

**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:

204447Orig1s000

STATISTICAL REVIEW(S)



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

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA #: 204,447 (original NDA)
Drug Name: Vortioxetine (Lu AA21004)
Indication(s): Major Depressive Disorder
Applicant: Takeda Global Research & Development Center
Dates: Submission date: October 1, 2012
PDUFA V: October 2, 2013
Review Priority: Standard
Biometrics Division: Division of Biometrics I
Statistical Reviewers: George Kordzakhia, Ph.D.
Concurring Reviewers: Peiling Yang, Ph.D., H.M. James Hung, Ph.D.
Medical Division: Division of Psychiatry Products
Clinical Team: Jenn Sellers, M.D., Reviewer
Jing Zhang, M.D., Team Leader
Project Manager: Hiren Patel, Pharm. D.

Keywords:

NDA review, mixed models, ANCOVA.

Table of Contents

1. EXECUTIVE SUMMARY	7
2. INTRODUCTION	8
2.1 OVERVIEW.....	8
2.2 DATA SOURCES	8
3. STATISTICAL EVALUATION	9
3.1 DATA AND ANALYSIS QUALITY	9
3.2 EVALUATION OF EFFICACY	9
3.2.1 <i>Study Design and Endpoints</i>	10
3.2.1.1 Study 11492A.....	10
3.2.1.2 Study 305.....	11
3.2.1.3 Study 13267A.....	11
3.2.1.4 Study 315.....	12
3.2.1.5 Study 316.....	13
3.2.1.6 Study 12541A (elderly study).....	13
3.2.1.7 Study 11985A (maintenance study).....	14
3.2.2 <i>Statistical Methodologies</i>	15
3.2.2.1 Study 11492A (non-US multiregional, including Canada).....	15
3.2.2.2 Study 305 (conducted outside North America).....	16
3.2.2.3 Study 13267A (conducted outside North America).....	17
3.2.2.4 Study 315 (conducted in US).....	18
3.2.2.5 Study 316 (conducted in US).....	18
3.2.2.6 Study 12541A (elderly study, multiregional study including US and Canada).....	19
3.2.2.7 Study 11985A (maintenance study, non-US, multiregional, including Canada).....	20
3.2.3 <i>Patient Disposition, Demographic and Baseline Characteristics</i>	20
3.2.3.1 Study 11492A (non-US multiregional, including Canada).....	20
3.2.3.2 Study 305 (conducted outside North America).....	21
3.2.3.3 Study 13267A (conducted outside North America).....	22
3.2.3.4 Study 315 (conducted in US).....	23
3.2.3.5 Study 316 (conducted in US).....	24
3.2.3.6 Study 12541A (elderly study, multiregional study including US and Canada).....	24
3.2.3.7 Study 11985A (maintenance study, non-US, multi-regional including Canada).....	25
3.2.4 <i>Efficacy Results and Conclusions</i>	26
3.2.4.1 Study 11492A (non-US, multiregional, including Canada).....	26
3.2.4.2 Study 305 (outside North America).....	27
3.2.4.3 Study 13267A (outside North America).....	28
3.2.4.4 Study 315 (US).....	29
3.2.4.5 Study 316 (US).....	30
3.2.4.6 Study 12541A (elderly study, multiregional trial including US and Canada).....	31
3.2.4.7 Study 11985A (maintenance study, non-US, multi-regional including Canada).....	33
3.3 EVALUATION OF SAFETY	39
4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	44
4.1 GENDER, RACE, AGE, AND GEOGRAPHIC REGION	44
4.1.1 <i>Study 11492A (non-US multiregional, including Canada)</i>	45
4.1.2 <i>Study 305 (outside North America)</i>	45
4.1.3 <i>Study 13267A (outside North America)</i>	46
4.1.4 <i>Study 315 (US)</i>	46
4.1.5 <i>Study 316 (US)</i>	47
4.1.6 <i>Study 12541A (elderly, multiregional including US and Canada)</i>	47
4.1.7 <i>Study 11985A (maintenance, multiregional, non US, multi-regional including Canada)</i>	48
4.2 OTHER SPECIAL/SUBGROUP POPULATIONS	48

5. SUMMARY AND CONCLUSIONS	49
5.1 STATISTICAL ISSUES	49
5.2 COLLECTIVE EVIDENCE	49
5.3 CONCLUSIONS AND RECOMMENDATIONS	54
APPENDIX A. BRIEF SUMMARY OF NEGATIVE AND FAILED STUDIES	55
APPENDIX B. BRIEF SUMMARY OF STUDY 308 (GAD)	58
APPENDIX C. PRIMARY ENDPOINT EFFICACY BY VISIT	59
APPENDIX D. INCIDENCE RATES OF TESD BY GENDER SUBGROUPS	61

LIST OF TABLES

Table 1. List of the efficacy studies and investigated doses	9
Table 2. Study 11492A Subject Disposition.....	20
Table 3. Study 11492A Demographic and Baseline Characteristics (Patients Treated).....	21
Table 4. Study 305 Subject Disposition.....	21
Table 5. Study 305 Demographic and Baseline characteristics (All randomized patients)	22
Table 6. Study 13267A Subject Disposition.....	22
Table 7. Study 13267A Demographic and Baseline Characteristics (Patients Treated).....	22
Table 8. Study 315 Subject Disposition.....	23
Table 9. Study 315 Demographic and Baseline Characteristics (All Randomized Patients).....	23
Table 10. Study 316 Subject Disposition (All Randomized Patients).....	24
Table 11. Study 316 Demographic and Baseline Characteristics (All Randomized Patients).....	24
Table 12. Study 12541A Subject Disposition (All Randomized Patients)	25
Table 13. Study 12541A Demographic and Baseline Characteristics (Patients Treated).....	25
Table 14. Study 11985A Subject Disposition (All Randomized Patients)	25
Table 15. Study 11985A Demographic and Baseline Characteristics (FAS)	26
Table 16. MADRS Total Score Change From Baseline at Week 6 (FAS, LOCF, ANCOVA).....	26
Table 17. CGI-I score at Week 6 (LS Mean Difference (SE [#]) from Placebo).....	27
Table 18. HAM-D24 Total Score Change From Baseline at Week 8 (FAS, MMRM).....	27
Table 19. Sponsor-Proposed Key Secondary Endpoints at Week 8 (LS Mean Difference (SE [#]) from Placebo).....	28
Table 20. MADRS Total Score Change From Baseline at Week 8 (LS Mean Difference (SE [#]) from Placebo).....	28
Table 21. MADRS Total Score Change From Baseline at Week 8 (FAS, MMRM).....	29
Table 22. Sponsor-Proposed Key Secondary Endpoints at Week 8 (LS Mean Difference (SE [#]) from Placebo).....	29
Table 23. MADRS Total Score Change From Baseline at Week 8 (FAS, MMRM).....	30
Table 24. Sponsor-Proposed Key Secondary endpoints at Week 8 (LS Mean Difference (SE [#]) from Placebo).....	30
Table 25. Study 316 MADRS Total Score Change From Baseline at Week 8 (FAS, MMRM).....	31
Table 26. Sponsor-Proposed Key Secondary Endpoints at Week 8 (LS Mean (SE [#]) Difference from Placebo).....	31
Table 27. Study 12541A HAM-D-24 Total Score Change From Baseline at Week 8 (FAS, LOCF, ANCOVA).....	32
Table 28. Sponsor-Proposed Key Secondary Endpoints at Week 8 (LS Mean Difference (SE [#]) from Placebo).....	32
Table 29. Secondary endpoints at Week 8 (LS Mean Difference (SE [#]) to Placebo).....	32
Table 30. Time to Relapse within 24 Weeks of Double-blind Period (FAS).....	33
Table 31. Summary of Percent of Patients in Subgroups based on Stabilization Duration	39
Table 32. Relapse Rates in Subgroups based on Stabilization Duration	39
Table 33. Number of patients with ASEX assessment at Baseline.....	42
Table 34. Incidence of TESD by Study in subjects without sexual dysfunction at baseline	43
Table 35. Incidence of TESD by study at two consecutive visits in subjects without sexual dysfunction at baseline.....	44
Table 36. MADRS Total Score Change From Baseline at Week 6 (FAS, LOCF, ANCOVA).....	45
Table 37. HAM-D24 Total Score Change From Baseline at Week 8 (FAS, MMRM).....	45
Table 38. MADRS Total Score Change From Baseline at Week 8 (FAS, MMRM).....	46
Table 39. MADRS Total Score Change From Baseline at Week 8 (FAS, MMRM).....	46
Table 40. MADRS Total Score Change From Baseline at Week 8 (FAS, MMRM).....	47
Table 41. HAM-D-24 Total Score Change From Baseline at Week 8 (FAS, LOCF, ANCOVA).....	47
Table 42. Time to relapse within 24 Weeks of Double-blind Period (FAS, MMRM)	48
Table 43. Primary efficacy results for positive efficacy studies	50
Table 44. Summary Results of CGI-I Analysis for Positive Acute Efficacy Studies	52
Table 45. Summary Results of SDS Analysis for Positive Acute Efficacy Studies	53
Table 46. MADRS Total Score Change From Baseline at Week 8 (FAS, LOCF, ANCOVA).....	55
Table 47. HAM-D-24 Total Score Change From Baseline at Week 6	56
Table 48. HAM-D24 Total Score Change From Baseline at Week 8 (FAS, LOCF).....	57
Table 49. Study 317 MADRS Total Score Change From Baseline at Week 8 (FAS, MMRM).....	57
Table 50. HAM-A Total Score Change From Baseline at Week 8 (FAS, LOCF, ANCOVA).....	58

Table 51. LS Mean (SE) Change From Baseline by Week in MADRS Total Score LS Mean (FAS, LOCF, ANCOVA).....	59
Table 52. LS Mean (SE) Change from Baseline by Week in HAM-D-24 Total Score (FAS, MMRM)	59
Table 53. LS Mean (SE) Change from Baseline by Week in MADRS Total Score (FAS, MMRM).....	59
Table 54. LS Mean (SE) Change from Baseline by Week in MADRS Total Score (FAS, MMRM).....	59
Table 55. LS Mean (SE) Change from Baseline by Week in MADRS Total Score (FAS, MMRM).....	60
Table 56. LS Mean (SE) Change from Baseline by Week in HAM-D-24 Total Score (FAS, LOCF, ANCOVA)	60
Table 57. Incidence of TESD by Study in Male subjects without sexual dysfunction at baseline	61
Table 58. Incidence of TESD by Study in Female subjects without sexual dysfunction at baseline.....	61
Table 59. Incidence of TESD by study at two consecutive visits in Male subjects without sexual dysfunction at baseline.....	62
Table 60. Incidence of TESD by study at two consecutive visits in Female subjects without sexual dysfunction at baseline.....	62

LIST OF FIGURES

Figure 1. Kaplan-Meier Curves of Time to Relapse in the Double-Blind Treatment Phase (curves from top to bottom: LuAA21004, Placebo).....	34
Figure 2. Kaplan-Meier Estimates of Relapse Probability in the Double-Blind Treatment Phase (curves from top to bottom: Placebo, LuAA21004).....	35
Figure 3. Empirical Cumulative Distribution Function curves of censoring time for all censored patient population (curves from top to bottom: LuAA21004, Placebo).....	36
Figure 4. Empirical Cumulative Distribution Function curves of time to event in the scenario without censoring (curves from top to bottom: Placebo, LuAA21004).....	37
Figure 5. Kaplan-Meier Estimates of Probability of Composite Event (curves from top to bottom: Placebo, LuAA21004).....	38

1. EXECUTIVE SUMMARY

The sponsor submitted an original New Drug Application (NDA 204-447) for vortioxetine (Lu AA21 004) tablets (5, 10, 15 and 20 mg) for the treatment of Major Depressive Disorder.

The efficacy of Lu AA21004 in the treatment of MDD has been evaluated in 10 short-term placebo-controlled studies (9 in adults and 1 in elderly subjects) and 1 long-term placebo-controlled relapse-prevention study. Out of 10 short-term studies, 5 studies were conducted exclusively in the United States.

The primary efficacy assessment tools in the short-term studies were either the MADRS or HAM-D24. The sponsor also pre-specified multiple key-secondary endpoints varying from study to study. The list of key secondary endpoints includes change from baseline in SDS, CGI-I, HAM-D24 response rate, MADRS remission rate, and change from baseline in HAM-D24 in subjects with baseline HAM-A ≥ 20 . As was indicated in the FDA Advice Letter dated September 14, 2010, the only acceptable key secondary endpoints are either CGI-I or SDS.

In the dose range of 5-20 mg, statistically significant treatment effect of Lu AA21004 with respect to placebo was demonstrated in at least one short-term study for each investigated dose (5mg, 10mg, 15mg, and 20mg). The doses lower than 20 mg were not statistically significantly better than placebo in US studies.

Based on the results of the long-term maintenance trial, treatment by Lu AA21004 (5mg or 10mg) was statistically superior to placebo in the relapse prevention. In the study protocol, the required duration of stability prior to randomization was only 2 weeks. Approximately 60% of patients in each treatment arm were stable for at least 4 weeks prior to randomization and less than 40% of the patients were stable for 6 weeks or more. Only a few patients in both treatment arms were stable for 10 weeks. At the End of Phase II Meeting (February 5, 2008) the Division pointed out that the stabilization period is too short. However, given that the protocol was submitted to the Division after this non-US study was initiated, it was too late to make changes.

The observed incidence of treatment emergent sexual dysfunction (TESD) for all studies combined had tendency to increase with the dose in the dose range of 5-20 mg. However, incidence rates varied substantially from study to study within each dose. In female subgroup the observed incidence rates were higher than in males in all treatment arms. The sponsor did not pursue any labeling claims pertaining to the TESD.

2. INTRODUCTION

2.1 Overview

The sponsor submitted an original New Drug Application (NDA 204-447) for vortioxetine (Lu AA21 004) tablets (5, 10, 15 and 20 mg) for the treatment of Major Depressive Disorder.

The efficacy of Lu AA21004 in the treatment of MDD has been evaluated in 10 short-term placebo-controlled studies (9 in adults and 1 in elderly subjects) and 1 long-term placebo-controlled relapse-prevention study. Out of 10 short-term studies, 5 studies were conducted exclusively in the United States, 2 studies were multiregional trials including Canada, 2 studies were multiregional trials conducted outside North America, and 1 trial was multiregional and included both, US and Canada.

2.2 Data Sources

The clinical study reports and data sets are submitted electronically. The network path for the submission is: <\\Cdsub5\evsprod\NDA204447\0000>. Primary analysis data sets and SAS programs are located at <\\Cdsub5\evsprod\NDA204447\0000\m5\datasets>. The sponsor's responses to Division's requests pertaining to treatment emergent sexual dysfunction and stabilization duration are located at <\\Cdsub1\evsprod\NDA204447\0021>, and <\\Cdsub1\evsprod\NDA204447\0022>.

3. STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The reviewer found the quality and integrity of the submitted data satisfying and acceptable for the review analysis.

3.2 Evaluation of Efficacy

The sponsor submitted clinical study reports for 10 short-term efficacy and 1 long-term placebo-controlled relapse-prevention study.

Table 1. List of the efficacy studies and investigated doses

Study Name	Investigated Dose					
	1mg	2.5mg	5mg	10mg	15mg	20mg
Acute Efficacy						
11492A	Region: multiregional, non US, includes Canada. Dates: 08/2006-08/2007					
			X	X		
11984A	Region: multiregional, non US, includes Canada. Dates: 02/2008-04/2009					
	X	X	X			
303	Region: US. Dates: 04/2008-11/2008					
		X				
304	Region: US. Dates: 04/2008-12/2008					
	X	X				
305	Region: multiregional, outside North America. Dates: 08/2008-08/2009					
	X		X	X		
13267A	Region: multiregional, outside North America. Dates: 05/2010-09/2011					
					X	X
315	Region: US. Dates: 06/2010-02/2012					
					X	X
316	Region: US. Dates: 07/2010-01/2012					
				X		X
317	Region: US. Dates: 08/2010-06/2012					
				X	X	
12541A (elderly)	Region: multiregional, includes US and Canada. Dates: 01/2009-02/2010					
			X			
Relapse Prevention						
11985A	Region: multiregional, non US, includes Canada. Dates: 12/2007-09/2009					
			Flexible dose: 5mg or 10mg			

Source: Reviewer's summary

Out of 10 short-term studies, 6 studies (11492A, 305, 13267A, 315, 316, 12541A) had positive efficacy results for at least one of the investigated doses. The reviewer evaluated efficacy results only for positive short-term studies and for the maintenance study.

3.2.1 Study Design and Endpoints

All short-term efficacy studies were multicenter, double-blind, randomized, placebo-controlled, parallel-group fixed-dose studies of 6- and 8-week duration. The primary endpoint was either change from baseline in HAM-D24 or MADRS. The sponsor also proposed multiple key-secondary endpoints varying from study to study. The list of key secondary endpoints includes change from baseline in SDS, CGI-I, HAM-D24 response rate, MADRS remission rate, and change from baseline in HAM-D24 in subjects with baseline HAM-A \geq 20. As was indicated in the Memorandum of Meeting Minutes (Type C meeting held on March 30, 2010) and in the FDA Advice Letter dated September 14, 2010, the only acceptable key secondary endpoints are either CGI-I or SDS. However, since both are used to assess the functional domain, it would be redundant to use both as key secondary endpoints. The sponsor acknowledged the Agency's comments, and agreed that analyses on the above endpoints (except SDS or CGI-I) are exploratory and are not intended for labeling. However, the sponsor decided to retain the above key secondary endpoints in the hierarchy to be able to prospectively assess them.

The relapse-prevention study was a multi-national, multi-centre, randomized, double-blind, placebo-controlled, relapse-prevention study with a 12-week open-label treatment period and a double-blind treatment period of at least 24 weeks. The primary efficacy variable was the time to relapse of MDD within the first 24 weeks of the double-blind period.

3.2.1.1 Study 11492A

The primary objective of the study was to evaluate the efficacy, safety, and tolerability of two fixed doses (5 and 10mg/day) of Lu AA21004 *versus* that of placebo after 6 weeks of treatment in patients with MDD.

The study was conducted outside US, specifically at 49 centers in 11 countries: Austria, Australia, Canada, Czech Republic, Finland, France, Italy, Malaysia, Slovakia, Spain, and Sweden. Patients who fulfilled all inclusion criteria and none of the exclusion criteria were randomized equally (1:1:1:1) to placebo, Lu AA21004 5mg/day, Lu AA21004 10mg/day, or venlafaxine 225 mg/day for a 6-week double-blind treatment period. The doses of Lu AA21004 were 5mg/day or 10mg/day for 6 weeks. The dose of venlafaxine was 75mg/day for 4 days, 150mg/day for the following 3 days, and 225mg/day for the remainder of the treatment period. The primary efficacy variable, MADRS total score, was assessed weekly during the double-blind treatment period.

The primary efficacy endpoint was the mean change from baseline in the MADRS total score at Week 6. The sponsor also included the mean change from baseline at Week 1 as a key secondary endpoint. In addition, the sponsor listed multiple exploratory secondary endpoints including CGI-I and HAMD-24 total score.

Reviewer's Remark: The mean change from baseline in the MADRS Total score at Week 1 is not an acceptable key secondary endpoint. The protocol was submitted to the Division after study was initiated. At the time of submission, the sponsor considered this study as a proof of concept study.

3.2.1.2 Study 305

The primary objective of the study was to evaluate the efficacy of 3 fixed doses of Lu AA21004 (1 mg, 5 mg, and 10 mg QD) versus placebo after 8 weeks of treatment in subjects with MDD.

The study was conducted outside North America, specifically in 14 countries: Australia, Croatia, France, Germany, Latvia, Lithuania, Malaysia, Netherlands, Poland, Republic of Korea, Russia, South Africa, Taiwan, and Ukraine. Out of 62 sites participating in the study, 48 sites randomized subjects. At Baseline, subjects were randomized equally (1:1:1:1) to 1 of the 4 treatment arms for an 8-week, double-blind treatment period. Subjects were seen weekly during the first 2 weeks of treatment and then every 2 weeks up to the end of the 8-week treatment period.

The primary endpoint for this study was the mean change from Baseline in the 24-item Hamilton Depression Scale (HAM-D24) total score after 8 weeks of treatment.

The sponsor proposed the following key secondary endpoints:

- Mean change from Baseline in SDS total score at week 8.
- Mean CGI-I at week 8.
- HAM-D24 response rate at week 8, with response defined as a $\geq 50\%$ decrease in HAM-D24 total score from Baseline.
- Mean change from baseline in HAM-D24 total score at week 8 in subjects with baseline HAM-A ≥ 20 .
- MADRS remission rate at week 8, with remission defined as a MADRS total score ≤ 10 .

3.2.1.3 Study 13267A

The primary objective of study 13267A was to evaluate the efficacy of two fixed doses of Lu AA21004 (15 and 20 mg/day) versus placebo as assessed by the change from baseline in MADRS after 8 weeks of treatment in subjects with MDD.

This study was conducted outside North America, specifically at 72 sites in 13 countries: Belgium, Estonia, Finland, France, Germany, Latvia, Lithuania, Norway, Russian Federation, Slovakia, South Africa, Sweden, and Ukraine. The study patients were randomized equally (1:1:1:1) to placebo, Lu AA21004 15mg/day, Lu AA21004 20mg/day, or duloxetine 60mg/day for 8 weeks of double-blind treatment. Duloxetine treatment arm was used as an active reference. Patients randomized to treatment with Lu AA21004 received a dose of 10mg/day during Week 1 (up titration); from Weeks 2 to 8 they received a dose of either 15 or 20mg/day. Patients were seen weekly during the first two weeks of treatment and then every two weeks up to the end of the 8 week treatment period.

The primary endpoint specified by the sponsor was mean change from baseline in MADRS total score at Week 8.

The sponsor defined the following key secondary endpoints in prioritized order:

- MADRS response at Week 8 (response defined as a $\geq 50\%$ decrease in the MADRS total score from baseline)
- Mean CGI-I score at Week 8
- Mean change from baseline in MADRS total score at Week 8 in patients with baseline HAM-A total score ≥ 20
- Mean change from baseline in HAM-A total score at Week 8
- Remission at Week 8 (remission defined as a MADRS total score ≤ 10)
- Mean change from baseline in SDS total score at Week 8.

3.2.1.4 Study 315

The primary objective of study 315 was to evaluate the efficacy of two fixed doses of LuAA21004, 15 and 20 mg QD, compared with placebo as assessed by the MADRS after 8 weeks of treatment in subjects with MDD.

The study enrolled patients at 58 sites in the United States. At the Baseline Visit subjects were randomized equally in a 1:1:1:1 ratio to 1 of the 4 treatment arms for an 8-week, double-blind treatment period to receive once daily oral doses of placebo, Lu AA21004 15 mg, Lu AA21004 20 mg, or duloxetine 60 mg which was used as a reference drug. Subject randomization was stratified by subject's baseline sexual function status (normal or abnormal decided by Arizona Sexual Experiences Scale [ASEX] score). Subjects assigned to either Lu AA21004 15 or 20 mg arms received Lu AA21004 10 mg QD (up-titration) for the first week of the Double-Blind Treatment Period and 15 or 20 mg QD for the remaining 7 weeks of treatment. Subjects visited the site weekly during the first 2 weeks of treatment and then every 2 weeks up to the end of the 8-week Treatment Period.

Mean change from Baseline in MADRS total score at week 8 was the primary endpoint. The following secondary endpoints were included by the sponsor in the multiple testing procedure:

- MADRS response at Week 8, with response defined as a $\geq 50\%$ decrease in the MADRS total score from Baseline.
- Mean CGI-I score at Week 8.
- Change from Baseline in MADRS total score at Week 8 in subjects with baseline HAM-A total score ≥ 20 .
- MADRS remission at Week 8, with remission defined as a MADRS total score ≤ 10 .
- Change from Baseline in SDS total score at Week 8.

3.2.1.5 Study 316

The primary objective of study 316 was to evaluate the efficacy of Lu AA21004 10 and 20 mg QD compared with placebo as assessed by the MADRS after 8 weeks of treatment in subjects with MDD.

The study enrolled patients at 37 sites in the United States. At Baseline (Day 0), subjects were randomized in a 1:1:1 ratio to 1 of 3 treatment arms to receive once daily oral doses of Lu AA21004 10 mg, Lu AA21004 20 mg, or placebo in the 8-week, double-blind treatment period. Subject randomization was stratified by subject's baseline sexual function status (normal or abnormal decided by Arizona Sexual Experiences Scale [ASEX] score). All subjects took their first dose of study medication in the morning of Day 1, the day after randomization. Subjects assigned to Lu AA21004 20 mg received Lu AA21004 10 mg for the first week of the Double-Blind Treatment Period and 20 mg for the remaining 7 weeks of treatment. Subjects visited the site weekly during the first 2 weeks of treatment and then every 2 weeks up to the end of the 8-week Treatment Period

Mean change from Baseline in MADRS total score at week 8 was the primary endpoint. The following secondary endpoints were included in the multiple testing procedure by the sponsor:

- MADRS response at Week 8, with response defined as a $\geq 50\%$ decrease in the MADRS total score from Baseline.
- Mean CGI-I score at Week 8.
- Change from Baseline in MADRS total score at Week 8 in subjects with baseline HAM-A total score ≥ 20 .
- MADRS remission at Week 8, with remission defined as a MADRS total score ≤ 10 .
- Change from Baseline in SDS total score at Week 8.

3.2.1.6 Study 12541A (elderly study)

The primary objective was to assess the efficacy of Lu AA21004 (5mg/day) *versus* placebo in the acute treatment of depression by means of the change from baseline in the 24-item Hamilton Depression Scale (HAM-D24D) total score after 8 weeks of double-blind treatment in *elderly* patients (65 years or older).

The study was conducted at 81 centers in 7 countries: 4 in Canada, 10 in Finland, 4 in France, 5 in Germany, 5 in Sweden, 9 in Ukraine, and 44 in the United States. Patients meeting the diagnostic criteria for recurrent MDD with at least one previous Major Depressive Episode (MDE) before the age of 60 years were seen for a Screening Visit. At baseline, eligible patients who continued to fulfill all the inclusion criteria and none of the exclusion criteria were equally randomized (1:1:1) to one of three treatment groups (Lu AA21004 5mg/day, duloxetine 60mg/day, or placebo) for the 8-week, double-blind treatment period. Patients were seen weekly during the first 2 weeks of treatment and then every 2 weeks up to the end of the 8-week treatment period.

The primary efficacy endpoint was mean change from baseline in HAM-D24 total score at Week 8. To evaluate the efficacy of Lu AA21004 during 8 weeks of double-blind treatment mean changes in HAM-D24 score at Weeks 6, 4, 2 and 1 were included by sponsor as key secondary endpoints.

Reviewer's Remark: The mean change from baseline in HAM-D24 score at Weeks 6, 4, 2 and 1 are not acceptable key secondary endpoints.

3.2.1.7 Study 11985A (maintenance study)

The primary objective was to evaluate the efficacy of Lu AA21004 (5 and 10 mg/day) in the prevention of relapse of Major Depressive Episodes (MDE). This was a non-US, randomized, double-blind, placebo-controlled, relapse-prevention study in patients with MDD. The study was conducted at 66 centers in 17 countries: 4 in Canada, 2 in Australia, 2 in Austria, 3 in Belgium, 6 in Finland, 6 in France, 6 in Germany, 5 in India, 3 in the Republic of Korea, 3 in Norway, 7 in Poland, 6 in South Africa, 5 in Sweden, 2 in Taiwan, 2 in Thailand, 3 in Turkey, and 1 in the United Kingdom.

The study consisted of two consecutive periods: a 12-week open-label treatment period with Lu AA21004 and a double-blind, fixed-dose, placebo-controlled treatment period of at least 24 weeks. From Week 8 of the Open-label Period, the dose was fixed. Patients in remission (MADRS total score ≤ 10 at both Weeks 10 and 12) were randomized to the double-blind, placebo-controlled, fixed-dose treatment period (Double-blind Period). Patients were randomized equally (1:1) to Lu AA21004 or placebo; patients randomized to Lu AA21004 continued on the dose that was fixed from Week 8 in the Open-label Period; patients randomized to placebo were switched abruptly to placebo. In the Double-blind Period, efficacy and safety data were collected at Weeks 1, 2, and 4 and then at 4-week intervals.

The primary efficacy variable was the time to relapse of MDD *within the first 24 weeks* of the Double-blind Period. Relapse was defined as a MADRS total score ≥ 22 or an unsatisfactory treatment effect (lack of efficacy) as judged by the investigator. The time to relapse was defined as: (date of relapse - date of Visit 8 [Baseline II / Randomization]) + 1 day.

Reviewer's Remark: The protocol was submitted to the Division after this non-US study was initiated. At the End of Phase II Meeting (February 5, 2008) the Division pointed out that the stabilization period of 12 weeks is too short. Patients need to be stable for at least 12 weeks before randomization.

3.2.2 Statistical Methodologies

In the short-term efficacy studies in adults and the elderly, the primary efficacy analysis methods proposed by the sponsor were either the mixed model for repeated measurements (MMRM) or the analysis of covariance (ANCOVA) model based on the last observation carried forward (LOCF).

All efficacy analyses in the short-term studies were performed on the full analysis set (FAS), which was defined as all randomized subjects who took at least 1 dose of study drug and who had at least 1 valid post-baseline measurement of the primary efficacy variable.

