	54.0 (0.95)	-0.6 (0.45)	301	52.9 (0.91)	0.9 (0.45)	296	53.7 (0.94)	0.3 (0.46)	0.146
CITT	319	53.7 (0.93)	-0.6 (0.45)	323	52.7 (0.87)	0.9 (0.45)	316	53.6 (0.91)	0.3 (0.46)	0.146
Endpoint ^b										
LOCF	301	52.0 (1.02)	-2.7 (0.56)	301	51.0 (0.98)	-1.0 (0.50)	296	53.3 (0.96)	0.0 (0.48)	<0.001
CITT	319	51.7 (0.99)	-2.7 (0.56)	323	51.0 (0.94)	-1.0 (0.50)	316	53.2 (0.93)	0.0 (0.48)	<0.001

^aPairwise comparison for no difference between GAL-CR and Placebo from ANOVA model with factors Treatment and Pooled Country (type III SS).

^bThe endpoint was defined as the last available observation up to 14 days after the last dose of study medication GAL-IR vs. Placebo at endpoint: p=0.018 (LOCF CITT).

Table 3.1.6.3.2 ADCS-ADL Change from Baseline (OC)

Timepoint	PLACEBO			GAL-IR			GAL-CR			P value ^a
	N	Mean (SE)	Change (SE)	N	Mean (SE)	Change (SE)	N	Mean (SE)	Change (SE)	
Baseline	308	54.5 (0.87)	--	310	52.0 (0.90)	--	303	53.5 (0.88)	--	
Week 8	294	53.8 (0.98)	-0.7 (0.45)	292	52.6 (0.93)	0.9 (0.42)	290	54.5 (0.94)	0.8 (0.41)	0.013
Week 12	281	54.2 (0.99)	-0.3 (0.46)	279	52.8 (0.95)	1.1 (0.47)	276	54.1 (0.94)	0.4 (0.48)	0.321
Week 26	258	52.4 (1.09)	-2.4 (0.60)	242	50.9 (1.12)	-1.0 (0.57)	245	53.9 (1.03)	0.0 (0.55)	0.003

^aPairwise comparison for no difference between GAL-CR and Placebo from ANOVA model with factors Treatment and Pooled Country (type III SS).
GAL-IR vs. Placebo at Week 26: p=0.088.

3.1.7 Reviewer's Analysis

The reviewer validated the applicant's analyses according to the protocol, and duplicated the p-values if keeping three digits. The applicant used OC for the primary analyses but the Agency pointed out to the applicant that LOCF is commonly used for the primary analyses in this indication. The primary analyses include ANOVA on change from baseline to Week 26 in ADAS-cog/11 and CMH on CIBIC-plus at Week 26. P-values for ADAS-cog/11 are .0001 using both LOCF and OC. However, p-values for CIBIC-plus are .216 using LOCF and .0859 using OC, respectively. The decision rule is to have significance on both ADAS-cog/11 and CIBIC-plus simultaneously, so the result of the current trial didn't achieve statistical significance.

The reviewer also validated analysis on ADCS-ADL which has p-value .0003 using LOCF and .0029 using OC. Although the test is statistically significant at .05, one should interpret the result with caution because the primary analyses, which spent all alpha, are not significant.

3.2 Evaluation of Safety

See Clinical Review by Dr. Ranjit Mani.

4. Findings in Special/Subgroup Populations

4.1 Gender, Race, and Age

Mean of change from baseline for ADAS-Cog/11 and mean of value at endpoint for CIBIC-PLUS are listed for gender and age groups. Since majority subjects are caucasian, no descriptive statistics on race is listed. A negative number indicates that GAL CR is better than PLACEBO numerically.

Table 4.1.1 ADAS-Cog/11 by Gender (ITT-LOCF)

Gender	PLACEBO			GAL CR			GAL - PLACEBO
	Change from Baseline			Change from Baseline			
	N	Mean	SD	N	Mean	SD	
Male	104	1.0	5.87	105	-1.5	5.43	-2.5
Female	192	1.2	5.59	186	-1.1	5.09	-2.2

Table 4.1.2 CIBIC-PLUS by Gender (ITT-LOCF)

Gender	PLACEBO			GAL CR			GAL - PLACEBO
	Value at endpoint			Value at endpoint			
	N	Mean	SD	N	Mean	SD	
Male	106	4.29	1.19	107	4.11	1.18	-0.18
Female	195	4.37	1.11	189	4.31	1.07	-0.06

Table 4.1.3 ADAS-Cog/11 by Age (ITT-LOCF)

Age	PLACEBO			GAL CR			GAL - PLACEBO
	Change from Baseline			Change from Baseline			
	N	Mean	SD	N	Mean	SD	
< 65	25	1.8	4.79	21	-1.8	5.83	-3.6
65-85	237	1.2	5.69	245	-1.0	5.18	-2.2
> 85	34	0.1	6.20	25	-2.8	4.84	-2.9

Table 4.1.4 CIBIC-PLUS by Age (ITT-LOCF)

Age	PLACEBO			GAL CR			GAL - PLACEBO
	Value at endpoint			Value at endpoint			
	N	Mean	SD	N	Mean	SD	
< 65	25	4.08	1.38	22	4.5	1.30	0.42
65-85	242	4.4	1.14	250	4.22	1.08	-0.18
> 85	34	4.15	0.93	24	4.21	1.28	0.6

4.2 Other Special/Subgroup Populations

4.2.1 Regional analysis

Means of change from baseline for ADAS-Cog/11 and value at endpoint for CIBIC-PLUS are listed for USA and non-USA groups.

Table 4.2.1 ADAS-Cog/11 by Region (ITT-LOCF)

Region	PLACEBO			GAL CR			GAL - PLACEBO
	Change from Baseline			Change from Baseline			
	N	Mean	SD	N	Mean	SD	
USA	201	1.5	5.72	197	-1.6	5.18	-3.1
Non-USA	95	0.3	5.51	94	-0.6	5.23	-0.9

Table 4.2.2 CIBIC-PLUS by Region (ITT-LOCF)

Region	PLACEBO			GAL CR			GAL - PLACEBO
	Change from Baseline			Change from Baseline			
	N	Mean	SD	N	Mean	SD	
USA	204	4.41	1.10	202	4.21	1.03	-0.2
Non-USA	97	4.21	1.21	94	4.31	1.27	0.1

5. Summary and Conclusions

5.1 Statistical Issues and Collective Evidence

The primary analyses include ANOVA on change from baseline to Week 26 in ADAS-cog/11 and CMH on CIBIC-plus at Week 26. P-values for ADAS-cog/11 are .0001 using both LOCF and OC. However, p-values for CIBIC-plus are .216 using LOCF and .0859 using OC, respectively. The decision rule is to have significance on both ADAS-cog/11 and CIBIC-plus simultaneously, so the result of the current trial didn't achieve statistical significance.

5.2 Conclusions and Recommendations

The data and analyses from the current submission doesn't support the applicant's claim, because CIBIC-plus, one of the two primary endpoints ADAS-cog/11 and CIBIC-plus, didn't achieve statistical significance.

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

/s/

Kun He
11/25/03 08:39:52 AM
BIOMETRICS

Kun Jin
11/25/03 11:36:42 AM
BIOMETRICS

Kooros Mahjoob
11/26/03 09:51:44 AM
BIOMETRICS