3.2.2.1 *Study 11492A (non-US multiregional, including Canada)*

Primary Efficacy

The primary efficacy statistical model was an analysis of covariance (ANCOVA) of the change from baseline in MADRS total score (FAS, last observation carried forward [LOCF]) with treatment and center as fixed factors and the baseline MADRS score as a covariate.

Four hypotheses were proposed by the sponsor as being part of the primary efficacy analysis, which was adjusted for multiplicity using a closed testing procedure that ranked hypotheses hierarchically and continued testing at the 5% level of significance as long as the previous hypothesis was rejected. The order of testing was:

- H1: No difference between 10mg Lu AA21004 and placebo at Week 6
- H2: No difference between 5mg Lu AA21004 and placebo at Week 6
- H3: No difference between 10mg Lu AA21004 and placebo at Week 1
- H4: No difference between 5mg Lu AA21004 and placebo at Week 1

Supportive Analysis

The primary efficacy analysis was repeated on observed cases (OC) data, using both an ANCOVA and a mixed model for repeated measurements (MMRM).

Efficacy Analysis of Secondary Endpoints (exploratory)

The change from baseline to each visit in all the secondary efficacy variables, except response and remission, was analyzed using an ANCOVA, adjusting for baseline score, center, and treatment. For CGI-I, the baseline CGI-S score was used for adjustment.

Reviewer's Remark: The mean change from baseline in the MADRS Total score at Week 1 is not an acceptable key secondary endpoint. The protocol was submitted to the Division in April 2007, after study was initiated (August 2006). At the time of submission, the sponsor considered this study as a proof of concept study.

3.2.2.2 *Study 305 (conducted outside North America)*

Primary Efficacy

The primary endpoint (Change from Baseline in HAM-D24 total score after 8 weeks of treatment) was analyzed using a mixed model for repeated measurements (MMRM). The model included the fixed categorical effects of treatment, site, visit, and treatment-by-visit interaction, as well as the continuous, fixed covariates of baseline HAM-D24 total score and baseline HAM-D24 total score-by-visit interaction. An unstructured covariance structure was used to model the within-patient errors. This analysis was performed using OC data.

Analyses of the sponsor-proposed key secondary efficacy variables

Changes from Baseline in the SDS and CGI-I scores were analyzed per visit using MMRM similar to the one described above for the primary variable. HAM-D24 response (defined as a $\geq 50\%$ decrease from Baseline in HAM-D24 total score) and MADRS remission (defined as a MADRS total score of 10 or less) were analyzed by logistic regression, adjusting for Baseline score and treatment using both LOCF and OC.

Reviewer's Remark: Since the only acceptable key secondary endpoints are either CGI-I or SDS, the statistical analyses methods for response and remission were not reviewed.

Multiplicity Adjustment

To control for multiplicity, a pre-specified sequential testing procedure was applied to compare 10 mg and 5 mg Lu AA21004 to placebo. The efficacy endpoints were tested in the following sequential order at significance level 0.05; as soon as an endpoint was non-significant at 0.05, the testing procedure stopped for all subsequent endpoints:

- Change from baseline in HAMD-24 total score at Week 8 (Lu AA21004 10mg vs Placebo, MMRM).
- Change from baseline in SDS total score at Week 8 (Lu AA21004 10mg vs Placebo, MMRM).
- CGI-I at Week 8 (Lu AA21004 10mg vs Placebo, MMRM).
- HAMD-24 response rate at Week 8 (Lu AA21004 10mg vs Placebo, LOCF).
- Change from baseline in HAMD-24 total score at Week 8 in subjects with baseline HAMA-A ≥ 20 (Lu AA21004 10mg vs Placebo, MMRM).
- MADRS remission rate at Week 8 (Lu AA21004 10mg vs Placebo, LOCF).
- Change from baseline in HAMD-24 total score at Week 8 (Lu AA21004 5mg vs Placebo, MMRM).
- Change from baseline in SDS total score at Week 8 (Lu AA21004 5mg vs Placebo, MMRM).
- CGI-I at Week 8 (Lu AA21004 5mg vs Placebo, MMRM).
- HAMD-24 response rate at Week 8 (Lu AA21004 5mg vs Placebo, LOCF).
- Change from baseline in HAMD-24 total score at Week 8 in subjects with baseline HAMA-A ≥ 20 (Lu AA21004 5mg vs Placebo, MMRM).
- MADRS remission rate at Week 8 (Lu AA21004 5mg vs Placebo, LOCF).

3.2.2.3 *Study 13267A (conducted outside North America)*

Primary Efficacy

Primary efficacy analysis was a mixed model for repeated measurements (MMRM) of the change from baseline in MADRS total score at Week 8. These analyses were performed on the full-analysis set (FAS), using observed cases (OC). The model included the fixed categorical effects of treatment, site, visit, and treatment-by-visit interaction, as well as the continuous, fixed covariates of baseline MADRS total score and baseline MADRS total score-by-visit interaction. An unstructured covariance structure was used to model the within-patient errors.

Analysis of the sponsor-proposed key secondary variables

The analyses of the continuous endpoints (MADRS and SDS total scores and CGI-I score) were performed using the same methodology as for the primary efficacy analysis (FAS, MMRM). For analyses of the CGI-I, the CGI-S score served as baseline. Response and remission were analyzed using logistic regression with treatment as factor and the baseline score as a covariate (FAS, last observation carried forward [LOCF]).

Reviewer's Remark: Since the only acceptable key secondary endpoints are either CGI-I or SDS, the statistical analyses methods for response and remission were not reviewed.

As sensitivity analyses, the primary and key secondary continuous endpoints were tested using an analysis of covariance (ANCOVA), adjusting for baseline score, site, and treatment, (FAS, OC and LOCF). For the analyses of the CGI-I, the CGI-S score served as baseline.

To adjust for multiplicity, the 15 and 20mg doses of Lu AA21004 were tested separately *versus* placebo in the primary and key secondary efficacy analyses at a Bonferroni-corrected significance level of $0.05/2 = 0.025$. The following sequence of hierarchically ordered primary and key secondary endpoints was used:

1. change from baseline at Week 8 in MADRS total score (primary)
2. response (defined as a $\geq 50\%$ decrease from baseline in MADRS total score) at Week 8
3. CGI-I score at Week 8
4. change from baseline at Week 8 in MADRS total score in patients with a baseline HAM-A total score ≥ 20
5. remission (defined as a MADRS total score ≤ 10) at Week 8
6. change from baseline at Week 8 in SDS total score

As soon as a hypothesis was rejected (that is, there was no statistically significant difference *versus* placebo at the 0.025 level of significance within a dose [15 or 20mg]), the testing procedure was stopped for all subsequent endpoints for that dose.

3.2.2.4 *Study 315 (conducted in US)*

Primary Analysis

Primary analysis was based on a mixed model for repeated measurements (MMRM) with treatment, center, week, treatment-by-week interaction, baseline MADRS total score-by-week as fixed effects.

Analysis of key secondary variables proposed by sponsor

CGI-S and CGI-I were analyzed as continuous variables by study visit using MMRM similar to the analysis of the primary variable, where the CGI-S Baseline was used as the covariate adjustment in the MMRM.

The responder and remission rates, including MADRS response and MADRS remission, were analyzed at all time points by logistic regression adjusting for baseline score and treatment by both LOCF and OC methods.

Reviewer's Remark: Since the only acceptable key secondary endpoints are either CGI-I or SDS, the statistical analyses methods for response and remission were not reviewed.

Multiple testing

To control the two-sided type I error over all the efficacy endpoints that are intended to support potential claims among the 2 Lu AA21004 doses, 15 mg and 20 mg, the efficacy endpoints were tested for each dose in the following sequential order at significance level 0.025; as soon as an endpoint was non-significant at 0.025, the testing procedure stopped for all subsequent endpoints for that dose:

- Change from Baseline in MADRS total score at Week 8 (MMRM).
- MADRS responders at Week 8 (LOCF).
- CGI-I at Week 8 (MMRM).
- Change from Baseline in MADRS total score at Week 8 in subjects with Baseline HAM-A ≥ 20 (MMRM).
- MADRS remissions at Week 8 (LOCF).
- Change from Baseline in SDS total score at Week 8 (MMRM).

3.2.2.5 *Study 316 (conducted in US)*

Primary Analysis

Change from Baseline in MADRS total score was the primary variable. Primary analysis was based on a mixed model for repeated measurements (MMRM) analysis of covariance with treatment, center, week, treatment-by-week interaction, baseline MADRS total score-by-week as fixed effects. An unstructured covariance matrix was assumed.

Analyses of key secondary variables proposed by sponsor

CGI-S and CGI-I were analyzed as continuous variables by study visit using MMRM similar to the analysis of the primary variable, where the CGI-S Baseline was used as the covariate adjustment in the MMRM. Change from Baseline in SDS total score was analyzed as continuous variables by study visit using MMRM where the relevant Baseline was used as the covariate adjustment.

The responder and remission rates, including MADRS response and MADRS remission, were analyzed at all time points by logistic regression adjusting for baseline score and treatment by both LOCF and OC methods.

Reviewer's Remark: Since the only acceptable key secondary endpoints are either CGI-I or SDS, the statistical analyses methods for response and remission were not reviewed.

Multiple testing

To control the two-sided type I error over all the efficacy endpoints that are intended to support potential claims among the 2 Lu AA21004 doses, 10 mg and 20 mg, the efficacy endpoints were tested for each dose in the following sequential order at significance level 0.025; as soon as an endpoint was non-significant at 0.025, the testing procedure stopped for all subsequent endpoints for that dose:

- Change from Baseline in MADRS total score at Week 8 (MMRM).
- MADRS responders at Week 8 (LOCF).
- CGI-I at Week 8 (MMRM).
- Change from Baseline in MADRS total score at Week 8 in subjects with Baseline HAM-A ≥ 20 (MMRM).
- MADRS remissions at Week 8 (LOCF)
- Change from Baseline in SDS total score at Week 8 (MMRM).

3.2.2.6 Study 12541A (elderly study, multiregional study including US and Canada)

Primary Efficacy Analysis

The primary efficacy analysis was an analysis of covariance (ANCOVA) of the change from baseline in HAM-D24 total score at Week 8 (FAS, last observation carried forward [LOCF]), with treatment and centre as factors and the baseline HAM-D24 total score as a covariate.

Analyses of sponsor-proposed key secondary efficacy endpoints

The changes from baseline in HAM-D24 total score at Weeks 6, 4, 2, and 1 were investigated using the same model as the one that was used in the primary efficacy analysis.

Multiple testing

To control the two-sided type I error over all the efficacy endpoints that are intended to support potential claims the efficacy endpoints were tested in the following sequential order at significance level 0.05. As soon as an endpoint was non-significant at 0.05, the testing procedure stopped for all subsequent endpoints:

- H1: No difference between Lu AA21004 and placebo at Week 8
- H2: No difference between Lu AA21004 and placebo at Week 6
- H3: No difference between Lu AA21004 and placebo at Week 4
- H4: No difference between Lu AA21004 and placebo at Week 2
- H5: No difference between Lu AA21004 and placebo at Week 1.

Reviewer's Remark: The mean change from baseline in HAM-D24 score at Weeks 6, 4, 2 and 1 are not acceptable key secondary endpoints.

3.2.2.7 Study 11985A (maintenance study, non-US, multiregional, including Canada)

The full analysis set (FAS) included all patients who completed the open-label treatment period, were randomized to the double-blind treatment period and who took at least one dose of the 1 dose of study drug in the double-blind period.

Primary Efficacy Analysis

The treatment groups were compared using a Cox model with an exact method to handle ties (based on the FAS). This analysis was supplemented by Kaplan-Meier plots. The log-rank test was performed as a supportive analysis.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

3.2.3.1 Study 11492A (non-US multiregional, including Canada)

Table 2 summarizes the subject disposition and analysis sets. Of the 429 subjects randomized, 426 subjects received treatment, and 360 (83.9%) completed the treatment. Full analysis set included 425 patients.

Table 2. Study 11492A Subject Disposition

Subjects	Placebo	Venlafaxine 225mg/day	Lu AA21004	
			5mg	10mg
Randomized	105	114	109	101
Treated	105	113	108	100
Completed	87 (82.9%)	93 (82.3%)	98 (90.7%)	82 (82.0%)
Early Termination	18 (17.1%)	20 (17.7%)	10 (9.3%)	18 (18.0%)
Full Analysis Set	105	112	108	100

Source: Clinical Study Report Panel 8 (pg. 53)

The demographic and baseline characteristics are summarized in Table 3. Approximately 55% of the patients in the venlafaxine group and 65% of the patients in the placebo and Lu AA21004 groups were women. The mean age of the patients was 43 years, ranging from 18 to 65 years, and the majority (92%) were Caucasian.

Table 3. Study 11492A Demographic and Baseline Characteristics (Patients Treated)

Characteristic	Placebo	Venlafaxine 225mg/day	Lu AA21004	
			5mg	10mg
Number of Patients (treated)	105	113	108	100
Sex				
Male	36 (34.3%)	51 (45.1%)	38 (35.2%)	34 (34.0)
Female	69 (65.7%)	62 (54.9%)	70 (64.8%)	66 (66.0)
Race				
Caucasian	98 (93.3%)	104 (92.0%)	101 (93.5%)	89 (89.0%)
Black	1 (1.0%)	0 (0.0%)	0 (0.0%)	1 (1.0%)
Asian	6 (5.7%)	8 (7.1%)	7 (6.5%)	8 (8.0%)
Other	0 (0.0%)	1 (0.9%)	0 (0.0%)	2 (2.0%)
Age Mean (SD)	42 (10.9)	45 (10.3)	43.8 (11.6)	42.3 (13.1)
Weight (kg) Mean (SD)	75.2 (15.4)	75.4 (16.4)	72.9 (18.7)	71.7 (15.6)
BMI (kg/m²) Mean (SD)	26.0 (5.1)	25.8 (4.7)	25.2 (5.0)	24.8 (4.8)

Source: Clinical Study Report Panel 14 (pg. 59), Table 24 (pg. 199)

3.2.3.2 Study 305 (conducted outside North America)

Table 4 summarizes the subject disposition and analysis sets. Of the 664 subjects screened, 560 were randomized. The percentage of subjects completing treatment in the 4 treatment groups was similar among treatments.

Table 4. Study 305 Subject Disposition

Subjects	Placebo	1mg	Lu AA21004	
			5mg	10mg
Randomized	140 (100%)	140 (100%)	140 (100%)	140 (100%)
Completed	127 (90.7 %)	127 (90.7%)	129 (92.1%)	122 (87.1%)
Early Termination	13 (9.3%)	13 (9.3%)	11 (7.9%)	18 (12.9%)
Full Analysis Set	139 (99.3%)	139 (99.3%)	139 (99.3%)	139 (99.3%)

Source: Clinical Study Report Table 10.a. (pg. 52)

The demographic and baseline characteristics are summarized in Table 5. Approximately 2/3 of subjects in each treatment group were women, and the mean age ranged from 45.4 years to 47.3 years. The majority of subjects in each group were Caucasian/White, although the 1 mg group had a greater percentage of White subjects (92.1%) than the other treatment groups.

Table 5. Study 305 Demographic and Baseline characteristics (All randomized patients)

Characteristic	Lu AA21004			
	Placebo	1mg	5mg	10mg
Sex				
Male	54 (38.6%)	47 (33.6%)	53 (37.9%)	55 (39.3%)
Female	86 (61.4%)	93 (66.4%)	87 (62.1%)	85 (60.7%)
Race				
Caucasian	120 (85.7%)	129 (92.1%)	120 (85.7%)	114 (81.4%)
Black	5 (3.6%)	1 (0.7%)	2 (1.4%)	2 (1.4%)
Asian	14 (10.0%)	8 (5.7%)	17 (12.1%)	23 (16.4%)
Other	1 (0.7%)	2 (1.4%)	1 (0.7%)	1 (0.7%)
Age (years): Mean (SD)	46.4 (12.26)	45.4 (11.89)	47.3 (11.95)	46.4 (12.27)
Weight (kg): Mean (SD)	75.2 (14.92)	75.6 (17.28)	75.4 (17.02)	74.6 (15.19)
BMI (kg/m²): Mean (SD)	26.4 (4.6)	26.5 (5.4)	26.4 (5.1)	26.2 (4.6)

Source: Clinical Study Report Table 10.b. (pg. 54-55)

3.2.3.3 Study 13267A (conducted outside North America)

Table 6 summarizes the subject disposition and analysis sets. Of the 608 subjects randomized, 607 subjects received treatment, and 506 (83.4%) completed the treatment. Full analysis set included 604 patients.

Table 6. Study 13267A Subject Disposition

Subjects	Placebo	Duloxetine	Lu AA21004	
			15mg	20mg
Randomized	158	147	152	151
Treated	158 (100%)	147 (100%)	151 (100%)	151 (100%)
Completed	133 (84.2%)	131 (89.1%)	117 (77.5%)	125 (82.8%)
Early Termination	25 (15.8%)	16 (10.9%)	34 (22.5%)	26 (17.2%)
Full Analysis Set	158	146	149	151

Source: Clinical Study Report Panel 11 (pg. 65)

The demographic and baseline characteristics are summarized in Table 7.

Table 7. Study 13267A Demographic and Baseline Characteristics (Patients Treated)

Characteristic	Placebo	Duloxetine	Lu AA21004	
			15mg	20mg
Number of Patients (treated)	158	147	151	151
Sex				
Male	48 (30.4%)	45 (30.6%)	54 (35.8%)	60 (39.7%)
Female	110 (69.6%)	102 (69.4%)	97 (64.2%)	91 (60.3%)
Race				
Caucasian	156 (98.7%)	144 (98.0%)	150 (99.3%)	146 (96.7%)
Black	2 (1.3%)	3 (2.0%)	0 (0.0%)	2 (1.3%)
Asian	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.7%)
Other	0 (0.0%)	0 (0.0%)	1 (0.7%)	2 (1.3%)
Age (years): Mean (SD)	48.1 (13.1)	45.6 (13.6)	47.0 (14.6)	46.2 (13.4)
Weight (kg): Mean (SD)	77.1 (17.5)	74.8 (17.5)	73.1 (14.7)	74.4 (14.9)
BMI (kg/m²): Mean (SD)	27.0 (6.0)	26.3 (5.6)	25.7 (4.7)	25.8 (4.2)

Source: Clinical Study Report Panel 17 (pg. 69)

Approximately two-thirds of the patients in the total patient population were women. There was an imbalance between the treatment groups with respect to the proportions of women, ranging from 60% (Lu AA21004 20mg) to 70% (placebo). The mean age of the patients was 47 years, ranging from 46 years (Lu AA21004 20mg and duloxetine) to 48 years (placebo), and the majority (>96%) were white.

3.2.3.4 Study 315 (conducted in US)

A total of 614 subjects were randomized at 58 sites. The number of subjects randomized by treatment group is presented in Table 8. A total of 144 subjects (23.5%) prematurely discontinued the study. Four subjects (2 placebo and 2 duloxetine 60 mg) were randomized but not treated; they discontinued prior to receiving study drug.

Table 8. Study 315 Subject Disposition

Subjects	Placebo	Duloxetine	Lu AA21004	
			15mg	20mg
Randomized	161 (100%)	152 (100%)	147 (100%)	154 (100%)
Treated	159 (98.8%)	150 (98.7%)	147 (100%)	154 (100%)
Full Analysis Set	153 (95.0%)	146 (96.1%)	145 (98.6%)	147 (95.5%)
Completed	129 (80.1%)	115 (75.7%)	113 (76.9%)	113 (73.4%)
Early Termination	32 (19.9%)	37 (24.3%)	34 (23.1%)	41 (26.6%)

Source: Clinical Study Report Table 10.b (pg. 70)

Demographic data are summarized in Table 9. Mean age (SD) was 42.9 (12.35) years, 26.2% of the subjects were male, and 73.8% were female. The majority of the subjects were Caucasian (76.5%), 22.1% were Black, and 1.1% were Asian.

Table 9. Study 315 Demographic and Baseline Characteristics (All Randomized Patients)

Characteristic	Placebo	Duloxetine	Lu AA21004	
			15mg	20mg
Number of Patients	161	152	147	154
Sex				
Male	45 (28.0%)	33 (21.7%)	43 (29.3%)	40 (26.0%)
Female	116 (72.0%)	119 (78.3%)	104 (70.7%)	114 (74.0%)
Race				
Caucasian	122 (75.8%)	119 (78.3%)	114 (77.6%)	115 (74.7%)
Black	37 (23.0%)	32 (21.1%)	31 (21.1%)	36 (23.4%)
Asian	1 (0.6%)	1 (0.7%)	2 (1.4%)	3 (1.9%)
Other	1 (0.6%)	0	0	0
Age (years): Mean (SD)	42.4 (12.6)	43.4 (12.2)	43.1 (12.3)	42.8 (12.4)
Weight (kg): Mean (SD)	88.6 (24.8)	87.2 (24.2)	87.4 (21.4)	87.4 (23.8)
BMI (kg/m²): Mean (SD)	31.1 (7.9)	31.5 (8.4)	31.3 (7.5)	30.9 (7.6)

Source: Clinical Study Report Table 11.b (pg. 74)

3.2.3.5 Study 316 (conducted in US)

A total of 462 subjects were randomized at 37 sites. The number of subjects randomized by treatment group is presented in Table 10. Seventy-seven subjects (16.7%) prematurely discontinued the study.

Table 10. Study 316 Subject Disposition (All Randomized Patients)

Subjects	Placebo	Lu AA21004	
		10mg	20mg
Randomized	157 (100%)	155 (100%)	150 (100%)
Treated (Safety)	157 (100%)	155 (100%)	150 (100%)
Full Analysis Set	155 (98.7%)	154 (99.4%)	148 (98.7%)
Completed	139 (88.5%)	124 (80.0%)	122 (81.3%)
Early Termination	18 (11.5%)	31 (20.0%)	28 (18.7%)

Source: Clinical Study Report Table 10.b (pg. 66)

Demographic data are summarized in Table 11. Mean age (SD) was 42.8 (12.23) years, 27.5% of the subjects were male and 72.5% were female. The majority of subjects were Caucasian (69.9%), 27.9% were Black, and 0.6% were Asian.

Table 11. Study 316 Demographic and Baseline Characteristics (All Randomized Patients)

Characteristic	Placebo	Lu AA21004	
		10mg	20mg
Number of Patients	N=157	N=155	N=150
Sex			
Male	47 (29.9%)	37 (23.9%)	43 (28.7%)
Female	110 (70.1%)	118 (76.1%)	107 (71.3%)
Race			
Caucasian	120 (76.4%)	106 (68.4%)	97 (64.7%)
Black	37 (23.6%)	43 (27.7%)	49 (32.7%)
Other	0 (0.0%)	6 (3.9%)	4 (2.7%)
Age (years): Mean (SD)	42.3 (11.6)	43.1 (12.0)	43.1 (13.1)
Weight (kg): Mean (SD)	88.3 (23.3)	89.0 (23.1)	87.1 (23.6)
BMI (kg/m²): Mean (SD)	31.3 (7.3)	31.9 (7.8)	30.8 (7.8)

Source: Clinical Study Report Table 11.b (pg. 71)

3.2.3.6 Study 12541A (elderly study, multiregional study including US and Canada)

The subject disposition is presented in Table 12. A total of 453 subjects were randomized, 61 withdrew and 392 patients completed the trial. Between 84% and 88% of the patients in each treatment group completed the study.

Table 12. Study 12541A Subject Disposition (All Randomized Patients)

Subjects	Placebo	Lu AA21004 5mg	Duloxetine 60mg
Randomized	145	157	151
Treated	145	156	151
Completed	128 (88.3%)	136 (87.2%)	128 (84.8%)
Withdrawn	17 (11.7%)	20 (12.8%)	23 (15.2%)
AEs	6	10	15
Lack of Efficacy	7	2	0
Other	4	8	8
Full Analysis Set	145	155	148

Source: Clinical Study Report Panel 11 (pg.60), Panel 12 (pg. 60)

Demographic data are summarized in Table 13. Mean age (SD) was 70.6 (4.9) years, 34.3% of the subjects were male and 65.7% were female. The overwhelming majority of subjects were Caucasian (94.7%), 4.0% were Black, and 0.7% were Asian.

Table 13. Study 12541A Demographic and Baseline Characteristics (Patients Treated)

Characteristic	Placebo	Lu AA21004 5mg	Duloxetine 60mg
Number of Patients	N=145	N=156	N=151
Sex			
Male	55 (37.9%)	49 (31.4%)	51 (33.8%)
Female	90 (62.1%)	107 (68.6%)	100 (66.2%)
Race			
Caucasian	139 (95.9)	145 (92.9%)	144 (95.4%)
Black	4 (2.8%)	8 (5.1%)	6 (4.0%)
Asian	1 (0.7%)	1 (0.6%)	1 (0.7%)
Other	1 (0.7%)	2 (1.3%)	0 (0.0%)
Age (years): Mean (SD)	70.3 (4.4)	70.5 (4.8)	70.9 (5.5)
Weight (kg): Mean (SD)	76.9 (16.2)	77.9 (15.9)	76.7 (16.4)

Source: Clinical Study Report Panel 18 (pg. 71)

3.2.3.7 Study 11985A (maintenance study, non-US, multi-regional including Canada)

The subject disposition is presented in Table 14. A total of 400 subjects were randomized, 171 withdrew, and 80 patients withdrew due to a relapse.

Table 14. Study 11985A Subject Disposition (All Randomized Patients)

Subjects	Placebo	Lu AA21004
Randomized	194	206
Treated	192	204
Completed	104 (53.6%)	125 (60.7%)
Withdrawn	90 (46.4%)	81 (39.3%)
AEs	5	16
Relapse	52	28
Other	33	37
Full Analysis Set	192	204

Source: Clinical Study Report Panel 14 (pg.63), Panel 15 (pg. 64)

Demographic data by treatment group are summarized in Table 15. Mean age (SD) was 44.9 (12.2) years, 36.9% of the subjects were male and 63.1% were female. The overwhelming majority of subjects were Caucasian (80.6%), the second largest racial subgroup were Asian (16.7%).

Table 15. Study 11985A Demographic and Baseline Characteristics (FAS)

Characteristic	Placebo	Lu AA21004
Number of Patients	N=192	N=204
Sex		
Male	72 (37.5%)	74 (36.3%)
Female	120 (62.5%)	130 (63.7%)
Race		
Caucasian	154 (80.2%)	165 (80.9%)
Black	2 (1.0%)	4 (2.0%)
Asian	34 (17.7%)	32 (15.7%)
Other	2 (1.0%)	3 (1.5%)
Age (years): Mean (SD)	45.1 (12.1)	44.8 (12.4)

Source: Clinical Study Report Panel 22 (pg. 77)

3.2.4 Efficacy Results and Conclusions

3.2.4.1 Study 11492A (non-US, multiregional, including Canada)

Primary Endpoint

The results of the primary efficacy analysis are presented in Table 16. In the primary efficacy analysis by ANCOVA (LOCF), Lu AA21004 10 mg and 5mg arms were statistically significantly better than placebo (both p-values < 0.0001) in reducing the MADRS total score at Week 6, with LS mean differences from placebo of -5.7 and -5.9 points respectively. At Week 1, both arms were not significantly different from placebo.

Table 16. MADRS Total Score Change From Baseline at Week 6 (FAS, LOCF, ANCOVA)

MADRS Total Score	Placebo (N=105)	Venlafaxine 225mg (N=112)	Lu AA21004	
			5mg (N=108)	10mg (N=100)
Week 6				
Mean Baseline Score (SD [*])	33.9 (2.7)	34.2 (3.1)	34.1 (2.6)	34.0 (2.8)
LS Mean Change from Baseline (SE [#])	-14.5 (1.03)	-20.9 (0.99)	-20.4 (1.01)	-20.2 (1.04)
Difference from Placebo (SE)		-6.4 (1.38)	-5.9 (1.39)	-5.7 (1.42)
p-value		<0.0001	<0.0001	<0.0001
Week 1				
LS Mean Change from Baseline (SE)	-5.0 (0.50)	-4.5 (0.48)	-5.3 (0.49)	-5.9 (0.51)
Difference from Placebo (SE)		0.5 (0.67)	-0.2 (0.67)	-0.8 (0.69)
p-value		0.414	0.749	0.238

Source: Clinical Study Report Table 42 (pg. 241), Panel 17 (pg.62)

Results confirmed by the reviewer

* Standard deviation, #Standard error

Secondary Endpoint CGI-I

The sponsor did not pre-specify any key secondary endpoints. All analysis results for the secondary endpoints were exploratory and were not adjusted for multiplicity. In the ANCOVA (LOCF) analysis of CGI-I score at Week 6, both Lu AA21004 arms appeared to be better than placebo (both p-values < 0.001). The analysis results are presented in Table 17.

Table 17. CGI-I score at Week 6 (LS Mean Difference (SE[#]) from Placebo.

Efficacy variable	Venlafaxine		Lu AA21004 5mg		Lu AA21004 10mg	
	Difference	p-value	Difference	p-value	Difference	p-value
CGI-I score	-0.7 (0.16)	<0.0001	-0.6 (0.16)	<0.001	-0.6 (0.16)	<0.001

Source: Clinical Study Report Table 349 (pg. 349)

[#]Standard error

Reviewer's Conclusion for Study 11492A

Lu AA21004 10mg and 5mg treatment arms were statistically significantly better than placebo in mean change from baseline in MADRS total score at Week 6.

3.2.4.2 Study 305 (outside North America)

Primary Endpoint

For the primary endpoint (change from Baseline in HAM-D24 total score after 8 weeks of treatment), based on the pre-specified sequential testing procedure, only the Lu AA21004 10 mg group was statistically significantly superior to placebo (p<0.001, MMRM). Even though the nominal p-values for Lu AA21004 1 mg and 5 mg dose groups were <0.001, these arms can not be considered statistically significantly different from placebo because the formal testing was stopped at an earlier step per the pre-specified testing sequence. The results of the primary efficacy analysis are presented in Table 18.

Table 18. HAM-D24 Total Score Change From Baseline at Week 8 (FAS, MMRM)

HAM-D-24 Total score	Lu AA21004			
	Placebo (N=139)	1mg (N= 139)	5mg (N=139)	10mg (N=139)
Mean Baseline Score (SD [*])	32.7 (4.4)	32.5 (5.2)	32.2 (5.0)	33.1 (4.8)
LS Mean Change from Baseline (SE [#])	-11.3 (0.74)	-14.8 (0.75)	-15.4 (0.74)	-16.2 (0.76)
Difference from Placebo (SE)	-	-3.5 (1.04)	-4.1 (1.04)	-4.9 (1.05)
p-value	-	<0.001	<0.001	<0.001

Source: Clinical Study Report Table 11.a (pg. 63) and Table 15.2.1.2.1.

Results confirmed by the reviewer

* Standard deviation, [#]Standard error

Sponsor-Proposed Key Secondary Endpoints

The 10mg dose was not statistically significantly different from placebo in the second variable in the testing hierarchy (the change from Baseline in SDS total score at Week 8). Since the formal testing was stopped at the second variable in the pre-specified order, none of the subsequent endpoints in the pre-specified testing hierarchy are considered statistically significantly different from placebo. The analysis results for all sponsor-proposed key secondary endpoints are presented in Table 19.

Table 19. Sponsor-Proposed Key Secondary Endpoints at Week 8 (LS Mean Difference (SE[#]) from Placebo).

Efficacy variable	Lu AA21004 1mg		Lu AA21004 5mg		Lu AA21004 10mg	
	Difference	p-value	Difference	p-value	Difference	p-value
Δ SDS total score	-0.1 (1.01)	0.963	-1.1 (0.99)	0.263	-1.5 (1.03)	0.135
CGI-I score	-0.5 (0.13)	<0.001	-0.5 (0.13)	<0.001	-0.6 (0.13)	<0.001
HAM-D24 response %	24.5	<0.001	22.3	<0.001	26.6	<0.001
Δ HAM-D24 total score (baseline HAM-A ≥20)*	-4.4 (1.47)	0.003	-4.3 (1.42)	0.003	-6.5 (1.49)	<0.001
MADRS remission %	9.4	0.062	12.3	0.015	10.1	0.026

Source: Clinical Study Report Tables 15.2.9.2.5, 15.2.5.2.5, 15.2.1.1.7, 15.2.1.9.3, and 15.2.2.2.13.

[#]Standard error

*Subgroup of patients with baseline HAM-A total score ≥20

Secondary Endpoint MADRS Total score

The sponsor did not pre-specify MADRS Total score as a key secondary variable. The analysis results were exploratory and were not adjusted for multiplicity. In the MMRM (FAS, OC) analysis of change from baseline in MADRS total score, all three Lu AA21004 arms were better than placebo at Week 8 (all p-values < 0.001). The analysis results are presented in Table 20.

Table 20. MADRS Total Score Change From Baseline at Week 8 (LS Mean Difference (SE[#]) from Placebo).

Efficacy variable	Lu AA21004 1mg		Lu AA21004 5mg		Lu AA21004 10mg	
	Difference	p-value	Difference	p-value	Difference	p-value
Δ MADRS total score	-4.0 (1.00)	<0.001	-4.2 (1.00)	<0.001	-4.8 (1.01)	<0.001

Source: Clinical Study Report Table 15.2.2.1.5

[#]Standard error

Reviewer's Conclusion for Study 305

Lu AA21004 10mg treatment arm was statistically significantly better than placebo in the mean change from baseline in HAM-D24 at Week 8. Based on the pre-specified sequential multiple testing procedure, other doses (1 mg and 5 mg) could not be shown effective, although the nominal p-values were very small.

3.2.4.3 Study 13267A (outside North America)

Primary Endpoint

The results of the primary efficacy analysis are presented in Table 21. Both doses of Lu AA21004 were statistically significantly superior to placebo in mean change from baseline in MADRS total score at Week 8 (FAS, MMRM), with a mean treatment difference to placebo of -5.5 (Lu AA21004 15mg) and -7.1 points (Lu AA21004 20mg).

Table 21. MADRS Total Score Change From Baseline at Week 8 (FAS, MMRM)

MADRS Total Score	Placebo (N=158)	Duloxetine (N=146)	Lu AA21004	
			15mg (N=149)	20mg (N=151)
Mean Baseline Score (SD [*])	31.5 (3.6)	31.2 (3.5)	31.8 (3.4)	31.2 (3.4)
LS Mean Change from Baseline (SE [#])	-11.7 (0.76)	-21.2 (0.77)	-17.2 (0.79)	-18.8 (0.78)
LS Mean Difference from Placebo (SE)	-	-9.5 (1.07)	-5.5 (1.09)	-7.1 (1.08)
p-value	-	<0.0001	<0.0001	<0.0001

Source: Clinical Study Report Table 30 (pg. 264)

Results confirmed by the reviewer

* Standard deviation, #Standard error

Sponsor-Proposed Key Secondary Endpoints

Both doses of Lu AA21004 were statistically significantly superior to placebo in CGI-I score and change from baseline in SDS score. Since the only acceptable key secondary endpoints are either CGI-I or SDS, the statistical analysis results on the other secondary endpoints proposed by the sponsor are considered exploratory.

Table 22. Sponsor-Proposed Key Secondary Endpoints at Week 8 (LS Mean Difference (SE[#]) from Placebo).

Efficacy variable	Duloxetine		Lu AA21004 15mg		Lu AA21004 20mg	
	Difference	p-value	Difference	p-value	Difference	p-value
MADRS response %	41.7 %	<0.001	24.7%	<0.001	29.3%	<0.001
CGI-I score	-1.1 (0.13)	<0.001	-0.7 (0.13)	<0.001	-1.0 (0.13)	<0.001
Δ MADRS total score (baseline HAM-A ≥20) [*]	-8.7	<0.001	-5.2	0.001	-6.4	<0.001
MADRS remission %	35.1 %	<0.001	15.9%	0.002	19.4%	<0.001
Δ SDS total score	-6.9 (1.13)	<0.001	-3.2 (1.16)	0.005	-3.9 (1.11)	0.001

Source: Clinical Study Report Panel 21 (pg. 77)

Results for CGI-I and SDS are confirmed by the reviewer

#Standard error

*Subgroup of patients with baseline HAM-A total score ≥20

Reviewer's Conclusion for Study 13267A

At Week 8, Lu AA21004 15mg and 20mg were statistically significantly better than placebo in mean change from baseline in MADRS total score, in the mean CGI-I score, and in change from baseline in SDS score.

3.2.4.4 Study 315 (US)

Primary Endpoint

The results of the primary efficacy analysis are presented in Table 23. In the primary efficacy analysis by MMRM, Lu AA21004 20 mg was statistically significantly better (p=0.023) than placebo in reducing the MADRS total score at Week 8, with an LS mean difference from placebo of -2.8 points. Lu AA21004 15 mg was not statistically significantly different from placebo at Week 8.

Table 23. MADRS Total Score Change From Baseline at Week 8 (FAS, MMRM)

MADRS Total Score	Placebo (N=153)	Duloxetine 60mg (N= 146)	Lu AA21004	
			15mg (N=145)	20mg (N=147)
Mean Baseline Score (SD [*])	31.5 (4.2)	32.8 (4.3)	31.9 (4.1)	32.0 (4.4)
LS Mean Change from Baseline (SE [#])	-12.8 (0.83)	-16.9 (0.88)	-14.3 (0.89)	-15.6 (0.88)
Difference from Placebo (SE)		-4.1 (1.21)	-1.5 (1.21)	-2.8 (1.21)
p-value		<0.001	0.224	0.023

Source: Clinical Study Report Table 11.f (pg. 81)

Results confirmed by the reviewer

* Standard deviation, #Standard error

Sponsor-Proposed Key Secondary Endpoints

Since only Lu A21004 20mg dose arm was statistically significantly superior to placebo in the primary endpoint, the sequential testing was applied to the secondary endpoints within this dose only. Overall, neither of the two doses of Lu AA21004 separated from placebo in CGI-I and change from baseline in SDS. The statistical analysis results on the other secondary endpoints proposed by the sponsor are considered exploratory as they were not accepted as the key secondary endpoints. The analysis results are presented in Table 24.

Table 24. Sponsor-Proposed Key Secondary endpoints at Week 8 (LS Mean Difference (SE[#]) from Placebo).

Efficacy variable	Duloxetine		Lu AA21004 15mg		Lu AA21004 20mg	
	Difference	p-value	Difference	p-value	Difference	p-value
MADRS response %	15.6%	0.004	4.9%	0.348	5.0%	0.332
CGI-I score	-0.3 (0.14)	0.014	-0.1 (0.14)	0.400	-0.2 (0.14)	0.177
Δ MADRS total score (baseline HAM-A ≥20) [*]	-4.1 (2.28)	0.078	0.9 (2.29)	0.684	-0.6 (2.42)	0.797
MADRS remission %	-0.8%	0.728	0.1%	0.845	2.5%	0.503
Δ SDS total score	-1.99 (1.12)	0.078	-0.05 (1.11)	0.962	-0.88 (1.10)	0.427

Source: Clinical Study Report Table 15.2.1.13.1, Table 15.2.4.1.5, Table 15.2.1.9.5, 15.2.1.14.1, and 15.2.6.1.5

Results for CGI-I and SDS are confirmed by the reviewer

#Standard error

*Subgroup of patients with baseline HAM-A total score ≥20

Reviewer's Conclusion for Study 315

Lu AA21004 20 mg treatment arm was statistically significantly better than placebo in mean change from baseline in MADRS total score at Week 8.

3.2.4.5 Study 316 (US)

Primary Endpoint

The results of the primary efficacy analysis are presented in Table 25. In the primary efficacy analysis by MMRM, Lu AA21004 20 mg was statistically significantly better (p=0.002) than placebo in reducing the MADRS total score at Week 8, with an LS mean difference from placebo of -3.6 points. Lu AA21004 10 mg was not statistically significantly different from placebo at Week 8.

Table 25. Study 316 MADRS Total Score Change From Baseline at Week 8 (FAS, MMRM)

Characteristic	Placebo (N=155)	Lu AA21004	
		10mg (N=154)	20mg (N=148)
Mean Baseline Score (SD [*])	32.0 (4.0)	32.3 (4.5)	32.4 (4.3)
LS Mean Change from Baseline (SE [#])	-10.8 (0.81)	-13.0 (0.83)	-14.4 (0.85)
Difference from Placebo (SE)		-2.2 (1.15)	-3.6 (1.16)
p-value		0.058	0.002

Source: Clinical Study Report Table 11.f (pg. 81)

Results confirmed by the reviewer

* Standard deviation, #Standard error

Sponsor-Proposed Key Secondary Endpoints

Based on the pre-specified multiple testing procedure, there were no statistically significant findings for the secondary endpoints although the nominal p-values for CGI-I and change from baseline in SDS were statistically significant at 0.025 level. The analysis results are presented in Table 26.

Table 26. Sponsor-Proposed Key Secondary Endpoints at Week 8 (LS Mean (SE[#]) Difference from Placebo).

Efficacy variable	Lu AA21004 10mg		Lu AA21004 20mg	
	Difference	p-value	Difference	p-value
MADRS response %	5.4%	0.301	10.8%	0.044
CGI-I score	-0.2 (0.13)	0.119	-0.3 (0.13)	0.024
Δ MADRS total score (baseline HAM-A ≥20) [*]	-4.3 (1.89)	0.025	-7.3 (1.85)	<0.001
MADRS remission %	7.2 %	0.093	8.1%	0.059
Δ SDS total score	-1.39 (1.04)	0.183	-2.4 (1.07)	0.025

Source: Clinical Study Report Table 15.2.1.13.1, Table 15.2.4.1.5, Table 15.2.1.9.5, 15.2.1.14.1, and 15.2.6.1.5

Results for CGI-I and SDS are confirmed by the reviewer

#Standard error

*Subgroup of patients with baseline HAM-A total score ≥20

Reviewer's Conclusion for Study 316

Lu AA21004 20mg treatment arm was statistically significantly better than placebo in mean change from baseline in MADRS total score at Week 8.

3.2.4.6 Study 12541A (elderly study, multiregional trial including US and Canada)

Primary Endpoint

The results of the primary efficacy analysis are presented in Table 27. In the primary efficacy analysis by ANCOVA (LOCF), Lu AA21004 5 mg was statistically significantly better (p=0.0011) than placebo in reducing the HAM-D-24 total score at Week 8, with an LS mean difference from placebo of -3.3 points.

Table 27. Study 12541A HAM-D-24 Total Score Change From Baseline at Week 8 (FAS, LOCF, ANCOVA)

HAM-D-24 Total Score	Placebo N=145	Lu AA21004 5mg (N=155)	Duloxetine 60mg (N=148)
Mean Baseline Score (SD [*])	29.4 (5.06)	29.2 (5.02)	28.5 (4.90)
LS Mean Change from Baseline (SE [#])	-10.3 (0.76)	-13.7 (0.74)	-15.8 (0.75)
Difference from Placebo (SE)		-3.3 (1.01)	-5.5 (1.03)
p-value		0.0011	<0.0001

Source: Clinical Study Report Table 32 (pg. 229) and Table 34.(pg. 231)

Results confirmed by the reviewer

* Standard deviation, #Standard error

Sponsor-Proposed Key Secondary Efficacy

Lu AA21004 treatment arm was statistically significantly superior to placebo arm (p = 0.024) at Week 6, but not at Week 4; the testing strategy was therefore stopped at Week 4. The analysis results are presented in Table 28

Table 28. Sponsor-Proposed Key Secondary Endpoints at Week 8 (LS Mean Difference (SE[#]) from Placebo).

HAM-D24 Total Score	Lu AA21004 5 mg		Duloxetine 60mg	
	Difference	p-value	Difference	p-value
Week 6	-2.1 (0.94)	0.024	-4.2 (0.96)	<0.0001
Week 4	-1.1 (0.85)	0.213	-3.3 (0.86)	<0.001
Week 2	-0.3 (0.70)	0.688	-1.25 (0.72)	0.083
Week 1	-0.42 (0.55)	0.448	0.14 (0.56)	0.797

Source: Clinical Study Report Table 34 (pg. 231)

Results confirmed by the reviewer

#Standard error

Secondary Endpoints (exploratory)

Exploratory analysis of CGI-I score and change from baseline in MADRS total score are given in Table 29. The endpoints were analyzed using an ANCOVA (LOCF) model, adjusting for baseline score, center, and treatment. For CGI-I, an ANCOVA of the absolute scores at each visit was performed, using the baseline CGI-S score for adjustment.

Table 29. Secondary endpoints at Week 8 (LS Mean Difference (SE[#]) to Placebo).

Efficacy Variable	Lu AA21004 5 mg		Duloxetine 60mg	
	Difference	p-value	Difference	p-value
Δ MADRS total score	-4.3 (1.03)	<0.0001	-6.8 (1.05)	<0.0001
CGI-I	-0.6 (0.13)	<0.0001	-0.8 (0.13)	<0.001

Source: Clinical Study Report Panel 23 (pg. 80)

#Standard error

Reviewer's Conclusion for Study 12541A

Lu AA21004 5mg treatment arm was statistically significantly better than placebo in mean change from baseline in HAM-D-24 total score at Week 8.

3.2.4.7 Study 11985A (maintenance study, non-US, multi-regional including Canada)

The time to relapse within the first 24 weeks of the Double-blind Period was compared between the treatment groups using the Cox proportional hazard model. The Lu AA21004 (5mg or 10 mg) arm was superior to placebo arm in the maintenance effect (hazard ratio of placebo to drug: 2.01, p-value of 0.0035). The proportion of patients who relapsed was lower in the AA21004 group (13%) than in the placebo group (26%). The results of the primary efficacy analysis are presented in Table 30. The log-rank test also applied by the sponsor gave a p-value of 0.003. This reviewer confirmed the sponsor's primary efficacy results.

Table 30. Time to Relapse within 24 Weeks of Double-blind Period (FAS).

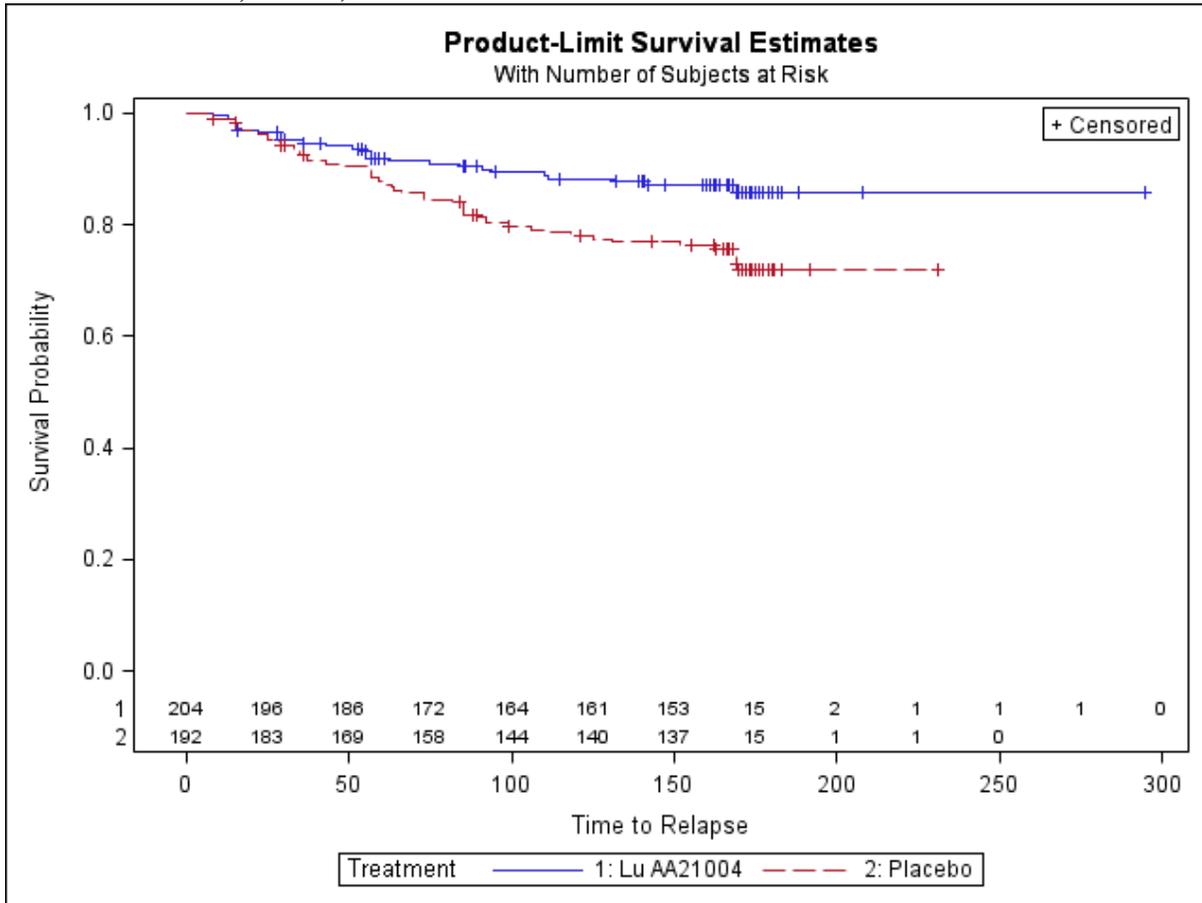
Treatment Arm	Number of Patients	Number of Relapses	Relapse Rate	Cox Proportional Hazard		Log-rank
				Hazard Ratio (plb./drug)	p-value	p-value
Placebo	192	50	26 %	2.01	0.0035	0.003
LuAA21004	204	27	13.2%			

Source: Clinical Study Report Panel 28 (pg. 86)

Results confirmed by the reviewer

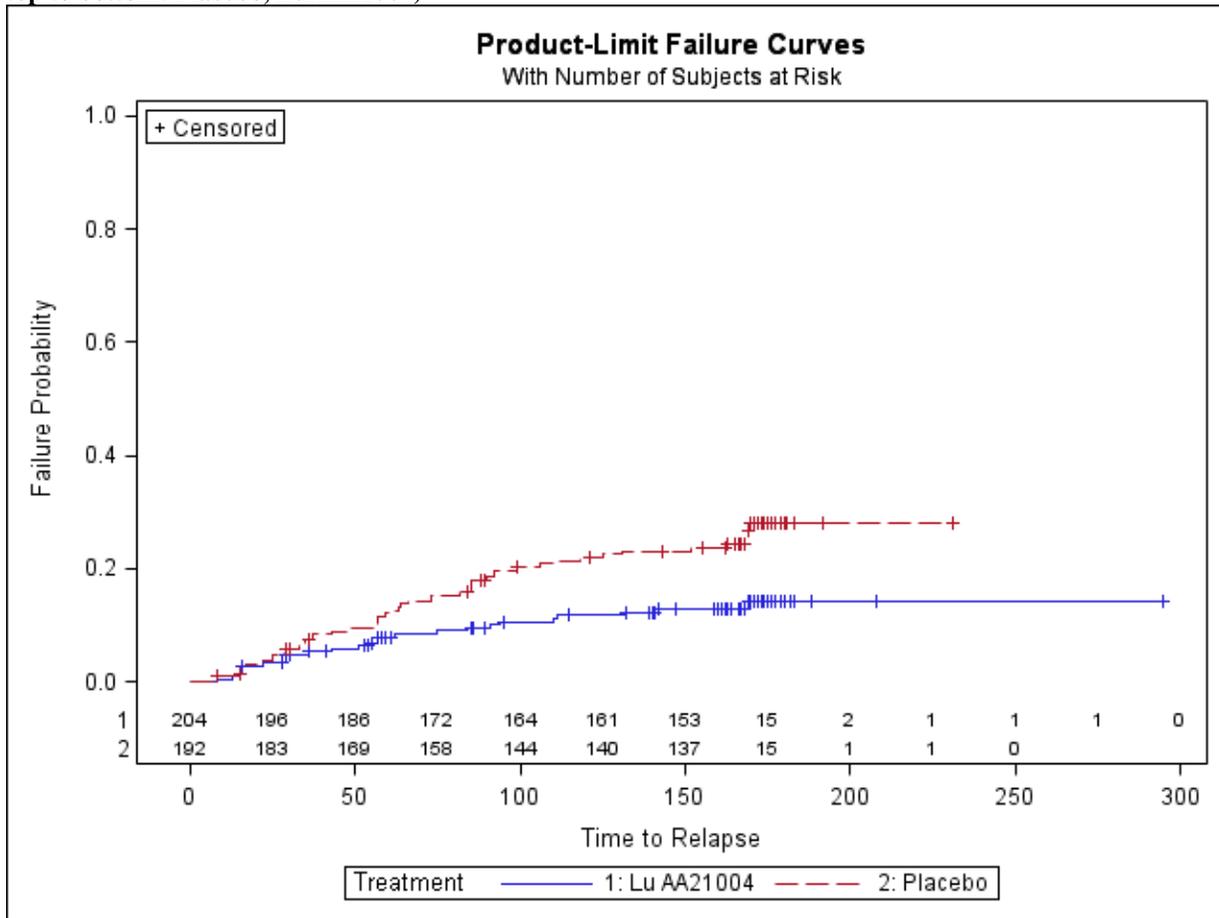
The Kaplan-Meier curves for time to relapse support that the observed relapse rate was lower in the Lu AA21004 treatment group than in placebo treatment group during the entire double-blind period (see Figure 1).

Figure 1. Kaplan-Meier Curves of Time to Relapse in the Double-Blind Treatment Phase (curves from top to bottom: LuAA21004, Placebo)



[Source: Reviewer's results]

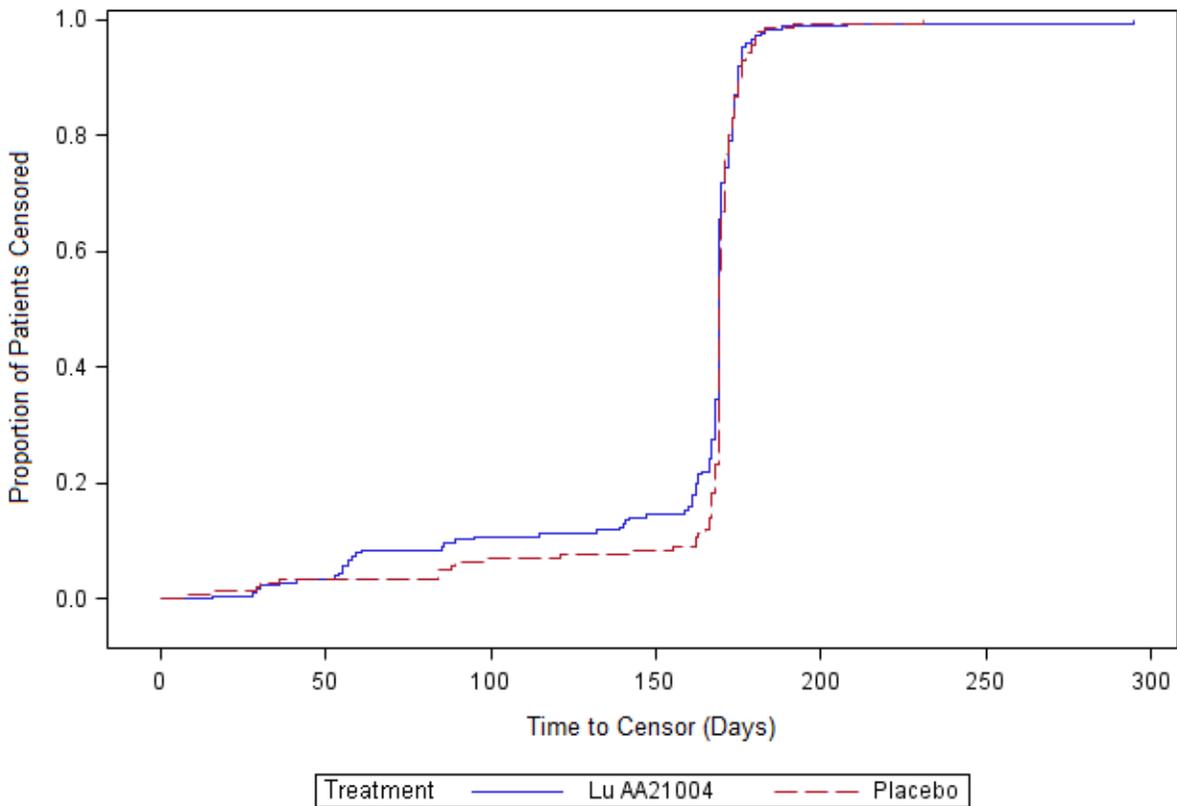
Figure 2. Kaplan-Meier Estimates of Relapse Probability in the Double-Blind Treatment Phase (curves from top to bottom: Placebo, LuAA21004)



[Source: Reviewer's results]

This reviewer explored empirical cumulative distribution functions (CDF) of time to censor for subpopulation of patients who had no intervention for mood episode. In Figure 3, for all censored patients the CDF curves indicate the proportion of patients in each treatment arm who were censored by a given day. For example, by Day 75, approximately 5% of patients were censored in the placebo group and approximately 10% were censored in Lu AA21004 arm. The plot suggests that time to censor was numerically longer in the placebo treatment arm compared with Lu AA21004 arm.

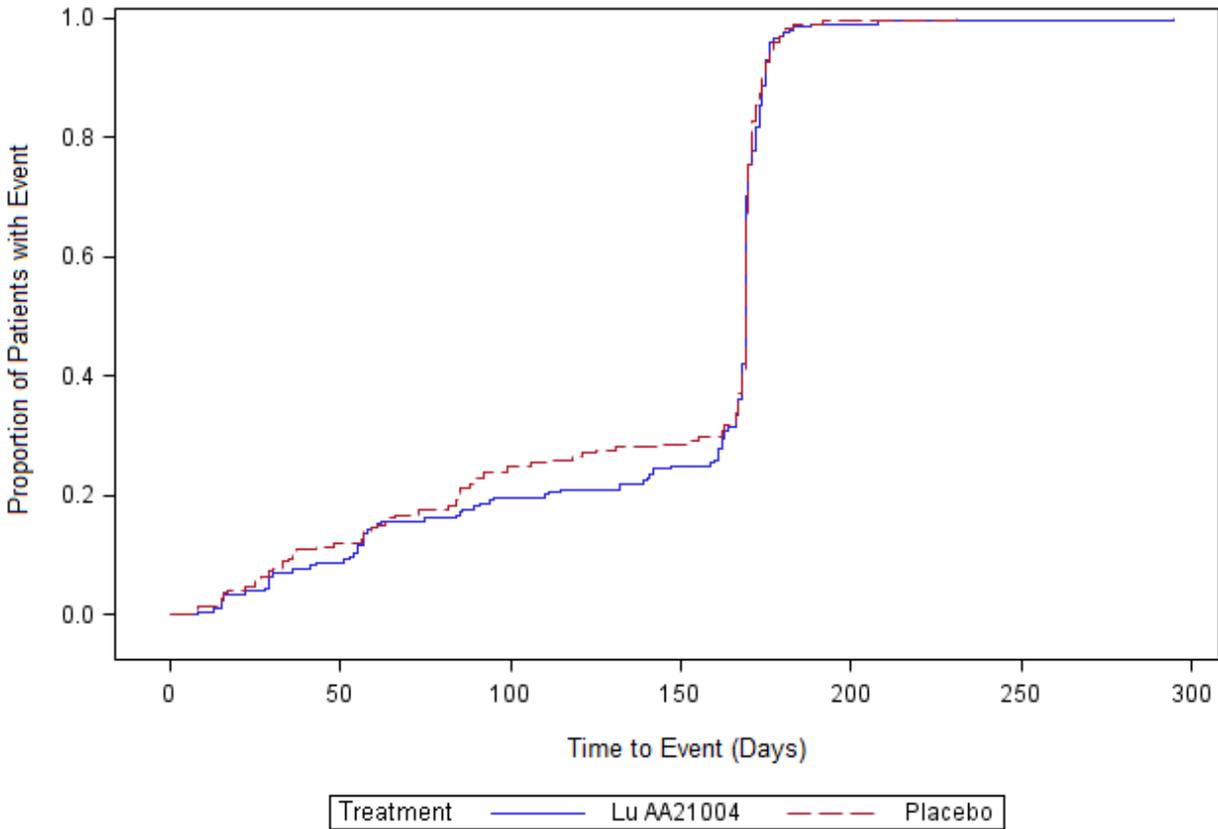
Figure 3. Empirical Cumulative Distribution Function curves of censoring time for all censored patient population (curves from top to bottom: LuAA21004, Placebo)



[Source: Reviewer’s results]

To explore potential impact of censoring distribution on the primary efficacy results, this reviewer considered a hypothetical “worst case” scenario. In this scenario, it was assumed that all patients who were censored actually had an event (relapse) at the time of censoring. The empirical distribution function of time to event for this hypothetical model suggested that numerically time to event was longer in the Lu AA21004 arm which supports results of primary efficacy analysis (see Figure 4).

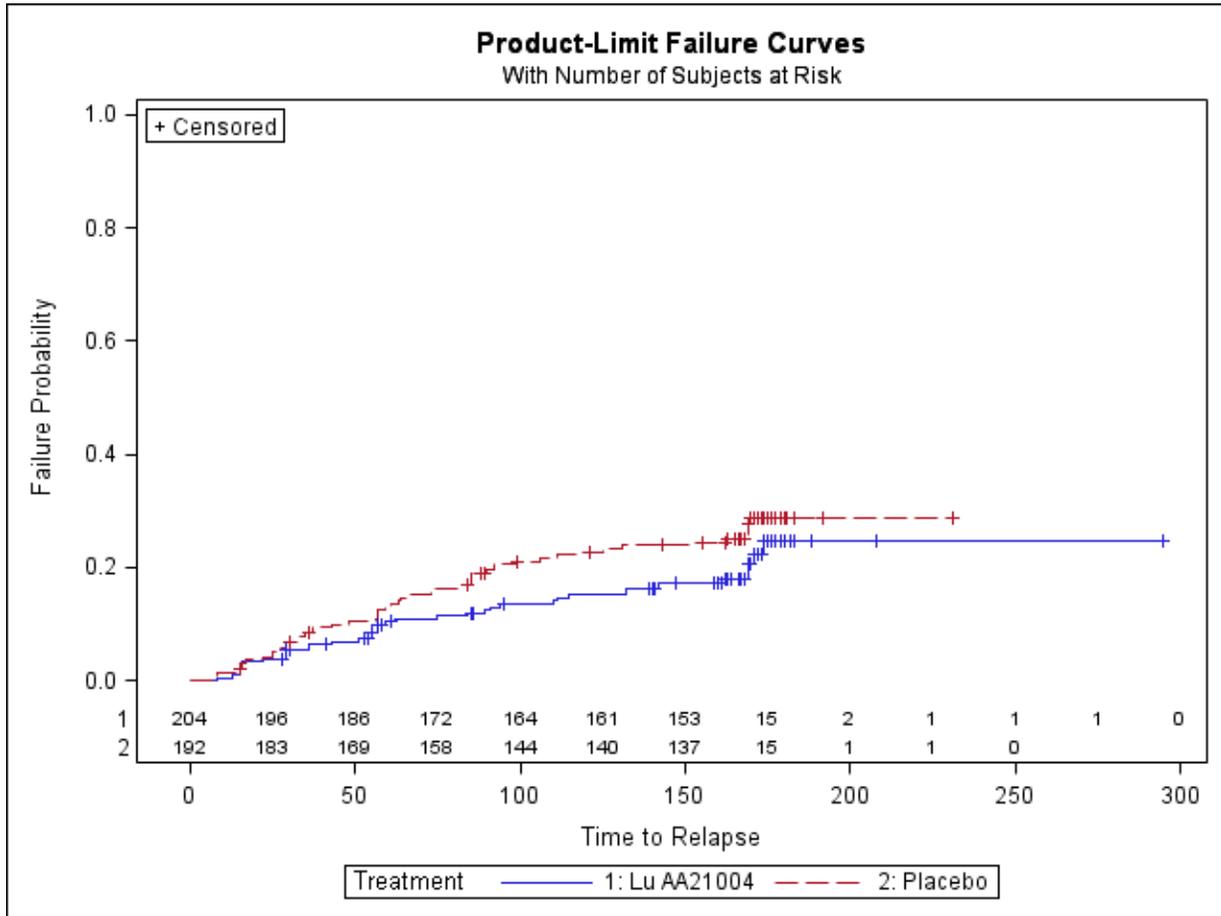
Figure 4. Empirical Cumulative Distribution Function curves of time to event in the scenario without censoring (curves from top to bottom: Placebo, LuAA21004)



[Source: Reviewer's results]

Another exploratory time to event analysis was considered for a composite event defined as either relapse or adverse event. In this exploratory analysis, numerical results were also in favor of Lu AA21004. Kaplan-Meier based curves for probability of composite event are depicted in Figure 5.

Figure 5. Kaplan-Meier Estimates of Probability of Composite Event (curves from top to bottom: Placebo, LuAA21004)



[Source: Reviewer’s results]

Duration of Stabilization

At the End of Phase II Meeting (February 5, 2008) the Division pointed out that the stabilization period of 12 weeks is too short. Patients need to be stable for at least 12 weeks before randomization. The Division requested the sponsor to explore the actual stabilization durations for each patient.

Table 31 includes numbers and percentages of patients in subgroups based on stabilization duration using criteria of sustained MADRS total score ≤ 10 . Both treatment arms had similar percentages of patients in the subgroups based on the length of stabilization prior to randomization. Approximately 60% of patients in each treatment arm were stable for at least 4 weeks prior to randomization and less than 40% of the patients were stable for 6 weeks or more. Only a few patients in both treatment arms were stable for 10 weeks.

Table 31. Summary of Percent of Patients in Subgroups based on Stabilization Duration

Stabilization Duration (Weeks)		≥ 2	≥ 4	≥ 6	≥ 8	≥ 10
Placebo	N=192	191 (99.5%)	115 (59.9%)	69 (35.9%)	34 (17.7%)	5 (2.6%)
LuAA21004	N=204	202 (99.0%)	126 (61.8%)	78 (38.2%)	26 (12.7%)	8 (3.9%)

Source: Sponsor's Response to FDA Request for Information, Table 1.1 ([\Cdsesub1\evsprod\NDA204447\0022](#))

Table 32 provides relapse rates and the hazard ratio in the remission duration subgroups. The results of the primary analyses and the relapse rates were consistent in the subgroups.

Table 32. Relapse Rates in Subgroups based on Stabilization Duration

Stabilization Duration (Weeks)		≥ 2	≥ 4	≥ 6	≥ 8	≥ 10
Relapse Rate	Primary Analysis					
Placebo	50/192 (26%)	50/191 (26%)	28/115 (24%)	18/69 (26%)	13/34(38%)	1/5 (20%)
LuAA21004	27/204 (13%)	26/202 (13%)	14/126 (11%)	9/78 (12%)	2/26 (8%)	0/8 (0%)
Hazard Ratio (plb./drug)	2.01	2.07	2.29	2.49	6.08	NA
p-value	0.0035	0.0026	0.0114	0.0256	0.0176	NA

Source: Sponsor's Response to FDA Request for Information, Table 1.1 ([\Cdsesub1\evsprod\NDA204447\0022](#))

Reviewer's conclusion for Study 11985A

Treatment by Lu AA21004 (5mg or 10mg) demonstrated statistically significant effect versus placebo in the relapse prevention based on Cox proportional hazard model (p=0.0035).

3.3 Evaluation of Safety

Analysis of Arizona Sexual Experience Scale (ASEX)

The ASEX score was assessed in 7 short-term studies: MDD Studies 11984A (negative, failed), 13267A (positive), 304 (negative), 315(positive), 316(positive), and 317 (negative) and in general anxiety disorder Study 308 (summary of study design and efficacy results are included in Appendix B). In studies 315, 316, and 317, the subject randomization was stratified by baseline sexual dysfunction status.

(b) (4)

Lu AA21004 is non-inferior to placebo with respect to treatment emergent sexual dysfunction.

Sponsor-proposed Analyses of ASEX

The main analysis was to assess treatment-emergent sexual dysfunction (TESD) based on the subset of subjects who do not have dysfunction at baseline. The non-inferiority was to be established through comparing the upper bound of the two-sided 95% confidence interval for the difference of the incidence rates between Lu AA21004 and placebo in subjects who developed sexual dysfunction at any time during the study period with a margin of 10 percentage points. Each subject who did not have sexual dysfunction at baseline, as based on ASEX scores, was evaluated for a shift to having sexual dysfunction. Within each treatment group *the percentage of*

subjects who shift (develop TESD) at any time during the study period was to be calculated. The presence of sexual dysfunction was defined as follow:

- ASEX total score of ≥ 19 , or
- score of ≥ 5 on any item, or
- score of ≥ 4 on any 3 items.

For the Phase III studies with sexual function assessed through ASEX, a weighted average of estimates with the weight being the sample sizes will be used for calculating the pooled incidence rates of TESD within each treatment groups. The formula for calculating the weighted average of estimate in a treatment group is as follow:

$$\hat{P} = \frac{N_1 * \hat{P}_1 + N_2 * \hat{P}_2 + .. + N_k * \hat{P}_k}{N_1 + N_2 + ... + N_k}$$

Where N_i and P_i , $i=1, 2, \dots, k$, are the number of subjects enrolled and estimated incidence rate for the treatment group from i^{th} study.

Confidence intervals (95% two-sided) for the differences between the incidence rates for each treatment group and placebo were to be constructed using the normal approximation to the binomial (Percent Treatment Group minus Percent Placebo).

Sponsor-proposed Non-inferiority Margin

Ten percentage points (10%) considering a placebo event rate of approximately 30% (according to Dr. Delgado).

Missing Data for ASEX Scale

When calculating the ASEX total score, if two or more individual items are missing, the total score was set to missing. If one individual item is missing, then the total score was calculated using a SAS function CEIL, as $CEIL[(\text{sum of non-missing items}) \times (\text{total number of items}) / (\text{number of non-missing items})]$. If the responses for the ASEX questionnaire were missing at a visit, or the subject did not answer the ASEX questionnaire, then the sexual dysfunction status was not determined for the subject at that visit. Data from other visits were used to evaluate whether the subject has TESD during the study.

Incidence Rates of Shifting from Normal to Abnormal in at least Two Consecutive Visits

As a supportive analysis, for each treatment group within the subgroup of no dysfunction at baseline in a study, the percent of subjects who transition to having sexual dysfunction measured in at least two consecutive visits during the study period was to be calculated. Confidence intervals (95% two-sided) for the differences of the incidence rates between each treatment group and placebo/duloxetine were to be constructed.

FDA Communication with the sponsor pertaining to analysis of sexual dysfunction

Division provided comments on ASEX analysis during Type C Meeting held on March 30, 2010. Additional comments were provided by the Division in the Advice/ Information request dated July 28, 2010. Brief summary of major recommendations is included below:

- The Division had no objection to conducting main analysis on the subgroup of subjects with no sexual dysfunction at baseline. However, the Division pointed out that patient randomization would need to be stratified by sexual function status at baseline.
- Division noted that the definition of shifting from normal to abnormal is based on any single visit and confirmation from additional visits is not required. Based on this definition, a patient is categorized as having shifted to "abnormal" even if he/she has only one visit showing an "abnormal" status. Division also noted that a more reasonable definition would require at least two consecutive visits with an abnormal rating.
- The Division conveyed to the sponsor that the only comparisons of interest would be for those doses that are shown to be effective. Analysis may be done by pooling trials with similar designs together given that trial is included as a factor in the statistical model, but it should not be done by pooling all effective doses together because the severity of adverse events generally depends on dose levels.
- The Division did not agree to the proposed NI margin since its clinical relevancy had not been justified.

Incidence rates of TESD

Table 33 provides overall summary of the number of patients with ASEX assessments and the number of patients without sexual dysfunction (SD) at baseline in studies 11984A, 13267A, 304, 308 (GAD study), 315, 316, and 317.

For all MDD studies combined, approximately 30% of subjects with ASEX assessment at Baseline did not have sexual dysfunction at Baseline. This was consistent across all LuAA21004 dose groups, duloxetine group and placebo group. In the GAD Study 308, the percentage of patients without sexual dysfunction was approximately 50%. Sample sizes (number of subjects without SD at Baseline) within studies, except GAD Study 308, were generally less than 50 subjects per treatment.

Table 33. Number of patients with ASEX assessment at Baseline

Study Name	Placebo	Lu AA21004					Duloxetine 60mg
		2.5mg	5mg	10mg	15mg	20mg	
11984A (non US, Can)							
Safety Set	148	155	157	151			155
ASEX at BL (a)	77	79	77	75			77
Dysfunction at BL	60	55	63	54			61
No Dysfunction at BL	16	24	14	21			15
13267A (outside North America)							
Safety Set	158				151	151	147
ASEX at BL (a)	158				150	151	147
Dysfunction at BL	131				109	116	104
No Dysfunction at BL	27				41	35	42
304 (US)							
Safety Set	151	149	153				150
ASEX at BL (a)	150	149	153				148
Dysfunction at BL	108	99	103				96
No Dysfunction at BL	42	49	48				50
315* (US)							
Safety set	159				147	154	150
ASEX at BL (a)	159				147	154	150
Dysfunction at BL	101				101	105	102
No Dysfunction at BL	58				45	45	47
316* (US)							
Safety set	157			155		150	
ASEX at BL (a)	157			155		150	
Dysfunction at BL	105			104		100	
No Dysfunction at BL	52			51		50	
317* (US)							
Safety Set	160			154	151		
ASEX at BL (a)	160			154	151		
Dysfunction at BL	121			112	109		
No Dysfunction at BL	39			42	42		
308 (US)							
Safety Set	155	156	155	156			154
ASEX at BL (a)	155	156	155	156			153
Dysfunction at BL	72	76	76	77			78
No Dysfunction at BL	82	80	79	79			75
All studies combined							
Safety Set	1088	460	465	616	449	455	756
ASEX at BL (a)	1016	384	385	540	448	455	675
Dysfunction at BL	698	230	242	347	319	321	441
No Dysfunction at BL	316	153	141	193	129	134	230
(Male, Female)	(169, 147)	(72, 81)	(72, 69)	(89, 104)	(70, 59)	(65, 69)	(111, 119)

* Patient randomization was stratified by ASEX status at baseline.

(a) Subjects with at least one ASEX question answered regardless of whether baseline dysfunction status could be determined.

Source: Sponsor's Response to FDA Request Table 4.1 ([\\Cdsub1\evsprod\NDA204447\0021](#))

Table 34 provides overall summary of the incidence of the sexual dysfunction at any visit during double-blind treatment period in short-term efficacy studies. The incidence rates for each dose varied substantially from study to study. For example, the incidence rate of TESD for 10mg dose ranged from 24.7% (in Study 11984A) to 48.0% (in Study 316). Similarly the incidence rates in 20mg dose groups varied from 35% to 65%.

Incidence rates for 2.5 mg and 5mg were investigated in three studies (11984A, 304 and 308) and in all three studies the observed incidence rates for 2.5mg were substantially higher. In the dose range of 5 to 20 mg, the observed incidence of TESD for all studies combined increased with the dose from 25.7% (5 mg) to 46.1% (20 mg). For the lower dose of 2.5mg, the observed incidence rate of 42.4% was noted. The observed incidence rates for 2.5mg, 15mg and 20mg were larger than placebo by more than 10 percentage points. In the female subgroup the observed incidence rates were higher than in males by 5-10% in all treatment arms. The proportions of the males and females varied from arm to arm but on average were balanced among treatment arms. Tables with incidence rates for males and females by study are provided in the Appendix D.

Table 34. Incidence of TESD by Study in subjects without sexual dysfunction at baseline

Study Name	Placebo	Lu AA21004					Duloxetine 60mg
		2.5mg	5mg	10mg	15mg	20mg	
11984A (non US, Can)	7/15 (46.7%)	11/23 (47.8%)	3/14 (21.4%)	10/21 (47.6%)			9/15 (60.0%)
13267A (out. N. America)	12/27 (44.4%)				22/40 (55.0%)	23/35 (65.7%)	18/41 (43.9%)
304 (US)	14/42 (33.3%)	25/49 (51.0%)	18/48 (37.5%)				23/49 (46.9%)
308 (US) [GAD]	20/81 (24.7%)	28/79 (35.4%)	14/74 (18.9%)	19/77 (24.7%)			34/74 (45.9%)
315 (US)	21/58 (36.2%)				16/45 (35.6%)	16/45 (35.6%)	25/47 (53.2%)
316 (US)	14/50 (28.0%)			24/50 (48.0%)		20/48 (41.7%)	
317 (US)	11/36 (30.6%)			14/42 (33.3%)	16/41 (39.0%)		
All studies	99/309 (32.0%)	64/151 (42.4%)	35/136 (25.7%)	67/190 (35.3%)	54/126 (42.9%)	59/128 (46.1%)	109/226 (48.2%)
All studies: Males	49/167 (29.3%)	27/72 (37.5%)	15/69 (21.7%)	26/89 (29.2%)	27/68 (39.7%)	26/61 (42.6%)	50/108 (46.3%)
All studies: Females	50/142 (35.2%)	37/79 (46.8%)	20/67 (29.9%)	41/101 (40.6%)	27/58 (46.6%)	33/67 (49.3%)	59/118 (50.0%)

Source: Integrated Summary of Safety Table 5.3.2.1 (pg. 4409), Sponsor's Response to FDA Request Tables 4.2 and 4.3

Incidence rates of TESD using alternative definition (meeting the criterion in at least two consecutive visits).

Data on the incidence of sexual dysfunction observed at two consecutive visits is presented in Table 35. Naturally, for this definition of TESD the range and the variability of the incidence rates were lower compared with the definition that uses TESD shift at any visit. Similar trends for the incidence rates were observed. In the dose range of 5 to 20 mg, the observed incidence of TESD increased with the dose from 18.9% (5 mg) to 31.7% (20 mg). For the lower dose of

2.5mg, the observed incidence rate (22.7%) was higher than the incidence rates for 5mg and 10mg doses. Overall, the incidence rates for the same dose varied substantially from study to study.

In female subgroup the observed incidence rates were higher than in males in all treatment arms. On average, the proportions of the males and females were balanced among treatment arms.

Table 35. Incidence of TESD by study at two consecutive visits in subjects without sexual dysfunction at baseline.

Study Name	Placebo	Lu AA21004					Duloxetine
		2.5mg	5mg	10mg	15mg	20mg	60mg
11984A (non US, Can)	3/15 (20.0%)	9/23 (39.1%)	3/14 (21.4%)	7/21 (33.3%)			6/15 (40.0%)
13267A (out N. America)	7/27 (25.9%)				12/40 (30.0%)	14/34 (41.2%)	8/41 (19.5%)
304 (US)	6/38 (15.8%)	12/43 (27.9%)	13/47 (27.7%)				14/44 (31.8%)
308 [GAD]	9/79 (11.4%)	11/75 (14.7%)	9/71 (12.7%)	12/73 (16.4%)			15/63 (23.8%)
315 (US)	11/57 (19.3%)				11/43 (25.6%)	12/45 (26.7%)	14/45 (31.1%)
316 (US)	9/48 (18.8%)			13/47 (27.7%)		14/47 (29.8%)	
317 (US)	4/33 (12.1%)			7/39 (17.9%)	9/41 (22.0%)		
All studies	49/297 (16.5%)	32/141 (22.7%)	25/132 (18.9%)	39/180 (21.7%)	32/124 (25.8%)	40/126 (31.7%)	57/208 (27.4%)
All studies: Males	22/162 (13.6%)	15/67 (22.4%)	11/67 (16.4%)	17/86 (19.8%)	13/67 (19.4%)	17/59 (28.8%)	26/99 (26.3%)
All studies: Females	27/135 (20.0%)	17/74 (23.0%)	14/65 (21.5%)	22/94 (23.4%)	19/57 (33.3%)	23/67 (34.3%)	31/109 (28.4%)

Source: Integrated Summary of Safety Table 5.3.2.2.4 (pg. 4416), Sponsor's Response to FDA Request Tables 4.2 and 4.3

Conclusion

The observed incidence of treatment emergent sexual dysfunction (TESD) for all studies combined had tendency to increase with the dose in the dose range of 5-20 mg. However, incidence rates varied substantially from study to study within each dose. In female subgroup the observed incidence rates were higher than in males in all treatment arms. The sponsor did not pursue any labeling claims pertaining to the TESD.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

This section contains the sponsor's and the reviewer's results of the exploratory subgroup analysis for all positive short-term studies and for relapse prevention study 11985A.

4.1.1 Study 11492A (non-US multiregional, including Canada)

Table 36 displays reviewer’s subgroup analysis by gender, race, and geographic region. There were only 5 patients (in total) in the racial subgroups other than Caucasian and Asian, and, thus, the subgroup analysis was not conducted in these subgroups. Overall, in all subgroups except Asian racial subgroup, both Lu AA21004 treatment arms were numerically better than placebo. Subgroup analysis by age was not performed since there were no patients older than 65 years.

Table 36. MADRS Total Score Change From Baseline at Week 6 (FAS, LOCF, ANCOVA)

Subgroups	Placebo	Venlafaxine	Lu AA21004	
			5mg	10mg
Gender: Male	N=36	N=51	N=38	N=34
LS Mean Change from Baseline (SE)	-14.2 (1.85)	-19.7 (1.60)	-20.0 (1.77)	-20.7 (1.92)
Gender: Female	N=69	N=61	N=70	N=66
LS Mean Change from Baseline (SE)	-15.2 (1.40)	-22.4 (1.45)	-21.5 (1.34)	-20.7 (1.37)
Race: Caucasian	N=98	N=103	N=101	N=89
LS Mean Change from Baseline (SE)	-14.0 (1.08)	-21.0 (1.06)	-20.4 (1.07)	-20.3 (1.13)
Race: Asian	N=6	N=8	N=7	N=8
LS Mean Change from Baseline (SE)	-16.0 (3.81)	-12.0 (3.53)	-15.5 (3.69)	-17.1 (2.69)
Region: North America (Canada)	N=9	N=8	N=8	N=7
LS Mean Change from Baseline (SE)	-17.2 (3.68)	-12.7 (3.87)	-19.9 (3.80)	-20.0 (4.06)
Region: outside of North America	N=96	N=104	N=100	N=93
LS Mean Change from Baseline (SE)	-14.3 (1.06)	-21.6 (1.02)	-20.5 (1.04)	-20.2 (1.07)

Source: Reviewer’s results

4.1.2 Study 305 (outside North America)

Table 37 displays the sponsor’s subgroup analysis by gender and reviewer’s subgroup analyses by race and age. The non-Caucasian racial subgroups were combined for the analysis because these subgroups had only a few patients. In all gender and age subgroups, and in the Caucasian racial subgroup, all three Lu AA21004 treatment arms were numerically better than placebo.

Table 37. HAM-D24 Total Score Change From Baseline at Week 8 (FAS, MMRM)

Subgroups	Placebo	Lu AA21004		
		1mg	5mg	10mg
Gender: Male	N=54	N=46	N=52	N=55
LS Mean Change from Baseline (SE)	-8.7 (1.21)	-14.5 (1.31)	-14.6 (1.23)	-15.7 (1.22)
Gender: Female	N=85	N=93	N=87	N=84
LS Mean Change from Baseline (SE)	-13.0 (0.95)	-15.2 (0.92)	-16.0 (0.94)	-16.4 (0.97)
Race: Caucasian	N=120	N=128	N=119	N=113
LS Mean Change from Baseline (SE)	-10.3 (0.79)	-14.8 (0.77)	-15.1 (0.80)	-16.4 (0.81)
Race: Other	N=19	N=11	N=20	N=26
LS Mean Change from Baseline (SE)	-16.1 (2.03)	-12.3 (2.86)	-16.0 (2.1)	-14.0 (2.16)
Age: 65 years or younger	N=129	N=134	N=132	N=130
LS Mean Change from Baseline (SE)	-11.4 (0.78)	-14.8 (0.77)	-15.4 (0.77)	-16.3 (0.79)
Age: older than 65 years	N=10	N=5	N=7	N=9
LS Mean Change from Baseline (SE)	-11.1 (2.35)	-18.8 (3.37)	-14.7 (3.05)	-13.4 (2.69)

Source: Clinical Study Report Table 11.b (pg.65), and reviewer’s results

4.1.3 Study 13267A (outside North America)

Table 38 displays sponsor's subgroup analysis by gender and reviewer's subgroup analysis by age. In both gender subgroups and in the age subgroup consisting of patients who are not older than 65 years, both Lu AA21004 treatment arms and the duloxetine arm were numerically better than placebo. Since there were only 11 patients (in total) in the racial subgroups other than Caucasian, the subgroup analysis by race was not performed by this reviewer.

Table 38. MADRS Total Score Change From Baseline at Week 8 (FAS, MMRM)

Subgroups	Placebo	Duloxetine 60mg	Lu AA21004	
			15mg	20mg
Gender: Male	N=48	N=44	N=52	N=60
LS Mean Change from Baseline (SE)	-11.6 (1.33)	-22.8 (1.38)	-16.4 (1.31)	-18.6 (1.21)
Gender: Female	N=110	N=102	N=97	N=91
LS Mean Change from Baseline (SE)	-11.5 (0.94)	-20.4 (0.95)	-17.5 (1.01)	-18.33 (1.05)
Age: younger than 65 years	N=142	N=137	N=135	N=136
LS Mean Change from Baseline (SE)	-11.2 (0.79)	-21.2 (0.79)	-17.3 (0.84)	-18.9 (0.82)
Age: 65 years or older	N=16	N=10	N=17	N=15
LS Mean Change from Baseline (SE)	-16.7 (2.57)	-21.6 (3.06)	-14.3 (2.29)	-16.8 (2.48)

Source: Clinical Study Report Table 11.h (pg. 83-84) and reviewer's results.

4.1.4 Study 315 (US)

Table 39 displays sponsor's subgroup analysis by gender and by race in Study 315. There were only 8 patients (in total) in the racial subgroups other than Caucasian and Black, and, thus, the subgroup analysis was not conducted in these subgroups. Overall, in all subgroups except male gender subgroup, both Lu AA21004 treatment arms and the duloxetine arm were numerically better than placebo. The overwhelming majority of patients were younger than 65 years. Among patients older than 65 years, there were only 4 patients in each of the two Lu AA21004 arms and 6 patients in the placebo arm, and, hence, subgroup analysis by age was not performed by this reviewer.

Table 39. MADRS Total Score Change From Baseline at Week 8 (FAS, MMRM)

Subgroups	Placebo	Duloxetine 60mg	Lu AA21004	
			15mg	20mg
Gender: Male	N=42	N=32	N=43	N=37
LS Mean Change from Baseline (SE)	-13.8 (1.52)	-17.3 (1.78)	-12.1 (1.58)	-12.8 (1.70)
Gender: Female	N=111	N=114	N=102	N=110
LS Mean Change from Baseline (SE)	-12.1 (0.99)	-17.0 (1.02)	-14.7 (1.07)	-16.4 (1.03)
Race: Caucasian	N=118	N=115	N=113	N=111
LS Mean Change from Baseline (SE)	-12.5 (0.92)	-16.1 (1.00)	-14.1 (1.00)	-15.0 (1.01)
Race: Black	N=33	N=30	N=30	N=33
LS Mean Change from Baseline (SE)	-13.2 (2.17)	-19.7 (2.13)	-15.3 (2.24)	-18.1 (2.07)

Source: Clinical Study Report Table 11.h (pg. 83-84)

4.1.5 Study 316 (US)

Table 40 displays sponsor's subgroup analysis by gender and race in Study 316. There were only 10 patients (in total) in the racial subgroups other than Caucasian and Black, and, thus, the subgroup analysis was not conducted in these subgroups. Overall, in all subgroups, both Lu AA21004 treatment arms were numerically better than placebo. There were only 8 patients (in total) of age older than 65, and, hence, subgroup analysis by age was not performed by this reviewer.

Table 40. MADRS Total Score Change From Baseline at Week 8 (FAS, MMRM)

Subgroups	Placebo	Lu AA21004	
		10mg	20mg
Gender: Male	N=46	N=37	N=41
LS Mean Change from Baseline (SE)	-9.6 (1.5)	-10.8 (1.6)	-15.0 (1.6)
Gender: Female	N=109	N=117	N=107
LS Mean Change from Baseline (SE)	-11.3 (0.99)	-13.8 (0.99)	-14.1 (1.01)
Race: Caucasian	N=118	N=106	N=97
LS Mean Change from Baseline (SE)	-11.4 (0.94)	-13.7 (1.04)	-14.7 (1.06)
Race: Black	N=37	N=42	N=47
LS Mean Change from Baseline (SE)	-10.6 (1.62)	-12.3 (1.49)	-14.6 (1.48)

Source: Clinical Study Report Table 11.h (pg. 80-81)

4.1.6 Study 12541A (elderly, multiregional including US and Canada)

Table 41 displays reviewer's subgroup analysis by gender, race and geographic region for the primary efficacy endpoint (HAM-D-24). There were only 6 patients (in total) in the racial subgroups other than Caucasian and Black, and, thus, the subgroup analysis was not conducted in these subgroups. Overall, in all subgroups except Black racial subgroup, the Lu AA21004 treatment arm and the duloxetine arm were numerically better than placebo. The North America region included USA and Canada with 44 and 4 sites respectively. The observed treatment effect of Lu AA21004 compared with Placebo in North America was very small (approximately 0.6).

Table 41. HAM-D-24 Total Score Change From Baseline at Week 8 (FAS, LOCF, ANCOVA)

Subgroups	Placebo	Duloxetine 60mg	Lu AA21004 5mg
Gender: Male	N=55	N=49	N=49
LS Mean Change from Baseline (SE)	-11.4 (1.35)	-16.6 (1.54)	-13.5 (1.54)
Gender: Female	N=90	N=99	N=106
LS Mean Change from Baseline (SE)	-10.6 (8.7)	-15.3 (8.18)	-14.8 (10.04)
Race: Caucasian	N=139	N=141	N=144
LS Mean Change from Baseline (SE)	-9.9 (0.78)	-16.0 (0.77)	-13.5 (0.77)
Race: Black	N=4	N=6	N=8
LS Mean Change from Baseline (SE)	-17.3 (5.05)	-7.0 (3.58)	-14.2 (2.9)
Region: North America (US and Canada)	N=58	N=62	N=61
LS Mean Change from Baseline (SE)	-10.2 (1.24)	-12.8 (1.19)	-10.8 (1.21)
Region: outside North America	N=87	N=86	N=94
LS Mean Change from Baseline (SE)	-10.5 (0.95)	-17.9 (0.96)	-15.5 (0.92)

Source: Reviewer's Results

4.1.7 Study 11985A (maintenance, multiregional, non US, multi-regional including Canada)

Table 42 displays sponsor’s subgroup analysis by gender and by race and reviewer’s subgroup analysis by geographic region. There were only 11 patients (in total) in the racial subgroups other than Caucasian and Asian, and, thus, the subgroup analysis was not conducted in those subgroups. Overall, in all subgroups except Asian racial subgroup, Lu AA21004 treatment arm was numerically better than placebo.

Table 42. Time to relapse within 24 Weeks of Double-blind Period (FAS, MMRM)

Subgroups	Placebo	Lu AA21004	Hazard Ratio
Gender: Male	N= 72	N=74	
Number of Relapsing Patients (%)	18 (25%)	9 (12%)	2.00
Gender: Female	N=120	N=130	
Number of Relapsing Patients (%)	32 (27%)	18 (14%)	2.01
Race: Caucasian	N=154	N=165	
Number of Relapsing Patients (%)	46 (30%)	21 (13%)	2.47
Race: Asian	N=34	N=32	
Number of Relapsing Patients (%)	3 (9%)	6 (19%)	0.42
Region: North America (Canada)	N=11	N=14	
LS Mean Change from Baseline (SE)	7 (64%)	3 (21%)	4.13
Region: outside North America	N=181	N=190	
LS Mean Change from Baseline (SE)	43 (24%)	24 (13%)	1.89

Source: Clinical Study Report Panel 31 (pg. 90), Reviewer’s results

4.2 Other Special/Subgroup Populations

No other subgroups were analyzed.

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

There were no issues with the statistical analysis methods in the efficacy studies.

5.2 Collective Evidence

Short-term efficacy

Primary Efficacy Results for the short-term studies

The acute efficacy of Lu AA21004 in the treatment of MDD has been evaluated in 10 short-term placebo-controlled studies (9 in adults and 1 in elderly subjects). The primary efficacy results of six positive short-term studies are summarized in Table 43. The primary efficacy assessment tools in these studies were either the MADRS or HAM-D24. In four of the six positive studies the primary efficacy variable was MADRS total score. In the other two studies, it was HAMD-24 total score. The treatment effect of Lu AA21004 with respect to placebo was demonstrated in 5mg, 10mg, 15mg, and 20mg doses. The only dose that did not have replication of the efficacy was 15mg.

In the two positive studies conducted solely in US (studies 315 and 316) no dose lower than 20 mg was shown statistically significantly better than placebo. Although the treatment effect of 5 mg was demonstrated in study 12541A (elderly patients), the observed LuAA21004 effect appeared essentially no different from the placebo effect in North American sites, which accounted for approximately 40% of the randomized patients as shown in Table 41. Also, out of four negative studies three studies were conducted exclusively in the US. It is not clear whether the difference in the effective dose range between North America region (US and Canada) and the regions outside North America can be attributed to any factor.

Table 43. Primary efficacy results for positive efficacy studies

Study Number	Primary Endpoint	Dose				
		1mg	5mg	10mg	15mg	20mg
11492A	MADRS	Region: non-US, multi-regional including Canada. Study Conduct Dates: 08/2006-08/2007				
	LS mean difference (SE)		-5.9 (1.4)	-5.7 (1.4)		
	p-value (unadjusted)		<0.001	<0.001		
	Significance (MCP adjusted)		Yes	Yes		
305	HAMD-24	Region: outside North America. Study Conduct Dates: 08/2008-08/2009				
	LS mean difference (SE)	-3.5 (1.0)	-4.1 (1.0)	-4.9 (1.0)		
	p-value (unadjusted)	<0.001	<0.001	<0.001		
	Significance (MCP adjusted)	N/A[#]	No*	Yes		
13267A	MADRS	Region: outside North America. Study Conduct Dates: 05/2010-09/2011				
	LS mean difference (SE)				-5.5 (1.1)	-7.1 (1.1)
	p-value (unadjusted)				<0.001	<0.001
	Significance (MCP adjusted)				Yes	Yes
315	MADRS	Region:US. Study Conduct Dates: 06/2010-02/2012				
	LS mean difference (SE)				-1.5 (1.21)	-2.8 (1.21)
	p-value (unadjusted)				0.224	0.023
	Significance (MCP adjusted)				No	Yes
316	MADRS	Region:US. Study Conduct Dates: 07/2010-01/2012				
	LS mean difference (SE)			-2.2 (1.15)		-3.6 (1.16)
	p-value (unadjusted)			0.058		0.002
	Significance (MCP adjusted)			No		Yes
12541A (elderly)	HAMD-24	Region: multi-regional, including US and Canada. Study Conduct Dates: 01/2009-02/2010				
	LS mean difference (SE)		-3.3			
	p-value (unadjusted)		0.001			
	Significance (MCP adjusted)		Yes			

Source: Reviewer's summary based on sponsor's clinical study reports

*Since 10mg dose was not statistically significantly different from placebo in the first key secondary variable in the testing sequence (SDS), the formal testing was stopped according to the pre-specified hierarchical multiple testing procedure. None of the subsequent null hypotheses (including hypotheses associated with 5mg dose) in the pre-specified testing hierarchy are considered statistically significantly different from placebo.

[#] 1mg arm was not formally tested against placebo.

Key secondary efficacy results

The sponsor pre-specified different multiple key-secondary endpoints and different hierarchical testing procedures varying from study to study. The list includes change from baseline in SDS, CGI-I, HAM-D24 response rate, MADRS remission rate, and change from baseline in HAM-D24 in subjects with baseline HAM-A ≥ 20 . As was indicated in the FDA Advice Letter dated September 14, 2010, the only acceptable key secondary endpoints are either CGI-I or SDS. However, it would be redundant to use both as key secondary endpoints. The CGI-I was assessed in all positive efficacy studies whereas SDS was not assessed in studies 11492A and 12541A. Also, in the studies where CGI-I was considered as an exploratory endpoint, the efficacy results in CGI-I supported primary efficacy findings. Thus, in this program, CGI-I would be more informative as a key secondary endpoint. Summary of efficacy results on both endpoints (CGI-I and SDS) are provided below.

Summary of efficacy results on CGI-I:

Summary of efficacy results in CGI-I is presented in Table 44. In four out of six positive short-term studies, CGI-I score was declared as a key secondary variable. Statistically significant treatment differences were shown in one study (13267A) at investigated doses of 15mg and 20mg.

Supportive results: In study 305, all three investigated doses (1mg, 5mg, and 10mg) had nominal p-values < 0.001 . However, these p-values could serve only as supportive evidence because of the pre-specified multiple testing hierarchy in this study. The sponsor pre-specified four key secondary endpoints and they were tested sequentially within each dose starting from 10mg dose. The testing stopped at SDS due to non-significance, and as a result, other key secondary endpoint were not tested. The 1mg dose arm was exploratory and was not included in the multiple testing procedure.

In addition, nominal p-values for CGI-I were very small in study 11492 (5 mg and 10 mg) and in study 12541A (5mg). In these studies CGI-I was considered exploratory variable and, thus, no formal statistical testing was conducted. However, the observed nominal p-values could serve as additional supportive evidence of efficacy with respect to CGI-I.

Table 44. Summary Results of CGI-I Analysis for Positive Acute Efficacy Studies

Study Number	CGI Endpoint status	Dose				
		1mg	5mg	10mg	15mg	20mg
11492A	Exploratory					
LS mean difference (SE)			-0.6 (0.16)	-0.6 (0.16)		
p-value (unadjusted)			<0.001	<0.001		
Significance (based on MCP)			N/A	N/A		
305	Key Secondary					
LS mean difference (SE)		-0.5(0.13)	-0.5 (0.13)	-0.6 (0.13)		
p-value (unadjusted)		<0.001	<0.001	<0.001		
Significance (based on MCP)*		N/A	No*	No*		
13267A	Key Secondary					
LS mean difference (SE)					-0.7 (0.13)	-1.0 (0.13)
p-value (unadjusted)					<0.001	<0.001
Significance (based on MCP)					Yes	Yes
315	Key Secondary					
LS mean difference (SE)					-0.1 (0.14)	-0.2 (0.14)
p-value (unadjusted)					0.400	0.177
Significance (based on MCP)					No	No
316	Key Secondary					
LS mean difference (SE)				-0.2 (0.13)		-0.3 (0.13)
p-value (unadjusted)				0.119		0.024
Significance (based on MCP)				No		No [#]
12541A	Exploratory					
LS mean difference (SE)			-0.6 (0.13)			
p-value (unadjusted)			<0.0001			
Significance (based on MCP)			N/A			

Source: Reviewer's summary based on sponsor's clinical study reports

*Since 10mg dose failed to beat placebo in the first key secondary variable in the sequence (SDS), the formal testing was stopped according to the pre-specified hierarchical testing. None of the subsequent endpoints in the pre-specified testing hierarchy are considered statistically significantly different from placebo.

[#]Since 20mg dose failed to beat placebo in the first key secondary variable in the sequence (MADRS response rate at Week 8), the formal testing was stopped according to the pre-specified hierarchical testing. None of the subsequent endpoints in the pre-specified testing hierarchy are considered statistically significantly different from placebo.

Summary of efficacy results on SDS:

In SDS score, the Lu AA21004 treatment was statistically significantly better than placebo in 15 mg and 20mg doses, but the statistical evidence was shown in one study only (study 13267A).

Summary of efficacy results in SDS is provided in Table 45.

Table 45. Summary Results of SDS Analysis for Positive Acute Efficacy Studies.

Study Number	CDS Endpoint status	Dose				
		1mg	5mg	10mg	15mg	20mg
11492A	Not assessed					
LS mean difference (SE)			N/A	N/A		
p-value (unadjusted)			N/A	N/A		
Significance (based on MCP)			N/A	N/A		
305	Key Secondary					
LS mean difference (SE)		-0.1 (1.01)	-1.1 (0.99)	-1.5 (1.03)		
p-value (unadjusted)		0.963	0.263	0.135		
Significance (based on MCP)		N/A	No	No		
13267A	Key Secondary					
LS mean difference (SE)					-3.2 (1.16)	-3.9 (1.11)
p-value (unadjusted)					0.005	<0.001
Significance (based on MCP)					Yes	Yes
315	Key Secondary					
LS mean difference (SE)					-0.1 (1.11)	-0.9 (1.10)
p-value (unadjusted)					0.962	0.427
Significance (based on MCP)					No	No
316	Key Secondary					
LS mean difference (SE)				-1.4 (1.04)		-2.4 (1.07)
p-value (unadjusted)				0.183		0.025
Significance (based on MCP)				No		No
12541A	Not assessed					
LS mean difference (SE)			N/A			
p-value (unadjusted)			N/A			
Significance (based on MCP)			N/A			

Source: Reviewer's summary based on sponsor's clinical study reports

Maintenance effect

Treatment by Lu AA21004 (5mg or 10mg) demonstrated statistically significant superiority over placebo in the relapse prevention based on Cox proportional hazard model ($p=0.0035$).

According to the study protocol, in order to get randomized patients were required to be stable during the last two weeks of the open-label phase (MADRS total score ≤ 10 at Weeks 10 and 12). The protocol was submitted to the Division after this non-US study was initiated. At the End of Phase II Meeting (February 5, 2008) the Division pointed out that the stabilization period of 12 weeks is too short. Based on criteria of sustained MADRS total score ≤ 10 , approximately 60% of patients in each treatment arm were stable for at least 4 weeks prior to randomization and less than 40% of the patients were stable for 6 weeks or more. Only a few patients in both treatment arms were stable for 10 weeks.

Sexual Dysfunction based on ASEX

The observed incidence of treatment emergent sexual dysfunction (TESD) for all studies combined had tendency to increase with the dose in the dose range of 5-20 mg. However, incidence rates varied substantially from study to study within each dose. In female subgroup the observed incidence rates were higher than in males in all treatment arms. The sponsor did not pursue any labeling claims pertaining to the TESD.

5.3 Conclusions and Recommendations

The treatment effect of Lu AA21004 with respect to placebo was demonstrated in 5mg, 10mg, 15mg, and 20mg doses). No dose lower than 20 mg was shown statistically significantly better than placebo in US studies.

Treatment by Lu AA21004 (5mg or 10mg) was statistically significantly superior to placebo in relapse prevention based on one maintenance study.

APPENDIX A. BRIEF SUMMARY OF NEGATIVE AND FAILED STUDIES

Study 11984A

Study Design

Primary objective was to evaluate the efficacy of three fixed dosages of Lu AA21004 (2.5, 5, or 10mg/day) *versus* placebo after 8 weeks of treatment in patients with MDD, the primary efficacy analysis considers only the 5 and 10 mg doses.

The study was conducted at 100 sites in 20 countries – Australia, Bulgaria, Canada, Czech Republic, Estonia, Finland, France, Hong Kong, India, Republic of Korea, Latvia, Lithuania, Malaysia, Philippines, Romania, Slovakia, Spain, Taiwan, Turkey, and Ukraine. This was a multi-national, randomized, double-blind, parallel-group, placebo-controlled, active-reference (duloxetine), fixed-dose study. Patients were randomized equally (1:1:1:1:1) to placebo, Lu AA21004 2.5mg/day, Lu AA21004 5mg/day, Lu AA21004 10mg/day, or duloxetine 60mg/day for 8 weeks of double-blind treatment (8-week Core Treatment Period). Efficacy and safety data were collected after 1 and 2 weeks and then at 2-week intervals.

The primary efficacy measure was the MADRS total score.

Primary Efficacy Analysis

The primary efficacy analysis was an analysis of covariance (ANCOVA) of the change from baseline in MADRS total score at Week 8 (FAS, last observation carried forward [LOCF]), with treatment and centre as factors and the baseline MADRS total score as a covariate. The results of the primary efficacy analysis are presented in Table 46. The mean treatment differences to placebo in all Lu AA21004 dose groups and in the duloxetine group were not statistically significant. Since there was no statistically significant treatment difference to placebo with the active reference, this study is considered failed.

Table 46. MADRS Total Score Change From Baseline at Week 8 (FAS, LOCF, ANCOVA)

HAMD-24 Total score	Placebo (N=145)	Duloxetine (N=149)	Lu AA21004		
			2.5mg (N= 155)	5mg (N=155)	10mg (N=151)
LS Mean Change from Baseline	-14.8 (0.82)	-16.8 (0.81)	-16.2 (0.79)	-16.5 (0.80)	-16.3 (0.80)
Difference from Placebo		-2.0 (1.14)	-1.4 (1.12)	-1.7 (1.13)	-1.5 (1.13)
p-value		0.074	0.219	0.132	0.185

Source: Clinical Study Report Panel 22 (pg. 82)

Study 303

Study Design

The primary objective of the study was to evaluate the efficacy of 5 mg Lu AA21004 QD compared with placebo after 6 weeks of treatment in subjects with MDD.

The study enrolled patients at 49 sites in the United States. It was a randomized, double-blind, placebo-controlled, parallel-group study. Eligible subjects were randomized (1:1) to treatment with either Lu AA21004 5 mg QD or placebo QD. After the Baseline Visit, subjects entered the 42-day (6-week) double-blind core treatment period. During the Treatment Period, subjects returned to the study site every week for evaluations.

The primary endpoints were the change from baseline in HAM-D24 total score at Week 6 and at each week assessed.

Primary Efficacy Analysis

The LS mean change from baseline in HAM-D24 total score at Week 6 was evaluated using analysis of covariance (ANCOVA), with treatment and center as fixed factors, baseline HAM-D24 as covariate, and using the LOCF technique. The results of the primary efficacy analysis are presented in Table 47. Lu AA21004 5 mg was not statistically significantly better than placebo.

Table 47. HAM-D-24 Total Score Change From Baseline at Week 6 (FAS, LOCF, ANCOVA)

	Placebo N=286	Lu AA21004 5mg N=292
LS Mean Change from Baseline (SE)	-13.9 (0.662)	-14.6 (0.650)
Difference from Placebo (SE)		-0.7 (0.887)
p-value		0.407

Source: Clinical Study Report Table 11.a (pg. 74)

Study 304

Study Design

The primary objective of the study was to evaluate the efficacy of 2 fixed doses of Lu AA21004 (2.5 and 5 mg QD) versus placebo after 8 weeks of treatment in subjects with MDD.

The study was conducted at 47 sites in the United States. It was a randomized, double-blind, placebo-controlled, duloxetine referenced, parallel-group study. At Baseline, eligible subjects were randomized, at a ratio of 1:1:1:1, to receive Lu AA21004 2.5 mg QD, Lu AA21004 5 mg QD, duloxetine 60 mg QD, or placebo QD during a 56-day (8-week) double-blind Treatment Period. Subjects were evaluated weekly during the first 2 weeks of treatment, and then every 2 weeks up to the end of the 8-week Treatment Period.

The primary efficacy endpoint was the least squares (LS) mean change from Baseline in HAM-D24 total score, after 8 weeks of treatment.

Primary Efficacy Analysis

Comparisons between the doses of Lu AA21004 (2.5 mg and 5 mg) and placebo was performed using analysis of covariance (ANCOVA), with treatment and center as fixed factors, baseline HAM-D24 as covariate, and using the LOCF technique. The results of the primary efficacy analysis are presented in Table 48. Neither of Lu AA21004 treatment arms was statistically significantly better than placebo in reducing the HAM-D24 total score at Week 8.

Table 48. HAM-D24 Total Score Change From Baseline at Week 8 (FAS, LOCF)

Characteristic	Placebo (N=149)	Duloxetine 60mg (N=149)	Lu AA21004	
			2.5mg (N=146)	5mg (N=153)
LS Mean Baseline Score (SE)	29.1 (0.44)	28.8 (0.43)	29.7 (0.43)	29.6 (0.43)
LS Mean Change from Baseline (SE)	-10.5 (0.76)	-13.5 (0.75)	-12.0 (0.74)	-11.1 (0.74)
Difference from Placebo (SE)		-3.0 (1.05)	-1.5 (1.04)	-0.6 (1.04)
p-value		0.005	0.138	0.577

Source: Clinical Study Report Table 11.a (pg.67)

Study 317

Study Design

The primary objective of study 316 was to evaluate the efficacy of Lu AA21004 10 and 15 mg QD compared with placebo as assessed by the MADRS after 8 weeks of treatment in subjects with MDD.

The study enrolled patients at 65 sites in the United States. At Baseline (Day 0), subjects were randomized in a 1:1:1 ratio to 1 of 3 treatment arms to receive once daily oral doses of Lu AA21004 10 mg, Lu AA21004 15 mg, or placebo in the 8-week, double-blind treatment period. Subject randomization was stratified by subject's baseline sexual function status (normal or abnormal decided by Arizona Sexual Experiences Scale [ASEX] score). All subjects took their first dose of study medication on Day 1, the day after randomization. Subjects assigned to the Lu AA21004 15 mg arm received Lu AA21004 10 mg for the first week of the Double-Blind Treatment Period and 15 mg for the remaining 7 weeks of treatment. Subjects visited the site weekly during the first 2 weeks of treatment and then every 2 weeks up to the end of the 8-week Treatment Period

Mean change from Baseline in MADRS total score at week 8 was the primary endpoint.

Primary Efficacy Analysis

Primary analysis was based on a mixed model for repeated measurements (MMRM) analysis of covariance with treatment, center, week, treatment-by-week interaction, baseline MADRS total score-by-week as fixed effects. An unstructured covariance matrix was assumed. The results of the primary efficacy analysis are presented in Table 49. Neither of Lu AA21004 treatment arms was statistically significantly better than placebo in reducing the MADRS total score at Week 8. Remark: Baseline LS means and P-values were from an ANOVA model with terms for treatment and center.

Table 49. Study 317 MADRS Total Score Change From Baseline at Week 8 (FAS, MMRM)

Characteristic	Placebo (N=149)	Lu AA21004	
		10mg (N=143)	15mg (N=142)
LS Mean Baseline Score (SE)	33.4 (0.36)	34.1 (0.37)	33.6 (0.37)
LS Mean Change from Baseline (SE)	-12.9 (1.04)	-13.7 (1.06)	-13.4 (1.09)
Difference from Placebo (SE)		-0.8 (1.49)	-0.5 (1.50)
p-value		0.597	0.745

Source: Clinical Study Report Table 11.f (pg. 76)

APPENDIX B. BRIEF SUMMARY OF STUDY 308 (GAD)

The study was conducted in 72 sites in the United States from June 2008 to February 2009. This was a Phase 3, multicenter, randomized, double-blind, placebo-controlled, active-referenced (duloxetine 60 mg QD), parallel-group study of 3 doses of Lu AA21004 (2.5, 5, or 10 mg) in acute treatment of adults with Generalized Anxiety Disorder. The double-blind treatment period was 8 week long. Subjects were evaluated weekly during the first 2 weeks of treatment, and then every 2 weeks up to the end of the 8-week treatment period. Subjects who completed the double-blind treatment period entered a 2-week double-blind taper-down/discontinuation period. Subjects were randomized in a 1:1:1:1;1 ratio to 1 of 5 treatment arms (approximately 150 patients per arm). Overall, 575 of 781 randomized (73.6%) subjects completed the study.

The primary efficacy endpoint was the mean change from baseline in the HAM-A total score, after 8 weeks of treatment.

Primary Efficacy Analysis

Comparisons between Lu AA21004 5 and 10 mg and placebo were performed on FAS using a mixed model for repeated measurements (MMRM). Primary efficacy results demonstrate that 8 weeks of treatment with Lu AA21004 (2.5, 5, and 10 mg) QD was not more effective than placebo for the acute treatment of subjects with GAD.

Table 50. HAM-A Total Score Change From Baseline at Week 8 (FAS, LOCF, ANCOVA)

HAM-D-24 Total score	Placebo (N=154)	Duloxetine (N=149)	Lu AA21004		
			2.5mg (N= 154)	5mg (N=148)	10mg (N=154)
LS Mean Change from Baseline	-11.3 (0.60)	-13.9 (0.64)	-12.2 (0.60)	-11.6 (0.61)	-11.7 (0.61)
Difference from Placebo		-2.6 (0.87)	-1.0 (0.84)	-0.3 (0.84)	-0.4 (0.85)
p-value		0.003	0.255	0.719	0.642

Source: Synopsis of Clinical Study Report Lu AA21004_308

APPENDIX C. PRIMARY ENDPOINT EFFICACY BY VISIT

Study 11492A

Table 51. LS Mean (SE) Change From Baseline by Week in MADRS Total Score LS Mean (FAS, LOCF, ANCOVA)

Visit/Week	Placebo	Venlafaxine 225mg	Lu AA21004	
			5mg	10mg
1	-5.0 (0.50)	-4.5 (0.48)	-5.3 (0.49)	-5.9 (0.51)
2	-8.9 (0.70)	-10.3 (0.67)	-10.6 (0.69)	-10.8 (0.71)
3	-10.7 (0.81)	-14.8 (0.79)	-14.9 (0.80)	-14.2 (0.83)
4	-13.1 (0.89)	-17.4 (0.86)	-17.1 (0.88)	-17.7 (0.90)
5	-14.1 (0.97)	-19.7 (0.94)	-19.6 (0.96)	-18.9 (0.990)
6	-14.5 (1.03)	-20.9 (0.99)	-20.4 (1.01)	-20.2 (1.04)

Source: Clinical Study Report Table 42 (pg. 241)

Study 305

Table 52. LS Mean (SE) Change from Baseline by Week in HAM-D-24 Total Score (FAS, MMRM)

Week	Placebo	Lu AA21004		
		1mg	5mg	10mg
1	-3.5 (0.32)	-2.8 (0.32)	-3.2 (0.32)	-3.2 (0.32)
2	-5.7 (0.48)	-7.3 (0.48)	-7.3 (0.49)	-7.4 (0.49)
4	-8.4 (0.63)	-11.1 (0.64)	-11.0 (0.64)	-11.9 (0.65)
6	-10.2 (0.69)	-13.4 (0.70)	-13.7 (0.70)	-15.2 (0.71)
8	-11.3 (0.74)	-14.8 (0.75)	-15.4 (0.74)	-16.2 (0.76)

Source: End of Text Tables and Figures, Table 15.2.1.1.8 (pg. 159-160)

Study 13267A

Table 53. LS Mean (SE) Change from Baseline by Week in MADRS Total Score (FAS, MMRM)

Week	Placebo	Duloxetine 60mg	Lu AA21004	
			15mg	20mg
1	-3.3 (0.31)	-3.2 (0.33)	-2.8 (0.33)	-2.7 (0.32)
2	-6.1 (0.48)	-8.7 (0.50)	-6.4 (0.50)	-7.8 (0.50)
4	-9.0 (0.60)	-13.8 (0.62)	-11.6 (0.62)	-12.6 (0.63)
6	-10.6 (0.71)	-18.4 (0.72)	-15.2 (0.74)	-16.2 (0.74)
8	-11.7 (0.76)	-21.2 (0.77)	-17.2 (0.79)	-18.8 (0.78)

Source: Clinical Study Report Table 30 (pg. 264)

Study 315

Table 54. LS Mean (SE) Change from Baseline by Week in MADRS Total Score (FAS, MMRM)

Week	Placebo	Duloxetine 60mg	Lu AA21004	
			15mg	20mg
1	-4.3 (0.47)	-5.2 (0.47)	-4.7 (0.48)	-4.9 (0.47)
2	-7.1 (0.59)	-9.2 (0.62)	-9.0 (0.62)	-8.9 (0.61)
4	-10.6 (0.72)	-12.1 (0.76)	-12.0 (0.76)	-12.5 (0.75)
6	-11.5 (0.78)	-14.6 (0.82)	-13.1 (0.83)	-14.8 (0.82)
8	-12.8 (0.83)	-16.9 (0.88)	-14.3 (0.89)	-15.6 (0.88)

Source: End of Text Tables and Figures, Table 15.2.1.1.5 (pg. 198-199)

Study 316**Table 55. LS Mean (SE) Change from Baseline by Week in MADRS Total Score (FAS, MMRM)**

Week	Placebo	Lu AA21004	
		10mg	20mg
1	-4.4 (0.50)	-4.8 (0.50)	-4.7 (0.51)
2	-6.9 (0.61)	-8.2 (0.61)	-8.1 (0.63)
4	-9.1 (0.71)	-11.1 (0.72)	-11.8 (0.74)
6	-9.9 (0.75)	-12.9 (0.77)	-13.4 (0.78)
8	-10.8 (0.81)	-13.0 (0.83)	-14.4 (0.85)

Source: End of Text Tables and Figures, Table 15.2.1.1.5 (pg. 178-180)

Study 12541A**Table 56. LS Mean (SE) Change from Baseline by Week in HAM-D-24 Total Score (FAS, LOCF, ANCOVA)**

Week	Placebo	Duloxetine 60mg	Lu AA21004 5mg
1	-3.6 (0.41)	-3.5 (0.41)	-4.0 (0.40)
2	-6.7 (0.53)	-7.9 (0.52)	-7.0 (0.51)
4	-9.0 (0.64)	-12.3 (0.63)	-10.1 (0.62)
6	-10.2 (0.71)	-14.4 (0.70)	-12.3 (0.69)
8	-10.3 (0.76)	-15.8 (0.75)	-13.7 (0.74)

Source: Clinical Study Report Table 34 (pg. 231)

APPENDIX D. INCIDENCE RATES OF TESD BY GENDER SUBGROUPS

1. Incidence of TESD at any visit

a) Males

Table 57. Incidence of TESD by Study in Male subjects without sexual dysfunction at baseline

Study Name	Placebo	Lu AA21004					Duloxetine 60mg
		2.5mg	5mg	10mg	15mg	20mg	
11984A (non US, Can)	3/9 (33.3%)	4/11 (36.4%)	3/11 (27.3%)	7/14 (50.0%)			7/11 (63.6%)
13267A (out. N. America)	4/14 (28.6%)				13/29 (44.8%)	14/23 (60.9%)	8/22 (36.4%)
304 (US)	10/30 (33.3%)	9/26 (34.6%)	8/22 (36.4%)				15/30 (50.0%)
308 [GAD]	10/39 (25.6%)	14/35 (40.0%)	4/36 (11.1%)	10/37 (27.0%)			12/31 (38.7%)
315 (US)	6/27 (22.2%)				6/22 (27.3%)	8/20 (40.0%)	8/14 (57.1%)
316 (US)	9/30 (30.0%)			9/22 (40.9%)		4/18 (22.2%)	
317 (US)	7/18 (38.9%)			0/16 (0.0%)	8/17 (47.1%)		
All studies	49/167 (29.3%)	27/72 (37.5%)	15/69 (21.7%)	26/89 (29.2%)	27/68 (39.7%)	26/61 (42.6%)	50/108 (46.3%)

Source: Sponsor's Response to FDA Request Table 4.2 ([\Cdsesub1\evsprod\NDA204447\0021](#))

b) Females

Table 58. Incidence of TESD by Study in Female subjects without sexual dysfunction at baseline

Study Name	Placebo	Lu AA21004					Duloxetine 60mg
		2.5mg	5mg	10mg	15mg	20mg	
11984A (non US, Can)	4/6 (66.7%)	7/12 (58.3%)	0/3 (0.0%)	3/7 (42.9%)			2/4 (50.0%)
13267A (out. N. America)	8/13 (61.5%)				9/11 (81.8%)	9/12 (75.0%)	10/19 (52.6%)
304 (US)	4/12 (33.3%)	16/23 (69.6%)	10/26 (38.5%)				8/19 (42.1%)
308 [GAD]	10/42 (23.8%)	14/44 (31.8%)	10/38 (26.3%)	9/40 (22.5%)			22/43 (51.2%)
315 (US)	15/31 (48.4%)				10/23 (43.5%)	8/25 (32.0%)	17/33 (51.5%)
316 (US)	5/20 (25.0%)			15/28 (53.6%)		16/30 (53.3%)	
317 (US)	4/18 (22.2%)			14/26 (53.8%)	8/24 (33.3%)		
All studies	50/142 (35.2%)	37/79 (46.8%)	20/67 (29.9%)	41/101 (40.6%)	27/58 (46.6%)	33/67 (49.3%)	59/118 (50.0%)

Source: Sponsor's Response to FDA Request Table 4.3 ([\Cdsesub1\evsprod\NDA204447\0021](#))

2. Incidence of TESD at two consecutive visits

a) Males

Table 59. Incidence of TESD by study at two consecutive visits in Male subjects without sexual dysfunction at baseline.

Study Name	Placebo	Lu AA21004					Duloxetine 60mg
		2.5mg	5mg	10mg	15mg	20mg	
11984A (non US, Can)	1/9 (11.1%)	3/11 (27.3%)	3/11 (27.3%)	6/14 (42.9%)			4/11 (36.4%)
13267A (out N. America)	1/14 (7.1%)				6/29 (20.7%)	9/22 (40.9%)	3/22 (13.6%)
304 (US)	5/27 (18.5%)	6/23 (26.1%)	7/22 (31.8%)				10/27 (37.0%)
308 [GAD]	4/38 (10.5%)	6/33 (18.2%)	1/34 (2.9%)	7/36 (19.4%)			5/25 (20.0%)
315 (US)	3/27 (11.1%)				2/21 (9.5%)	6/20 (30.0%)	4/14 (28.6%)
316 (US)	6/30 (20.0%)			4/22 (18.2%)		2/17 (11.8%)	
317 (US)	2/17 (11.8%)			0/14 (0.0%)	5/17 (29.4%)		
All studies	22/162 (13.6%)	15/67 (22.4%)	11/67 (16.4%)	17/86 (19.8%)	13/67 (19.4%)	17/59 (28.8%)	26/99 (26.3%)

Source: Sponsor's Response to FDA Request Table 4.2 ([\\Cdsub1\evsprod\NDA204447\0021](#))

b) Females

Table 60. Incidence of TESD by study at two consecutive visits in Female subjects without sexual dysfunction at baseline.

Study Name	Placebo	Lu AA21004					Duloxetine 60mg
		2.5mg	5mg	10mg	15mg	20mg	
11984A (non US, Can)	2/6 (33.3%)	6/12 (50.0%)	0/3 (0.0%)	1/7 (14.3%)			2/4 (50.0%)
13267A (out N. America)	6/13 (46.2%)				6/11 (54.5%)	5/12 (41.7%)	5/19 (26.3%)
304 (US)	1/11 (9.1%)	6/20 (30.0%)	6/25 (24.0%)				4/17 (23.5%)
308 [GAD]	5/41 (12.2%)	5/42 (11.9%)	8/37 (21.6%)	5/37 (13.5%)			10/38 (26.3%)
315 (US)	8/30 (26.7%)				9/22 (40.9%)	6/25 (24.0%)	10/31 (32.3%)
316 (US)	3/18 (16.7%)			9/25 (36.0%)		12/30 (40.0%)	
317 (US)	2/16 (12.5%)			7/25 (28.0%)	4/24 (16.7%)		
All studies	27/135 (20.0%)	17/74 (23.0%)	14/65 (21.5%)	22/94 (23.4%)	19/57 (33.3%)	23/67 (34.3%)	31/109 (28.4%)

Source: Sponsor's Response to FDA Request Table 4.3 ([\\Cdsub1\evsprod\NDA204447\0021](#))

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

/s/

GEORGE KORDZAKHIA
06/05/2013

PEILING YANG
06/05/2013
I concur with the review.

HSIEN MING J HUNG
06/05/2013

Statistical Review and Evaluation CARCINOGENICITY STUDIES



IND/NDA Number: NDA 204447
Drug name: Lu AA 21004
Indication(s): MDD in adults
Applicant: Takeda
Documents Reviewed: Electronic submission
Electronically submitted dataset
Dated: 2012-10-02
Review Priority: Normal
Biometrics Division: Division of Biometrics 6
Statistical Reviewer: Matthew Jackson, PhD
Concurring Reviewer: Karl Lin, PhD
Medical Division: Division of Psychiatry Products
Reviewing Pharmacologist: Antonia Dow, PhD
Project Manager: Hiren Patel
Keywords: Animal Mouse Rat Carcinogenicity

Contents

1	Summary of findings	6
1.1	Mouse study	6
1.2	Rat study	6
2	Mouse Study	7
2.1	Experimental design	7
2.2	Sponsor's analysis	7
2.2.1	Survival analysis	7
2.2.2	Tumor analysis	7
2.3	Reviewer's analysis	8
2.3.1	Survival analysis	8
2.3.2	Tumor analysis	10
2.3.3	Analysis of unexamined and autolytic organs	11
3	Rat Study	13
3.1	Experimental design	13
3.2	Sponsor's analysis	13
3.2.1	Survival analysis	13
3.2.2	Tumor analysis	13
3.3	Reviewer's analysis	15
3.3.1	Survival analysis	15
3.3.2	Tumor analysis	19
3.3.3	Analysis of unexamined and autolytic organs	20
4	Assessment of the validity of a negative study	21
4.1	Issues of concern when selecting the dose levels	21
4.2	Assessment of the validity of the mouse study	22
4.3	Assessment of the validity of the rat study	22
A	Tables from mouse study	23
A.1	Survival analysis	23
A.2	Tumor analysis	27
A.3	Unexamined and autolytic organs	46
A.4	Bodyweight changes	51
B	Tables from rat study	52
B.1	Survival analysis	52
B.2	Tumor analysis	57
B.3	Unexamined and autolytic organs	80
B.4	Weight changes	85

List of Tables

2.1	Sponsor’s list of tumor endpoints (mouse study)	8
2.2	Critical <i>p</i> -values used to determine statistical significance	11
3.1	Sponsor’s list of tumor endpoints (rat study)	14
3.2	Summary of notable neoplastic findings in rat study	15
A.1	Numbers of animals alive at certain timepoints (mouse study)	24
A.2	Results of log-rank tests of survival across all groups (mouse study)	25
A.3	Results of pairwise log-rank tests of survival between treated groups and vehicle control (mouse study)	26
A.4	Primary organs in female mouse experiment	28
A.5	Primary organs in male mouse experiment	29
A.6	Secondary organs in female mouse experiment	30
A.7	Secondary organs in male mouse experiment	31
A.8	Customized endpoints analyzed	32
A.9	Tumors reported in female mouse experiment	33
A.10	Tumors reported in male mouse experiment	37
A.11	Combination tumors reported in female mouse experiment	40
A.12	Combination tumors reported in male mouse experiment	42
A.13	Tumors reported significant in male mouse experiment	43
A.14	Combination tumors reported significant in male mouse experiment	44
A.15	Tumors reported secondary organs (mouse study)	45
A.16	Organs reported autolytic in female mouse experiment	47
A.17	Organs reported autolytic in male mouse experiment	48
A.18	Organs reported unexamined in female mouse experiment	49
A.19	Organs reported unexamined in male mouse experiment	50
A.20	Weight changes by group (mice)	51
B.1	Numbers of animals alive at certain timepoints (rat study)	53
B.2	Results of log-rank tests of survival across all groups (rat study)	54
B.3	Results of pairwise log-rank tests of survival between treated groups and vehicle control (rat study)	55
B.4	Results of log-rank test of survival across control groups	56
B.5	Primary organs in female rat experiment	58
B.6	Primary organs in male rat experiment	59
B.7	Secondary organs in female rat experiment	60
B.8	Secondary organs in male rat experiment	61
B.9	Tumors reported in female rat experiment	62
B.10	Tumors reported in male rat experiment	66
B.11	Combination tumors reported in female rat experiment	70
B.12	Combination tumors reported in male rat experiment	72
B.13	Tumors reported significant in female rat experiment	74
B.14	Tumors reported significant in male rat experiment	75
B.15	Combination tumors reported significant in female rat experiment	76

B.16	Combination tumors reported significant in male rat experiment	77
B.17	Tumors reported secondary organs (rat study)	78
B.18	Organs reported autolytic in female rat experiment	81
B.19	Organs reported autolytic in male rat experiment	82
B.20	Organs reported unexamined in female rat experiment	83
B.21	Organs reported unexamined in male rat experiment	84
B.22	Weight changes by group (rats)	85

List of Figures

2.1	Survival curves for female mice	9
2.2	Survival curves for male mice	9
3.1	Survival curves for female rats	16
3.2	Survival curves for male rats	16
3.3	Survival curves for control groups (female rat experiment)	17
3.4	Survival curves for control groups (male rat experiment)	18

Background

In this submission the sponsor included reports of two animal carcinogenicity studies, in mice and rats, to assess the carcinogenic potential of Lu AA 21004 when administered by gavage, once daily at appropriate drug levels for about 104 weeks. Results of this review have been discussed with the reviewing pharmacologist, Antonia Dow, PhD.

In this review, the phrase “dose response relationship” refers to the linear component of the effect of treatment, and not necessarily to a strictly increasing or decreasing mortality or tumor incidence rate as dose increases.

Chapter 1

Summary of findings

1.1 Mouse study

Both the female and male mouse experiments are negative. However, the male mouse experiment only narrowly misses statistical significance for hepatocellular tumors (the test of trend yielded a p -value of 0.0099 and the pairwise comparison between the high dose group and the control group yielded a p -value of 0.0611). In light of the positive finding for this endpoint in the female rat experiment, and the corresponding near positive finding for male rats, this result should not be dismissed out of hand.

Mortality was low, and there are no concerns about the dose level being excessive. Neither weight gain data or mortality data give any indication that the dose levels were close to the MTD in either sex. Crystalline material was found in the bile ducts of the high dose male animals, but no direct toxicity was observed.

Autolysis rates and unexamined organ rates were low, and are no cause for concern.

1.2 Rat study

The female rat experiment is positive for hepatocellular tumors and for rectal polypoid adenomas. The male rat experiment is negative, but only narrowly misses significance for hemangiomas and hemangiosarcomas, for hepatocellular tumors, and for histiocytic sarcomas.

The endpoint of hepatocellular tumors is the cause of considerable concern here, since the statistically significant finding in female rats makes it more reasonable to consider the near-significant findings in male rats and in male mice to be probable false negatives. This in turn leads to the suggestion that there is a genuine tumorigenic effect for this endpoint that acts across sexes and species.

However, it is also important to note that histological evidence (the analysis of which is outside the scope of this review) indicates that the hepatocellular tumors were an indirect consequence of a toxicity effect associated with Lu AA 21004, and which might not be relevant to humans receiving therapeutic doses.

The findings for rectal polypoid adenomas are also worrying. According to Antonia Dow, PhD, the vehicle for this study is known to be associated with intestinal tumors, but such an explanation does not explain the statistically significant dose response that was observed ($p = 0.0034$ for the test of trend and $p = 0.0448$ for the pairwise comparison between the high dose group and the vehicle control).

The female rat experiment was positive, and so necessarily adequate. The male rat experiment appears to have been adequate; mortality rates were low enough that there is no concern about excessive toxicity, and while there is no evidence of a dose related increase in mortality or reduction in weight gain, there is histological evidence of a toxicity effect.

Autolysis rates and unexamined organ rates were low, and are no cause for concern.

Chapter 2

Mouse Study

2.1 Experimental design

The study in mice consisted of two separate experiments; one in female mice, and one in male mice. The animals used in both experiments were Crl:CD1 (ICR) mice. Each experiment consisted of four groups of sixty animals. One group was the control group, who received the vehicle (an aqueous solution containing 15% hydroxypropoyl- β -cyclodextrin). The other three groups, designated the low, mid and high dose groups, received the test article, Lu AA 21004, in doses of 10, 30, and 100 mg/kg/day (female experiment) or 5, 15, and 50 mg/kg per day (male experiment). The test article (and the vehicle) were administered by gavage, for a daily dose volume of 10 mL/kg.

Both experiments were planned for 104 weeks. However, in week 102, after the number of surviving animals in the male control group dropped to twenty five, all surviving male animals were sacrificed. This means the male experiment was terminated two weeks early.

Animals were inspected visually for signs of ill health at least twice a day. Detailed physical exams, including palpation exams, were conducted weekly. In addition, a sequence of more detailed examinations was conducted before and after dosing, with decreasing frequency during the course of the study (daily during the first week of the study, twice a week in weeks 2–4, weekly in weeks 5–13, fortnightly during weeks 14–52, and every four weeks for the duration of the study). After death, whether due to natural causes or termination, all animals underwent a detailed necroscopy.

2.2 Sponsor's analysis

2.2.1 Survival analysis

The sponsor conducted a suite of log-rank tests of survival in each sex: a two tailed test of trend, and a one tailed pairwise test between each treated group and the control group. In addition, Kaplan-Meier curves were plotted.

In neither sex was the trend test found to yield significant results. The low and mid dose male groups were found to have statistically significantly increased mortality rates ($p = 0.049$ in each sex) relative to the control group, but the high dose group was not found to have a higher mortality rate than the control group ($p = 0.861$).

2.2.2 Tumor analysis

For each tumor type for which at least two tumor bearing animals of the same sex were found, and for a large number of combination endpoints (see table 2.1), a life table was constructed, and time-to-detection analyses, following Peto [6], were conducted. Specifically, for each endpoint, a one tailed trend test was conducted, together with a one tailed pairwise test of incidence between each treated group and the control group. Exact methods were used when the total number of tumor bearing animals was less than ten, otherwise asymptotic methods were used.

In female mice, the results for hemangioma and hemangiosarcomas combined yielded a significant trend test ($p = 0.046$), but none of the pairwise tests were significant.

In male mice, the endpoints of hepatocellular tumors and hepatocellular adenomas alone were found to yield significant trend tests ($p = 0.010$ and $p = 0.026$ respectively), but no pairwise comparisons were significant.

The sponsor accepted that in male mice the increased incidence of hepatocellular adenomas (and by extension of all hepatocellular tumors, although the sponsor noted that no increase was seen in the incidence of hepatocellular carcinomas) was dose related, but argued that this effect was due to a toxicity effect (build up of crystalline material) in the hepatobiliary system, and that dose levels below 15mg/kg/day (the mid dose level) for males and below 100 mg/kg/day (the high dose level for females) had been demonstrated safe.

The sponsor’s report does not discuss the ramifications of the hemangioma/hemangiosarcoma finding in female mice. Presumably this is because the results do not remain significant after making an adjustment for multiple testing, but this does not appear to have been stated explicitly. However, the results section of the report does say: “No statistically significant results were found for females.”

Table 2.1: Sponsor’s list of tumor endpoints (mouse study)

Female mice
Lungs/bronchi - Benign bronchioloalveolar adenoma
Lungs/bronchi - Malignant bronchioloalveolar adenocarcinoma
Lungs/bronchi - Benign bronchioloalveolar adenoma and malignant bronchioloalveolar adenocarcinoma combined
Liver - Benign hepatocellular adenoma
Liver - Benign haemangioma and malignant haemangiosarcoma combined
Thymus - Benign thymoma (lymphoid)
Thymus - Benign thymoma (lymphoid) and malignant thymoma combined
Pituitary (pars distalis) - Benign adenoma
Stomach - Benign squamous cell papilloma
Mammary areas - Malignant mammary adenocarcinoma
Skin - Malignant fibrosarcoma
Harderian glands - Benign adenoma
Ovaries - Benign luteoma
Ovaries - Benign granulosa cell tumour
Uterus - Benign haemangioma
Uterus - Benign haemangioma and malignant haemangiosarcoma combined
Uterus - Malignant endometrial stromal cell sarcoma
Uterus - Benign endometrial polyp
Uterus - Benign endometrial polyp and malignant endometrial stromal cell sarcoma combined
Uterus - Benign leiomyoma
Uterine cervix - Benign endometrial polyp
Haematopoietic tumour - Malignant lymphoma
Haematopoietic tumour - Malignant histiocytic sarcoma
Haematopoietic tumour - Malignant myeloid cell leukaemia
All tissues - Benign haemangioma
All tissues - Malignant haemangiosarcoma
All tissues - Benign haemangioma and malignant haemangiosarcoma combined
Male mice
Lungs/bronchi - Benign bronchioloalveolar adenoma
Lungs/bronchi - Malignant bronchioloalveolar adenocarcinoma
Lungs/bronchi - Benign bronchioloalveolar adenoma and malignant bronchioloalveolar adenocarcinoma combined
Liver - Benign hepatocellular adenoma
Liver - Malignant hepatocellular carcinoma
Liver - Benign hepatocellular adenoma and malignant hepatocellular carcinoma combined
Liver - Benign haemangioma
Liver - Benign haemangioma and malignant haemangiosarcoma combined
Adrenals - Benign subcapsular cell adenoma
Skin - Malignant fibrosarcoma
Harderian glands - Benign adenoma
Testes - Benign interstitial (Leydig) cell adenoma
Haematopoietic tumour - Malignant lymphoma
Haematopoietic tumour - Malignant histiocytic sarcoma
Haematopoietic tumour - Malignant myeloid cell leukaemia
All tissues - Benign haemangioma
All tissues - Malignant haemangiosarcoma
All tissues - Benign haemangioma and malignant haemangiosarcoma combined

2.3 Reviewer’s analysis

2.3.1 Survival analysis

The Kaplan-Meier survival plots are shown as figures 2.1 and 2.2. The numbers and proportions of animals surviving to various times are presented in table A.1. The results of log-rank tests of heterogeneity of survival and of dose response across the groups are presented in table A.2, and the results of log-rank survival tests comparing the treated groups with the vehicle control group are presented in table A.3.

Figure 2.1: Survival curves for female mice

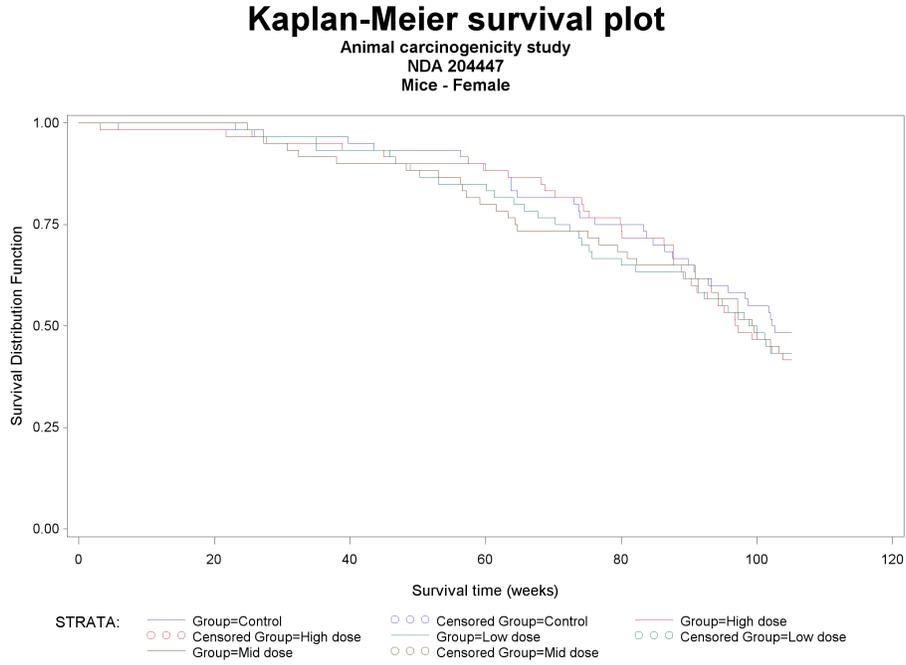
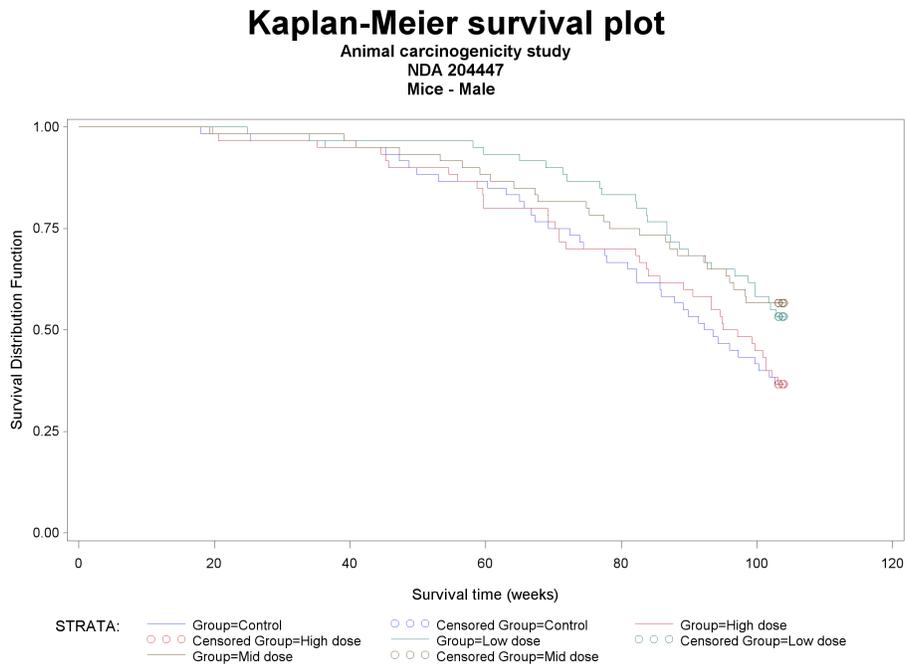


Figure 2.2: Survival curves for male mice



Commentary Neither statistical tests nor inspection of the Kaplan-Meier plots reveal any evidence of a dose related impact on survival in either the female or male mice.

The only indication of any dose related mortality is in the pairwise comparisons of survival of the low dose and mid dose male animals with the control animals; the p -value for the log-rank tests are, in these cases, 0.0389 (low versus control) and 0.0326 (mid versus control). However, in the absence of a significant increase in mortality in the high dose group, there seems no reason to conclude that these results are anything but statistical noise.

2.3.2 Tumor analysis

Endpoints

Analyses have been conducted using the sponsor's submitted dataset, and the sponsor's chosen nomenclature. In this dataset, organs or tissue types are described as being either tumorous, examined but found unusable due to autolysis, or unexamined. An organ that has been examined but was not found to be tumorous is not mentioned in the dataset.

From these data, we can infer the numbers of animals for which each organ or tissue type was examined, but only in those cases where at least one anomalous finding (i.e., a tumor was found, or a sample that was planned to be analyzed could not be, either because no sample was taken or because the sample was unusable due to autolysis) was reported. Organs which can thus be deduced to have been successfully analyzed in the majority of animals are, for the purposes of this review, considered *primary*. The lists of primary organs in the experiments on female and male mice respectively are presented in tables A.4 and A.5.

Organ or tissue types which were examined in only a few animals are considered *secondary*.

Secondary organs in the male and female mouse experiments are presented in tables A.6 and A.7 respectively.

Each tumor type found in a primary organ of at least one animal is considered a primary endpoint. In addition, in consultation with Antonia Dow, PhD, a list of combination endpoints has been drawn up. This list is presented in table A.8.

Statistical procedure

The tumor data were analyzed for dose response relationships and pairwise comparisons of tumor incidence in each of the treated groups versus the vehicle control group. Both the dose response relationship tests and pairwise comparisons were performed using the poly- k method described in the paper of Bailer and Portier[1] and developed in the paper of Bieler and Williams[2]. In this method, given a tumor type T , an animal h that lives the full study period (w_m) or dies before the terminal sacrifice with at least one tumor of type T gets a score of $s_h = 1$. An animal that dies at week w_h before the end of the study without such a tumor gets a score of

$$s_h = \left(\frac{w_h}{w_m} \right)^k < 1.$$

The adjusted group size is defined as $\sum_h s_h$. As an interpretation, an animal with score $s_h = 1$ can be considered as a whole animal while an animal with score $s_h < 1$ can be considered as a partial animal. The adjusted group size $\sum s_h$ is equal to N (the original group size) if all animals live up to the end of the study or if each animal develops at least one tumor of type T , otherwise the adjusted group size is less than N . These adjusted group sizes are then used for the dose response relationship (or the pairwise) tests using the Cochran-Armitage test. The test is repeated for each tumor type T .

One critical point to consider in the application of the poly- k test is the choice of the appropriate value of k , which depends on the relationship between tumor onset time and increased dose. For long term 104 week standard rat and mouse studies, a value of $k = 3$ is suggested in the literature, and so has been used in this review. For the calculation of p -values, the exact permutation method was used.

When testing so many endpoints, there is a danger of inflation of type I error. To control against this, the current draft guidance recommends making adjustments in the significance thresholds. In order to best manage the trade-off between control of type I and type II error, and to allow for the relative rarity of some tumors, it is recommended that a distinction be drawn between rare tumors (with a background incidence rate below 1%) and common tumors. For a study with two two year studies (one of mice, and one of rats), the currently proposed significance thresholds are given in table 2.2. It is expected that these adjustments will suffice to keep the submission-wide false positive rate at a nominal level of approximately 10%.

It should be noted that the FDA guidance for multiple testing for dose response relationship is based on a publication by Lin and Rahman [5]. In this work the authors investigated the use of this rule for Peto analysis. However, in a later work Rahman and Lin [7] showed that this rule for multiple testing for dose response relationship is also suitable for poly- k tests.

Since this is a study involving two species, it follows that for the comparisons of Lu AA 21004 with vehicle control, we use the thresholds for significance presented in table 2.2.

Table 2.2: Critical p -values used to determine statistical significance

Type of test	Rare tumor	Common tumor
Trend	0.025	0.005
Pairwise test between placebo and high dose	0.10	0.05

The results of the statistical analyses of tumor incidence in primary endpoints are presented in tables A.9 (female mice) and A.10 (male mice). The results of analyses of customized endpoints (see table A.8) are presented in tables A.11 and A.12.

Noteworthy results

None of the statistical tests conducted as part of the female mouse experiment yielded a p -value below 0.05. Individual tumor types in male mice for which tests yielding p -values below 0.05 were conducted are presented in table A.13, which is excerpted from table A.10. Combination tumor types for which tests yielding p -values below 0.05 were conducted are presented in table A.14, which is excerpted from table A.12.

Incidence rates for tumors found in secondary organs have not been analyzed statistically. Count data for such tumors are presented in table A.15.

Hepatocellular tumors in male mice The combined endpoint of hepatocellular adenomas and carcinomas was clearly a common tumor type; fourteen male control animals developed such tumors (ten developed hepatocellular adenomas alone, three developed hepatocellular carcinomas alone, and one animal (number 7) was reported as having developed both a hepatocellular adenoma and carcinoma). Neither the test of trend nor the pairwise test between the high dose and control groups yields results that are significant for common tumors; the p -value for the trend test is 0.0099 (above the threshold value of 0.005), and the pairwise test yields a p -value of 0.0611 (above the threshold value of 0.05). Considered in isolation, this result should therefore be found to be negative.

However, given that the results are nearly significant, and that tests of the same tumor types achieve significance in the female rat experiment and near significance in the male rat experiment, we should nonetheless be wary of this possible effect, the failure to achieve statistical significance notwithstanding.

2.3.3 Analysis of unexamined and autolytic organs

Unexamined animals

No animals have been reported as completely unexamined.

Organs reported autolytic

The numbers of organs found in female mice to be autolytic to the extent that analysis of collected tissue was not possible are presented in table A.16. The numbers of such organs found in male mice are presented in table A.17.

Among both male and female mice, the autolysis rates are generally low, and are no cause for concern.

Organs reported as unexamined

The numbers of animals with organs reported as being unexamined are presented in tables A.18 and A.19. Among both male and female mice, the rates at which organs have been reported are generally low, and are no cause for concern.

Chapter 3

Rat Study

3.1 Experimental design

The study in rats consisted of two separate experiments; one in female rats, and one in male rats. The animals used in both experiments were Wistar Han (CrI: WI (Han)) rats. Each experiment consisted of five groups of fifty five animals. One group was the vehicle control group, who received the vehicle (an aqueous solution containing 10% hydroxypropoyl- β -cyclodextrin with 4.4% glucose monohydrate), and one was a negative control group, who were dosed with water. The other three groups, designated the low, mid and high dose groups, received the test article, Lu AA 21004, in twice daily doses of 5, 15, and 40 mg/kg (female experiment) or 2, 7, and 20 mg/kg (male experiment). The test article (and the vehicle) were administered by gavage, for a dose volume of 5 mL/kg (twice daily).

Both experiments were conducted over a period of 104 weeks.

Animals were inspected visually for signs of ill health at least twice a day. Detailed physical exams, including palpation exams, were conducted weekly. In addition, a sequence of more detailed examinations was conducted before and after each dosing, with decreasing frequency during the course of the study (daily during the first week of the study, twice a week in weeks 2–4, weekly in weeks 5–13, fortnightly during weeks 14–52, and every four weeks for the duration of the study). After death, whether due to natural causes or termination, all animals underwent a detailed necroscopy.

3.2 Sponsor's analysis

3.2.1 Survival analysis

The sponsor conducted a suite of log-rank tests of survival in each sex: a two tailed test of trend, and a one tailed pairwise test between each treated group and the control group. In addition, Kaplan-Meier curves were plotted.

A significant trend of decreased survival with increased dose was noted in the female experiment ($p = 0.004$). None of the pairwise tests yielded significant results.

None of the tests in the male rat experiment were found to yield significant results.

3.2.2 Tumor analysis

For each tumor type for which at least two tumor bearing animals of the same sex were reported, and for a large number of combination endpoints (see table 3.1), a life table was constructed, and time-to-detection analyses, following Peto [6], were conducted. Specifically, for each endpoint, a one tailed trend test was conducted, together with a one tailed pairwise test of incidence between each treated group and the control group. Exact methods were used when the total number of tumor bearing animals found was less than ten, otherwise asymptotic methods were used.

Table 3.1: Sponsor's list of tumor endpoints (rat study)

Female rats
Liver - Benign hepatocellular adenoma
Liver - Benign hepatocellular adenoma and malignant hepatocellular carcinoma combined
Heart - Benign endocardial schwannoma
Mesenteric lymph node - Benign haemangioma
Mesenteric lymph node - Benign haemangioma and malignant haemangiosarcoma combined
Thymus - Benign thymoma
Thyroids - Benign C-cell adenoma
Thyroids - Benign C-cell adenoma and malignant C-cell carcinoma combined
Thyroids - Benign follicular cell adenoma
Thyroids - Malignant follicular cell carcinoma
Thyroids - Benign follicular cell adenoma and malignant follicular cell carcinoma combined
Adrenals - Benign cortical adenoma
Adrenals - Benign cortical adenoma and malignant cortical carcinoma combined
Adrenals - Benign pheochromocytoma and malignant pheochromocytoma combined
Pituitary (pars distalis) - Benign adenoma
Rectum - Benign polypoid adenoma
Mammary areas - Benign fibroadenoma
Mammary areas - Malignant adenocarcinoma
Mammary areas - Benign adenoma, benign fibroadenoma and malignant adenocarcinoma combined
Uterus - Benign endometrial polyp
Uterus - Benign endometrial adenoma
Uterus - Malignant endometrial adenocarcinoma
Uterus - Benign endometrial adenoma and malignant endometrial adenocarcinoma combined
Uterine cervix - Benign endometrial stromal polyp
Uterine cervix - Benign schwannoma and malignant schwannoma combined
Male rats
Liver - Benign hepatocellular adenoma
Liver - Malignant hepatocellular carcinoma
Liver - Benign hepatocellular adenoma and malignant hepatocellular carcinoma combined
Spleen - Malignant haemangiosarcoma
Spleen - Benign haemangioma and malignant haemangiosarcoma combined
Pancreas - Benign acinar cell adenoma
Pancreas - Benign Islet cell adenoma
Pancreas - Malignant Islet cell carcinoma
Pancreas - Benign Islet cell adenoma and malignant Islet cell carcinoma combined
Mesenteric lymph node - Benign haemangioma
Thymus - Benign thymoma
Thymus - Benign thymoma and malignant thymoma combined
Thyroids - Benign C-cell adenoma
Thyroids - Malignant C-cell carcinoma
Thyroids - Benign C-cell adenoma and malignant C-cell carcinoma combined
Thyroids - Benign follicular cell adenoma
Thyroids - Malignant follicular cell carcinoma
Thyroids - Benign follicular cell adenoma and malignant follicular cell carcinoma combined
Parathyroids - Benign chief cell adenoma
Adrenals - Benign cortical adenoma
Adrenals - Benign pheochromocytoma
Adrenals - Benign pheochromocytoma and malignant pheochromocytoma combined
Pituitary (pars distalis) - Benign adenoma
Mammary areas - Benign fibroma
Mammary areas - Benign fibroma and malignant fibrosarcoma combined
Skin - Benign keratoacanthoma
Skin - Benign fibroma
Skin - Benign basal cell tumour
Brain - Benign granular cell tumour
Brain - Benign granular cell tumour and malignant granular cell tumour combined
Epididymides - Malignant mesothelioma
Testes - Benign interstitial (Leydig) cell adenoma
Testes - Malignant mesothelioma
Preputial glands - Benign squamous cell papilloma
Preputial glands - Malignant squamous cell carcinoma
Preputial glands - Benign squamous cell papilloma and malignant squamous cell carcinoma combined
Haematopoietic tumour - Malignant lymphoma
Haematopoietic tumour - Malignant histiocytic sarcoma
All tissues - Benign haemangioma
All tissues - Malignant haemangiosarcoma
All tissues - Benign haemangioma and malignant haemangiosarcoma combined

The sponsor notes a number of significant (defined by an α -level of 0.05) results. They are summarized in table 3.2:

Table 3.2: Summary of notable neoplastic findings in rat study

Organ	Endpoint or Tumor	Sex	<i>p</i> -values	
			Trend test	Pairwise
Liver	Hepatocellular adenoma	Female	0.008	0.694
		Male	0.025	0.181
	All hepatocellular tumors	Female	0.010	0.054
		Male	0.013	0.019
Rectum	Polypoid adenoma	Female	0.002	0.032
Histiocytic sarcoma		Male	0.016	0.096
Mesenteric lymph node	Hemangioma	Male	0.003	0.012
		Male	0.003	0.012
		Male	0.015	0.023
Thyroid	Follicular adenoma	Male	0.050	0.255

The sponsor considers the increase in incidence in hepatocellular adenomas to be clinically meaningful, but suggests (or rather claims that the SAG panel suggests) that rather than being a consequence of the text article’s genotoxicity, the liver tumors are an indirect consequence of the build up of crystals in the hepatobiliary system. They argue that this effect is only observed at the mid and high dose levels in female rats, and at the high dose level in male rats, and therefore that at lower levels there is unlikely to be any corresponding effect.

The increase in hemangiomas of the mesenteric lymph node is acknowledged, but claimed to be specific to rats, “due to exaggerated susceptibility to angioproliferative stimuli in this species”. The increase in histiocytic sarcomas is dismissed as being random noise, and the findings for the thyroids are considered “incidental”. The increase in incidence of rectal tumors is attributed to the vehicle, although no explanation is given for why a significant dose response would be observed when all animals (except the water control group) received the same dose of this vehicle.

3.3 Reviewer’s analysis

3.3.1 Survival analysis

The Kaplan-Meier survival plots are shown as figures 3.1 and 3.2. The numbers and proportions of animals surviving to various times are presented in table B.1. The results of log-rank tests of heterogeneity of survival and of dose response across the groups are presented in table B.2, and the results of log-rank survival tests comparing the treated groups with the vehicle control group are presented in table B.3.

Commentary Among female rats, there is very strong evidence of a dose-related increase in mortality ($p = 0.0003$), although no one treated group has experienced significantly diminished survival compared to the controls. Among male rats, there is no indication of a relationship between dose and survival.

Comparison of control groups Kaplan-Meier plots of the control groups are shown as figures 3.3 and 3.4. The results of log-rank tests of survival between the control groups are presented in table B.4.

There is no statistically significant difference in survival between the water and vehicle control groups in either sex.

Figure 3.1

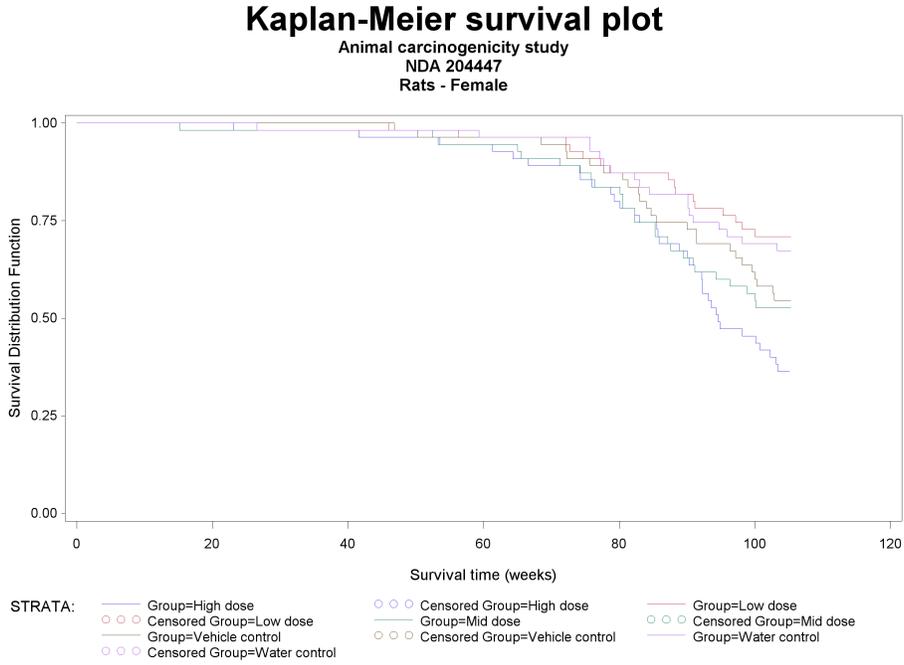


Figure 3.2

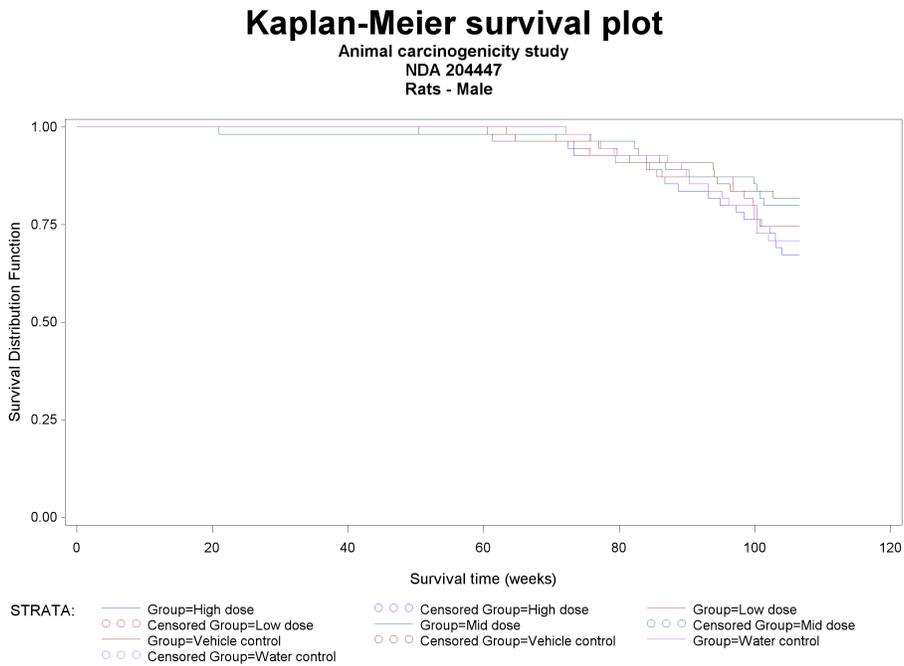


Figure 3.3: Survival curves for control groups (female rat experiment)

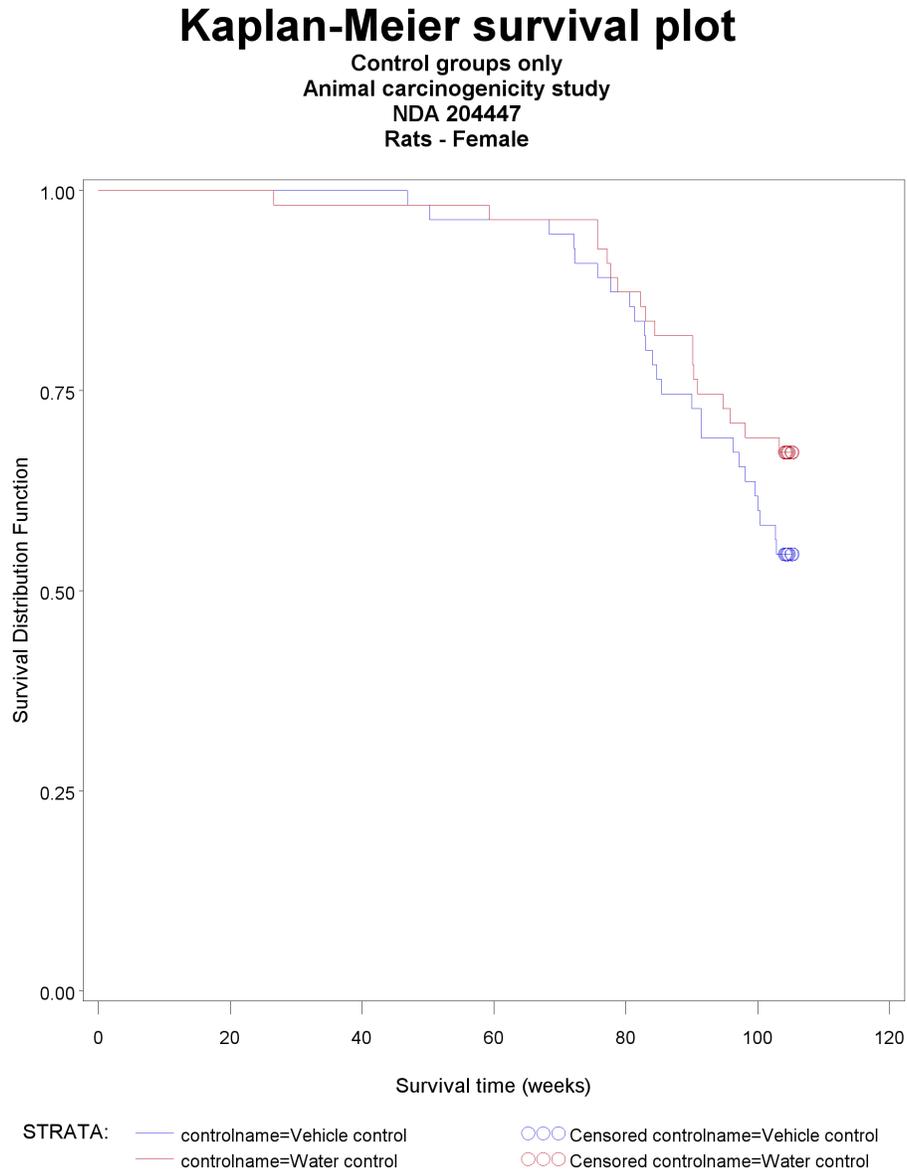
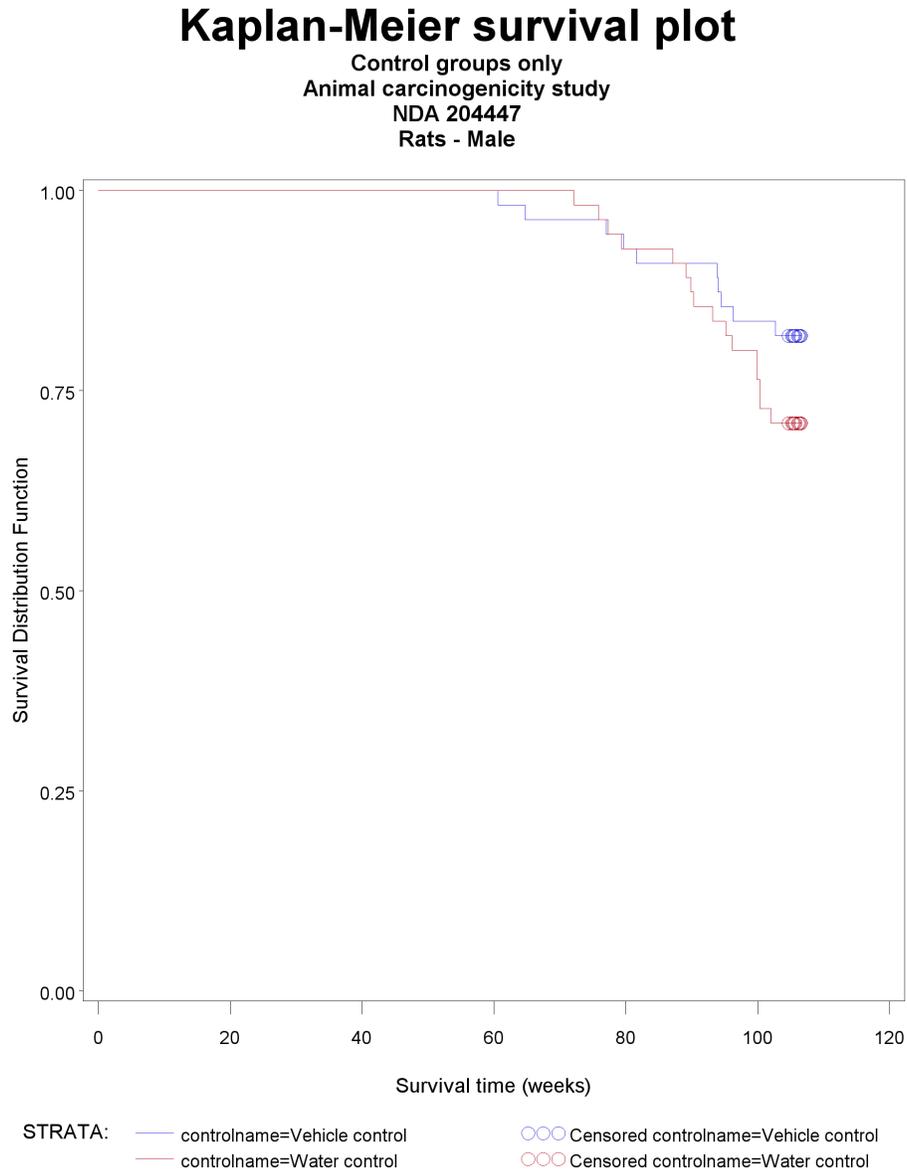


Figure 3.4: Survival curves for control groups (male rat experiment)



3.3.2 Tumor analysis

Endpoints

As in the mouse study, organs have been classed as either primary or secondary (see Section 2.3.2). The lists of organs adduced to be primary are presented in tables B.5 and B.6. Secondary organs in the female and male rat experiments are presented in tables B.7 and B.8 respectively.

The same customized endpoints have been analyzed as were considered in the mouse study (see table A.8).

Statistical procedure

The same statistical procedures are used to assess tumor incidence in rats as were used in mice (see Section 2.3.2). Note that the critical p -values used to determine significance are presented in table 2.2.

The results of the statistical analyses of tumor incidence in primary endpoints are presented in tables B.9 (female rats) and B.10 (male rats). The results of analyses of customized endpoints (see table A.8) are presented in tables B.11 and B.12.

Noteworthy results

Individual tumor types in female rats for which tests yielding p -values below 0.05 were conducted are presented in table B.13, which is excerpted from table B.9. Combination tumor types for which tests yielding p -values below 0.05 were conducted are presented in table B.15, which is excerpted from table B.11. Individual tumor types in male rats for which tests yielding p -values below 0.05 were conducted are presented in table B.14, which is excerpted from table B.10. Combination tumor types for which tests yielding p -values below 0.05 were conducted are presented in table B.16, which is excerpted from table B.12.

Incidence rates for tumors found in secondary organs have not been analyzed statistically. Count data for such tumors are presented in table B.17.

Hepatocellular tumors In the case of both the male and female rats, no hepatocellular tumors were reported in the control groups. In addition, only one animal of each sex in the water control groups developed such tumors. It is thus reasonable to consider these to be rare tumors. Accordingly, the results for the female rats are statistically significant ($p = 0.0103$ for the test of trend and 0.0474 for the pairwise comparison between the control and high dose group), and the results for the male rats are almost significant ($p = 0.0277$ for the test of trend and $p = 0.0252$ for the pairwise test).

Taken together, across the two sexes, and especially when viewed in conjunction with the near significant result for male mice, this provides strong evidence of a tumorigenic effect for hepatocellular tumors associated with Lu AA 21004.

However, there is doubt about the applicability of this effect to humans; after discussion with Antonia Dow, PhD, it seems reasonable to think that this effect is being caused only indirectly by Lu AA 21004, via some toxicological mechanism that is unlikely to be relevant to humans at normal therapeutic doses. It is beyond the scope of this review to assess such causal mechanisms. It is important to note, however, that even if this suggestion is correct, it does not allow us to conclude with any certainty that Lu AA 21004 does not directly cause hepatocellular tumors; the study should rather be viewed as inconclusive for these endpoints, as the toxicological effect would be masking any true carcinogenic effect.

Polypoid adenomas of the rectum Rectal polypoid adenomas are rare in rats, but even if they were not, so the fact that five such tumors were observed in the female rats (one in the mid dose group and four in the high dose group) would still mean that the resulting tests met the threshold for statistical significance (the trend test yielded a p -value of 0.0034, and the pairwise comparison yielded a p -value of $p = 0.0448$). This is therefore a positive finding.

In male rats, only one such tumor was found, in a mid dose male animal.

Hemangiomas and hemangiosarcomas When hemangiomas of the mesenteric lymph node are considered in isolation, the results for male animals are statistically significant ($p = 0.0040$ for the test of trend and $p = 0.0163$ for the comparison test). However, after discussion with Antonia Dow, PhD, it was decided that hemangiomas are only meaningful when combined across sites, and combined with hemangiosarcomas. When this combination endpoint is considered, the result narrowly fails to attain statistical significance; the test of trend yields a p -value of 0.0180, above the threshold of 0.0050, although the pairwise test $p = 0.0313$ remains significant. The results should therefore be considered a negative finding, but only narrowly so.

Histiocytic sarcomas Histiocytic sarcomas in male rats appear to be rare. Accordingly, the test of trend $p = 0.0173$ is statistically significant. However, the pairwise comparison between the high dose group and the vehicle control group is not quite statistically significant $p = 0.1137$ (just above the threshold value of 0.100). Nominally therefore, this is a negative finding. However, this result is close enough to statistical significance to at least warrant so some additional consideration before being dismissed.

3.3.3 Analysis of unexamined and autolytic organs

Unexamined animals

No animals have been reported as completely unexamined.

Organs reported autolytic

The numbers of organs found in female rats to be autolytic to the extent that analysis of collected tissue was not possible are presented in table B.18. The numbers of such organs found in male rats are presented in table B.19.

Among both male and female rats, the autolysis rates are generally low, and are no cause for concern.

Organs reported as unexamined

The numbers of animals with organs reported as being unexamined are presented in tables B.20 and B.21. Among both male and female rats, the rates at which organs have been reported are generally low, and are no cause for concern.

Chapter 4

Assessment of the validity of a negative study

4.1 Issues of concern when selecting the dose levels

The selection of an appropriate dose level for the high dose group is made difficult by the need to satisfy two competing imperatives: on the one hand, if the dose level is insufficiently high, then genuine carcinogenicity effects may not be apparent, but on the other hand, if the dose level is too high, then there is a risk of non-carcinogenic toxic effects killing the animals before they have a chance to demonstrate a carcinogenicity effect.

Haseman [4] suggested that a satisfactory balance between these two imperatives has been found when the following two conditions are both satisfied:

1. Were enough animals exposed, for a sustained amount of time, to the risk of late developing tumors?
2. Were dose levels high enough to pose a reasonable tumor challenge to the animals?

There is no consensus among experts regarding the number of animals and length of time at risk, although most carcinogenicity studies are designed to run for two years with fifty animals per treatment group. The following are some rules of thumb regarding these two issues as suggested by experts in this field:

Haseman [4] has done an investigation on the first issue. He gathered data from 21 studies using Fischer 344 rats and B6C3F1 mice conducted at the National Toxicology Program (NTP). It was found that, on the average, approximately 50% of the animals in the high dose group survived the two year study period. Also, in a personal communication with Dr. Karl Lin of Division of Biometrics-6, Haseman suggested that, as a rule of thumb, a 50% survival of 50 initial animals or 20 to 30 animals still alive in the high dose group, between weeks 80–90, would be considered as a sufficient number and adequate exposure. In addition Chu, Cueto and Ward [3], suggested that “to be considered adequate, an experiment that has not shown a chemical to be carcinogenic should have groups of animals with greater than 50% survival at one year.”

It appears, from these three sources that the proportions of survival at 52 weeks, 80–90 weeks, and two years are of interest in determining the adequacy of exposure and number of animals at risk.

Regarding the question of adequate dose levels, it is generally accepted that the high dose should be close to the maximum tolerated dose (MTD). In the paper of Chu, Cueto and Ward [3], the following criteria are mentioned for dose adequacy. A high dose is considered as close to MTD if any of the criteria is met:

1. A dose is considered adequate if there is a detectable loss in weight gain of up to 10% in a dosed group relative to the controls.

2. The administered dose is also considered an MTD if dosed animals exhibit clinical signs or severe histopathologic toxic effects attributed to the chemical.
3. In addition, doses are considered adequate if the dosed animals show a slight increased mortality compared to the controls.

4.2 Assessment of the validity of the mouse study

Both the female and male mouse experiments were negative, so it is appropriate to consider whether the dose levels were appropriate.

In both sexes, survival rates were good, with at least 32% of animals in each group surviving to the 90th week, and at least 22 surviving to termination. We can conclude that the dose levels were not excessive.

However, the evidence that the dose levels were suitably high is not so strong. In the absence of any indication of a mortality effect, it is appropriate to consider factors such as weight gain and histologically detected signs of toxicity. As shown in table A.20, there is no sign of any dose related reduction in weight gain; on the contrary, if anything, the dosed groups gained more weight than the vehicle control group.

So, the decision of whether the dose levels can be considered adequate hinges on the histological findings. However, assessment of these findings is outside the scope of this review, so no conclusion can be drawn here regarding the suitability of the dose levels.

4.3 Assessment of the validity of the rat study

The female rat experiment is positive for both polypoid adenomas of the rectum and hepatocellular tumors. It is therefore not necessary to worry about issues of statistical power in this case.

Despite some near-significant results (hepatocellular tumors, hemangiomas and hemangiosarcomas, and histiocytic sarcomas), the male rat experiment was negative, and so it is appropriate to consider the appropriateness of the dose regime.

Survival was generally very good (see table B.1), with at least 67% of each dose group surviving to termination. We can therefore conclude that the dose levels were not excessive. However, the evidence that the dose levels were adequate is somewhat weaker. There is no evidence of either a dose related increase in mortality (table B.2) or a dose related reduction in weight gain (table B.22). However, Antonia Dow, PhD has reported toxicological findings in the male rats that suggest that the dose levels were high enough for the animals to experience some dose related toxicity. It follows that the dose levels were probably adequate in this case.

Appendix A

Tables from mouse study

A.1 Survival analysis

APPEARS THIS WAY ON
ORIGINAL

Table A.1

**Survival rates at key times
NDA 204447
Animal carcinogenicity study
Mice**

<i>Species and Sex</i>	<i>Dose Group</i>	<i>Dose (mg per kg)</i>	<i>Number at start</i>	<i>Number alive after 52 weeks</i>	<i>Percentage alive after 52 weeks</i>	<i>Number alive after 78 weeks</i>	<i>Percentage alive after 78 weeks</i>	<i>Number alive after 90 weeks</i>	<i>Percentage alive after 90 weeks</i>	<i>Number sacrificed</i>	<i>Percentage sacrificed</i>	<i>Maximum survival (weeks)</i>
Mice - Female	Control	0	60	56	93%	45	75%	39	65%	29	48%	105
	Low dose	10	60	52	87%	40	67%	37	62%	26	43%	105
	Mid dose	30	60	53	88%	42	70%	39	65%	25	42%	105
	High dose	100	60	54	90%	46	77%	37	62%	26	43%	105
Mice - Male	Control	0	60	53	88%	40	67%	32	53%	22	37%	104
	Low dose	5	60	58	97%	50	83%	41	68%	32	53%	104
	Mid dose	15	60	56	93%	46	77%	41	68%	34	57%	104
	High dose	50	60	54	90%	42	70%	36	60%	22	37%	104

Table A.2

Log-rank tests of survival
NDA 204447
Animal carcinogenicity study
Mice

<i>Sex</i>	<i>Test of homogeneity: chi squared statistic</i>	<i>Test of homogeneity: degrees of freedom</i>	<i>Number of groups</i>	<i>Test of homogeneity: p-value</i>	<i>Test of trend (two tailed): p-value</i>	<i>Test of trend (one tailed): p-value</i>
Female	0.6947	3	4	0.8745	0.7983	0.3991
Male	8.2731	3	4	0.0407	0.3197	0.1599

Table A.3

**Pairwise comparisons (log-rank) of survival between treated groups and controls
NDA 204447
Animal carcinogenicity study
Mice**

<i>Species and Sex</i>	<i>Quantity</i>	<i>Low dose</i>	<i>Mid dose</i>	<i>High dose</i>
Mice - Female	Chi squared test statistic	0.4489	0.5906	0.2807
	p-value of comparison with control	0.5029	0.4422	0.5963
Mice - Male	Chi squared test statistic	4.2669	4.5692	0.0194
	p-value of comparison with control	0.0389	0.0326	0.8892

A.2 Tumor analysis

APPEARS THIS WAY ON ORIGINAL

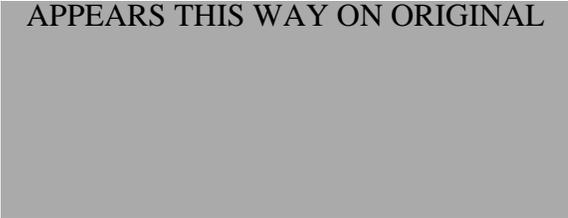


Table A.4

**Primary organs in study of female mice
NDA 204447
Animal carcinogenicity study**

<i>Organ or tissue name</i>
ADRENALS
BRAIN
CAECUM
GALL BLADDER
H-POIETIC TUMOUR
HARDERIAN GLANDS
LIVER
LN MESENTERIC
LUNGS + BRONCHI
MAMMARY
OVARIES
PANCREAS
PITUITARY
SKIN
SPLEEN
STOMACH
THYMUS
UTERINE CERVIX
UTERUS

Table A.5

**Primary organs in study of male mice
NDA 204447
Animal carcinogenicity study**

<u>Organ or tissue name</u>
ADRENALS
FEMUR INC. JOINT
H-POIETIC TUMOUR
HARDERIAN GLANDS
JEJUNUM
KIDNEYS
LIVER
LN MANDIBULAR
LUNGS + BRONCHI
PANCREAS
PITUITARY
SEMINAL VESICLES
SKELETAL MUSCLE
SKIN
SPLEEN
STOMACH
TESTES

Table A.6

**Secondary organs in study of female mice
NDA 204447
Animal carcinogenicity study**

Organ
or
tissue
name
BONE

Table A.7

**Secondary organs in study of male mice
NDA 204447
Animal carcinogenicity study**

<i>Organ or tissue name</i>
ADIPOSE TISSUE
HEAD

Table A.8

**Customized and combination endpoints analyzed
NDA 204447
Animal carcinogenicity study**

<i>Composite endpoint</i>
Bronchioloalveolar tumors
C-cell tumors
Cortical cell tumors
Cytadenomas, granulosa cell tumors and luteomas
Endometrial polyps and stromal cell polyps, and stromal cell sarcomas
Fibromas, fibrosarcomas, and sarcoms (NOS) of the skin
Follicular cell tumors
Glial cell tumors
Granular cell tumors
Haemangiomas and Haemangiosarcomas
Hepatocellular tumors
Islet cell tumors
Mammary adenocarcinomas and adenoacanthomas
Mammary adenocarcinomas and adenomas
Mammary adenomas and fibroadenomas
Mammary fibroadenomas and fibrosarcomas
Mesotheliomas
Pars distalis and pars intermedia tumors
Pheochromocytomas
Prostate tumors
Squamous cell tumors of the skin
Testicular tumors
Thymomas
Tubular cell tumors
Uterine and cervical adenomas, adenocarcinomas, and adenosquamous carcinomas

Table A.9

Table of reported tumors in Mouse Study
NDA 204447
Animal carcinogenicity study
Female mice

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Control Size = 60</i>	<i>Low dose Size = 60</i>	<i>Mid dose Size = 60</i>	<i>High dose Size = 60</i>
ADRENALS	PHAEOCHROMOCYTOMA	P-value of test of trend or comparison	.7410	.4881		
		Number of animals reported with tumor	0	1	0	0
BRAIN	MENINGEAL SARCOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
CAECUM	ADENOMA	P-value of test of trend or comparison	.2500			.4881
		Number of animals reported with tumor	0	0	0	1
GALL BLADDER	SARCOMA NOS	P-value of test of trend or comparison	.7312	.4756		
		Number of animals reported with tumor	0	1	0	0
H-POIETIC TUMOUR	HISTIOCYTIC SARCOMA	P-value of test of trend or comparison	.5270	.9826	.9839	.8518
		Number of animals reported with tumor	5	1	1	3
	MALIGNANT LYMPHOMA	P-value of test of trend or comparison	.2716	.4748	.1933	.3006
		Number of animals reported with tumor	13	14	19	17
MYELOID CELL LEUKAEMIA		P-value of test of trend or comparison	.9676	.1587	1	1
		Number of animals reported with tumor	1	4	0	0
HARDERIAN GLANDS	ADENOMA	P-value of test of trend or comparison	.1442	.4653	.4653	.2263
		Number of animals reported with tumor	2	3	3	5
LIVER	CHOLANGIOMA	P-value of test of trend or comparison	.2545			.4941
		Number of animals reported with tumor	0	0	0	1
	HAEMANGIOMA	P-value of test of trend or comparison	.4970		.4819	
		Number of animals reported with tumor	0	0	1	0
HAEMANGIOSARCOMA		P-value of test of trend or comparison	.4432	1	1	.7412
		Number of animals reported with tumor	1	0	0	1
HEPATOCELLULAR ADENOMA		P-value of test of trend or comparison	.2656	1	1	.6741
		Number of animals reported with tumor	2	0	0	2
LN MESENTERIC	HAEMANGIOSARCOMA	P-value of test of trend or comparison	.2561			.5000
		Number of animals reported with tumor	0	0	0	1
LUNGS + BRONCHI	BRONCHIOLOALVEOLAR ADENOCARCINOMA	P-value of test of trend or comparison	.3901	.4726	.2809	.4911

Table A.9

**Table of reported tumors in Mouse Study
NDA 204447
Animal carcinogenicity study
Female mice**

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Control Size = 60</i>	<i>Low dose Size = 60</i>	<i>Mid dose Size = 60</i>	<i>High dose Size = 60</i>
		Number of animals reported with tumor	1	2	3	2
	BRONCHIOLOALVEOLAR ADENOMA	P-value of test of trend or comparison	.7928	.6345	.7342	.8482
		Number of animals reported with tumor	10	9	8	7
MAMMARY	ADENOACANTHOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
	MAMMARY ADENOCARCINOMA	P-value of test of trend or comparison	.2646	.7286	.7350	.4913
		Number of animals reported with tumor	1	1	1	2
OVARIES	CYSTADENOMA	P-value of test of trend or comparison	.5000		.4819	
		Number of animals reported with tumor	0	0	1	0
	GRANULOSA CELL TUMOUR	P-value of test of trend or comparison	.4408	.2231		.4941
		Number of animals reported with tumor	0	2	0	1
	LUTEOMA	P-value of test of trend or comparison	.8282	.7883	.4590	.9391
		Number of animals reported with tumor	3	2	4	1
PANCREAS	ISLET CELL ADENOMA	P-value of test of trend or comparison	.8864	1	.8606	1
		Number of animals reported with tumor	2	0	1	0
PITUITARY	ADENOMA, PARS DISTALIS	P-value of test of trend or comparison	.1886		.4819	.4881
		Number of animals reported with tumor	0	0	1	1
SKIN	BASAL CELL TUMOUR	P-value of test of trend or comparison	.7394	.4819		
		Number of animals reported with tumor	0	1	0	0
	FIBROSARCOMA	P-value of test of trend or comparison	.0647		.4819	.2471
		Number of animals reported with tumor	0	0	1	2
	SARCOMA NOS	P-value of test of trend or comparison	.9333	.7346	1	1
		Number of animals reported with tumor	1	1	0	0
SPLEEN	HAEMANGIOSARCOMA	P-value of test of trend or comparison	.4970		.4819	
		Number of animals reported with tumor	0	0	1	0
STOMACH	SQUAMOUS CELL PAPILLOMA	P-value of test of trend or comparison	.6167	.4819	.4819	
		Number of animals reported with tumor	0	1	1	0

Table A.9

**Table of reported tumors in Mouse Study
NDA 204447
Animal carcinogenicity study
Female mice**

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Control Size = 60</i>	<i>Low dose Size = 60</i>	<i>Mid dose Size = 60</i>	<i>High dose Size = 60</i>
THYMUS	MALIGNANT THYMOMA	P-value of test of trend or comparison	.2564			.4938
		Number of animals reported with tumor	0	0	0	1
	THYMOMA (LYMPHOID)	P-value of test of trend or comparison	.8869	.8652	.6630	1
		Number of animals reported with tumor	2	1	2	0
UTERINE CERVIX	ADENOCARCINOMA	P-value of test of trend or comparison	.2561			.5000
		Number of animals reported with tumor	0	0	0	1
	ENDOMETRIAL POLYP	P-value of test of trend or comparison	.6243	.4940	.4878	
		Number of animals reported with tumor	0	1	1	0
UTERUS	LEIOMYOMA	P-value of test of trend or comparison	.7439	.4878		
		Number of animals reported with tumor	0	1	0	0
	DECIDUOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
ENDOMETRIAL ADENOCARCINOMA	ENDOMETRIAL ADENOCARCINOMA	P-value of test of trend or comparison	.4455	1	1	.7471
		Number of animals reported with tumor	1	0	0	1
	ENDOMETRIAL ADENOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
ENDOMETRIAL POLYP	ENDOMETRIAL POLYP	P-value of test of trend or comparison	.8456	.6972	.7120	.9105
		Number of animals reported with tumor	6	5	5	3
	ENDOMETRIAL STROMAL CELL SARCOMA	P-value of test of trend or comparison	.3949	1	.7410	.7471
		Number of animals reported with tumor	1	0	1	1
GRANULAR CELL TUMOUR	GRANULAR CELL TUMOUR	P-value of test of trend or comparison	.7394	.4819		
		Number of animals reported with tumor	0	1	0	0
	HAEMANGIOMA	P-value of test of trend or comparison	.0930	.7410	1	.2991
		Number of animals reported with tumor	1	1	0	3
HAEMANGIOSARCOMA	HAEMANGIOSARCOMA	P-value of test of trend or comparison	.4970		.4819	
		Number of animals reported with tumor	0	0	1	0

Table A.9

**Table of reported tumors in Mouse Study
NDA 204447
Animal carcinogenicity study
Female mice**

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Control Size = 60</i>	<i>Low dose Size = 60</i>	<i>Mid dose Size = 60</i>	<i>High dose Size = 60</i>
	LEIOMYOMA	P-value of test of trend or comparison	.6214	.4639	.2727	.7412
		Number of animals reported with tumor	1	2	3	1

Table A.10

**Table of reported tumors in Mouse Study
NDA 204447
Animal carcinogenicity study
Male mice**

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Control Size = 60</i>	<i>Low dose Size = 60</i>	<i>Mid dose Size = 60</i>	<i>High dose Size = 60</i>
ADRENALS	CORTICAL ADENOMA	P-value of test of trend or comparison	.4136	1	1	.7597
		Number of animals reported with tumor	1	0	0	1
	SUBCAPSULAR CELL ADENOMA	P-value of test of trend or comparison	.1770	.2130	.2815	.1576
		Number of animals reported with tumor	2	6	5	6
FEMUR INC. JOINT	HAEMANGIOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
H-POIETIC TUMOUR	HISTIOCYTIC SARCOMA	P-value of test of trend or comparison	.1133	.5529	.5366	.2597
		Number of animals reported with tumor	0	1	1	2
	MALIGNANT LYMPHOMA	P-value of test of trend or comparison	.9738	.9324	.6127	.9967
		Number of animals reported with tumor	7	4	8	1
	MALIGNANT MAST CELL TUMOUR	P-value of test of trend or comparison	.2335			.5065
		Number of animals reported with tumor	0	0	0	1
	MYELOID CELL LEUKAEMIA	P-value of test of trend or comparison	.8220	.3086		
		Number of animals reported with tumor	0	2	0	0
HARDERIAN GLANDS	ADENOMA	P-value of test of trend or comparison	.7804	.9879	.8638	.9536
		Number of animals reported with tumor	13	6	10	7
JEJUNUM	ADENOMA	P-value of test of trend or comparison	.5000		.5366	
		Number of animals reported with tumor	0	0	1	0
KIDNEYS	TUBULAR ADENOMA	P-value of test of trend or comparison	.2335			.5065
		Number of animals reported with tumor	0	0	0	1
	TUBULAR CARCINOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
LIVER	CHOLANGIOMA	P-value of test of trend or comparison	.2335			.5065
		Number of animals reported with tumor	0	0	0	1
	HAEMANGIOMA	P-value of test of trend or comparison	.1654	.9570	.7066	.5000
		Number of animals reported with tumor	3	1	3	4
	HAEMANGIOSARCOMA	P-value of test of trend or comparison	.4970		.5366	

Table A.10

Table of reported tumors in Mouse Study
NDA 204447
Animal carcinogenicity study
Male mice

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Control Size = 60</i>	<i>Low dose Size = 60</i>	<i>Mid dose Size = 60</i>	<i>High dose Size = 60</i>
		Number of animals reported with tumor	0	0	1	0
	HEPATOCELLULAR ADENOMA	P-value of test of trend or comparison	.0283	.5851	.8423	.0964
		Number of animals reported with tumor	11	13	9	18
	HEPATOCELLULAR CARCINOMA	P-value of test of trend or comparison	.1754	.8442	1	.5000
		Number of animals reported with tumor	4	3	0	5
LN MANDIBULAR	HAEMANGIOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
LUNGS + BRONCHI	BRONCHIOLOALVEOLAR ADENOCARCINOMA	P-value of test of trend or comparison	.9905	.1010	.5918	.9745
		Number of animals reported with tumor	4	11	5	1
	BRONCHIOLOALVEOLAR ADENOMA	P-value of test of trend or comparison	.6511	.8495	.4769	.8014
		Number of animals reported with tumor	18	17	22	16
PANCREAS	ISLET CELL ADENOMA	P-value of test of trend or comparison	.7771	.5542		
		Number of animals reported with tumor	0	1	0	0
PITUITARY	ADENOMA, PARS DISTALIS	P-value of test of trend or comparison	.7725	.5476		
		Number of animals reported with tumor	0	1	0	0
SEMINAL VESICLES	CARCINOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
	GRANULAR CELL TUMOUR	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
SKELETAL MUSCLE	HAEMANGIOSARCOMA	P-value of test of trend or comparison	.2381			.5128
		Number of animals reported with tumor	0	0	0	1
SKIN	FIBROSARCOMA	P-value of test of trend or comparison	.8783	.3245	.7532	.8891
		Number of animals reported with tumor	2	5	2	1
SPLEEN	HAEMANGIOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	4	0	0	0
STOMACH	HAEMANGIOMA	P-value of test of trend or comparison	.7725	.5476		
		Number of animals reported with tumor	0	1	0	0

Table A.10

**Table of reported tumors in Mouse Study
NDA 204447
Animal carcinogenicity study
Male mice**

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Control Size = 60</i>	<i>Low dose Size = 60</i>	<i>Mid dose Size = 60</i>	<i>High dose Size = 60</i>
	SQUAMOUS CELL PAPILLOMA	P-value of test of trend or comparison	.9493	.7983	1	1
		Number of animals reported with tumor	1	1	0	0
TESTES	INTERSTITIAL (LEYDIG) CELL ADENOMA	P-value of test of trend or comparison	.8483	.9835	.9802	.9727
		Number of animals reported with tumor	4	1	1	1

Table A.11

Table of reported tumors in Mouse Study
NDA 204447
Animal carcinogenicity study
Female mice
Composite endpoints

<i>Composite endpoint</i>	<i>Quantity</i>	<i>Control Size = 60</i>	<i>Low dose Size = 60</i>	<i>Mid dose Size = 60</i>	<i>High dose Size = 60</i>
Bronchioloalveolar tumors	P-value of test of trend or comparison	.7941	.6235	.5215	.8367
	Number of animals reported with tumor	11	10	11	8
Cytadenomas, granulosa cell tumors and luteomas	P-value of test of trend or comparison	.7687	.4590	.3023	.8126
	Number of animals reported with tumor	3	4	5	2
Endometrial polyps and stromal cell polyps, and stromal cell sarcomas	P-value of test of trend or comparison	.8262	.6775	.5772	.8873
	Number of animals reported with tumor	7	6	7	4
Fibromas, fibrosarcomas, and sarcoms (NOS) of the skin	P-value of test of trend or comparison	.2695	.7346	.7346	.5000
	Number of animals reported with tumor	1	1	1	2
Granular cell tumors	P-value of test of trend or comparison	.7394	.4819		
	Number of animals reported with tumor	0	1	0	0
Haemangiomas and Haemangiosarcomas	P-value of test of trend or comparison	.0566	.8659	.4543	.2073
	Number of animals reported with tumor	2	1	3	5
Hepatocellular tumors	P-value of test of trend or comparison	.2656	1	1	.6741
	Number of animals reported with tumor	2	0	0	2
Islet cell tumors	P-value of test of trend or comparison	.8864	1	.8606	1
	Number of animals reported with tumor	2	0	1	0
Mammary adenocarcinomas and adenoacanthomas	P-value of test of trend or comparison	.4079	.8610	.8659	.6832
	Number of animals reported with tumor	2	1	1	2
Mammary adenocarcinomas and adenomas	P-value of test of trend or comparison	.2646	.7286	.7350	.4913
	Number of animals reported with tumor	1	1	1	2
Pars distalis and pars intermedia tumors	P-value of test of trend or comparison	.1886		.4819	.4881
	Number of animals reported with tumor	0	0	1	1
Pheochromocytomas	P-value of test of trend or comparison	.7410	.4881		
	Number of animals reported with tumor	0	1	0	0
Squamous cell tumors of the skin	P-value of test of trend or comparison	.6214	.4819	.4819	
	Number of animals reported with tumor	0	1	1	0

Table A.11

**Table of reported tumors in Mouse Study
NDA 204447
Animal carcinogenicity study
Female mice
Composite endpoints**

<i>Composite endpoint</i>	<i>Quantity</i>	<i>Control Size = 60</i>	<i>Low dose Size = 60</i>	<i>Mid dose Size = 60</i>	<i>High dose Size = 60</i>
Thymomas	P-value of test of trend or comparison	.6463	.8659	.6551	.8706
	Number of animals reported with tumor	2	1	2	1
Uterine and cervical adenomas, adenocarcinomas, and adenosquamous carcinomas	P-value of test of trend or comparison	.2682	1	1	.6831
	Number of animals reported with tumor	2	0	0	2

Table A.12

**Table of reported tumors in Mouse Study
NDA 204447
Animal carcinogenicity study
Male mice
Composite endpoints**

<i>Composite endpoint</i>	<i>Quantity</i>	<i>Control Size = 60</i>	<i>Low dose Size = 60</i>	<i>Mid dose Size = 60</i>	<i>High dose Size = 60</i>
Bronchioloalveolar tumors	P-value of test of trend or comparison	.9313	.4381	.5522	.9024
	Number of animals reported with tumor	21	27	25	17
Cortical cell tumors	P-value of test of trend or comparison	.4136	1	1	.7597
	Number of animals reported with tumor	1	0	0	1
Fibromas, fibrosarcomas, and sarcoms (NOS) of the skin	P-value of test of trend or comparison	.8783	.3245	.7532	.8891
	Number of animals reported with tumor	2	5	2	1
Granular cell tumors	P-value of test of trend or comparison	1	1	1	1
	Number of animals reported with tumor	1	0	0	0
Haemangiomas and Haemangiosarcomas	P-value of test of trend or comparison	.4278	.9917	.9343	.8259
	Number of animals reported with tumor	7	2	4	5
Hepatocellular tumors	P-value of test of trend or comparison	.0099	.6693	.9511	.0611
	Number of animals reported with tumor	14	15	9	22
Islet cell tumors	P-value of test of trend or comparison	.7771	.5542		
	Number of animals reported with tumor	0	1	0	0
Pars distalis and pars intermedia tumors	P-value of test of trend or comparison	.7725	.5476		
	Number of animals reported with tumor	0	1	0	0
Squamous cell tumors of the skin	P-value of test of trend or comparison	.9493	.7983	1	1
	Number of animals reported with tumor	1	1	0	0
Testicular tumors	P-value of test of trend or comparison	.8483	.9835	.9802	.9727
	Number of animals reported with tumor	4	1	1	1
Tubular cell tumors	P-value of test of trend or comparison	.4136	1	1	.7597
	Number of animals reported with tumor	1	0	0	1

Table A.13

Table of tumors reported significant ($\alpha < 0.05$) in at least one arm - Mouse Study
NDA 204447
Animal carcinogenicity study
Male mice

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Control</i>	<i>Low dose</i>	<i>Mid dose</i>	<i>High dose</i>
LIVER	HEPATOCELLULAR ADENOMA	P-value of test of trend or comparison	.0283	.5851	.8423	.0964
		Number of animals reported with tumor	11	13	9	18
		Poly-3 adjusted incidence rate	27%	27%	20%	43%
		95% CI for poly-3 adjusted incidence rate (%)	(13.9,42.9)	(14.9,41.8)	(9.36,34.6)	(27,59.0)
		Poly-3 adjusted number of animals at risk	41.3	48.3	45.6	42.2

Table A.14

Table of tumors reported significant ($\alpha < 0.05$) in at least one arm - Mouse Study
NDA 204447
Animal carcinogenicity study
Male mice
Composite endpoints

<i>Composite endpoint</i>	<i>Quantity</i>	<i>Control</i>	<i>Low dose</i>	<i>Mid dose</i>	<i>High dose</i>
Hepatocellular tumors	P-value of test of trend or comparison	.0099	.6693	.9511	.0611
	Number of animals reported with tumor	14	15	9	22
	Poly-3 adjusted incidence rate	33%	31%	20%	52%
	95% CI for poly-3 adjusted incidence rate (%)	(19.1,49.5)	(18.3,46.3)	(9.36,34.6)	(35.5,68.0)
	Poly-3 adjusted number of animals at risk	42.3	48.5	45.6	42.4

**Tumor counts for organs reported widely analyzed
NDA 204447
Animal carcinogenicity study
Mice**

<i>Species and Sex</i>	<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Control</i>	<i>Low dose</i>	<i>Mid dose</i>	<i>High dose</i>
Mice - Female	BONE	OSTEOMA	Number of tumors found	0	0	2	0
			Number of animals examined	1	3	4	3
Mice - Male	ADIPOSE TISSUE	LIPOMA	Number of tumors found	0	0	1	0
			Number of animals examined	0	3	4	5
	HEAD	FIBROSARCOMA	Number of tumors found	0	0	0	1
			Number of animals examined	1	1	0	2

Table A.15

A.3 Unexamined and autolytic organs

APPEARS THIS WAY ON ORIGINAL



Table A.16

**Organs reported as autolytic
NDA 204447
Animal carcinogenicity study
Female Mice**

<i>Organ or tissue name</i>	<i>Control(count)</i>	<i>Control(%)</i>	<i>Low dose(count)</i>	<i>Low dose(%)</i>	<i>Mid dose(count)</i>	<i>Mid dose(%)</i>	<i>High dose(count)</i>	<i>High dose(%)</i>	<i>Total(count)</i>	<i>Total(%)</i>
CAECUM	1	1.7%	1	1.7%	1	1.7%	2	3.3%	5	2.1%
GALL BLADDER	1	1.7%	4	6.7%	7	12%	4	6.7%	16	6.7%
STOMACH	1	1.7%	1	0.4%

Table A.17

**Organs reported as autolytic
NDA 204447
Animal carcinogenicity study
Male Mice**

<i>Organ or tissue name</i>	<i>Control(count)</i>	<i>Control(%)</i>	<i>Low dose(count)</i>	<i>Low dose(%)</i>	<i>Mid dose(count)</i>	<i>Mid dose(%)</i>	<i>High dose(count)</i>	<i>High dose(%)</i>	<i>Total(count)</i>	<i>Total(%)</i>
JEJUNUM	3	5.0%	3	5.0%	2	3.3%	1	1.7%	9	3.8%

Table A.18

**Organs reported as unexamined
NDA 204447
Animal carcinogenicity study
Female Mice**

<i>Organ or tissue name</i>	<i>Control(count)</i>	<i>Control(%)</i>	<i>Low dose(count)</i>	<i>Low dose(%)</i>	<i>Mid dose(count)</i>	<i>Mid dose(%)</i>	<i>High dose(count)</i>	<i>High dose(%)</i>	<i>Total(count)</i>	<i>Total(%)</i>
BONE	59	98%	57	95%	56	93%	57	95%	229	95%
HARDERIAN GLANDS	1	1.7%	.	.	1	0.4%
LN MESENTERIC	1	1.7%	1	0.4%
OVARIES	.	.	1	1.7%	1	0.4%
PANCREAS	1	1.7%	.	.	1	0.4%
PITUITARY	1	1.7%	2	3.3%	.	.	1	1.7%	4	1.7%
SKIN	1	1.7%	1	0.4%
THYMUS	5	8.3%	5	8.3%	5	8.3%	4	6.7%	19	7.9%
UTERINE CERVIX	2	3.3%	1	1.7%	3	1.3%

Table A.19

**Organs reported as unexamined
NDA 204447
Animal carcinogenicity study
Male Mice**

<i>Organ or tissue name</i>	<i>Control(count)</i>	<i>Control(%)</i>	<i>Low dose(count)</i>	<i>Low dose(%)</i>	<i>Mid dose(count)</i>	<i>Mid dose(%)</i>	<i>High dose(count)</i>	<i>High dose(%)</i>	<i>Total(count)</i>	<i>Total(%)</i>
ADIPOSE TISSUE	60	100%	57	95%	56	93%	55	92%	228	95%
HARDERIAN GLANDS	1	1.7%	1	1.7%	2	0.8%
HEAD	59	98%	59	98%	60	100%	58	97%	236	98%
PANCREAS	1	1.7%	1	0.4%
PITUITARY	1	1.7%	.	.	1	0.4%
SPLEEN	1	1.7%	.	.	1	0.4%

A.4 Bodyweight changes

Table A.20: Weight changes by group (mice)

Sex	Vehicle control	Lu AA 21004					
	Δ_{CP}	Δ_L	$\frac{\Delta_L}{\Delta_{CP}} - 1$	Δ_M	$\frac{\Delta_M}{\Delta_{CP}} - 1$	Δ_H	$\frac{\Delta_H}{\Delta_{CP}} - 1$
Female	12.1	13.3	10%	14.1	17%	13.2	9%
Male	15.4	14.9	-3%	15.9	3%	18.4	19%

Appendix B

Tables from rat study

B.1 Survival analysis

APPEARS THIS WAY ON ORIGINAL



Table B.1

**Survival rates at key times
NDA 204447
Animal carcinogenicity study
Rats**

<i>Species and Sex</i>	<i>Dose Group</i>	<i>Dose (mg per kg)</i>	<i>Number at start</i>	<i>Number alive after 52 weeks</i>	<i>Percentage alive after 52 weeks</i>	<i>Number alive after 78 weeks</i>	<i>Percentage alive after 78 weeks</i>	<i>Number alive after 90 weeks</i>	<i>Percentage alive after 90 weeks</i>	<i>Number sacrificed</i>	<i>Percentage sacrificed</i>	<i>Maximum survival (weeks)</i>
Rats - Female	Water control	0	55	54	98%	49	89%	45	82%	37	67%	105
	Vehicle control	0	55	53	96%	48	87%	41	75%	30	55%	105
	Low dose	5	55	54	98%	49	89%	45	82%	39	71%	105
	Mid dose	15	55	54	98%	46	84%	36	65%	29	53%	105
	High dose	40	55	53	96%	46	84%	37	67%	20	36%	105
Rats - Male	Water control	0	55	55	100%	52	95%	48	87%	39	71%	107
	Vehicle control	0	55	55	100%	52	95%	50	91%	45	82%	107
	Low dose	2	55	55	100%	51	93%	48	87%	41	75%	107
	Mid dose	7	55	54	98%	53	96%	49	89%	44	80%	107
	High dose	20	55	54	98%	51	93%	46	84%	37	67%	107

Table B.2

**Log-rank tests of survival
NDA 204447
Animal carcinogenicity study
Rats**

<i>Sex</i>	<i>Test of homogeneity: chi squared statistic</i>	<i>Test of homogeneity: degrees of freedom</i>	<i>Number of groups</i>	<i>Test of homogeneity: p-value</i>	<i>Test of trend (two tailed): p-value</i>	<i>Test of trend (one tailed): p-value</i>
Female	15.8373	4	5	0.0032	0.0005	0.0003
Male	3.9697	4	5	0.4101	0.2514	0.1257

Table B.3

**Pairwise comparisons (log-rank) of survival between treated groups and controls
NDA 204447
Animal carcinogenicity study
Rats**

<i>Species and Sex</i>	<i>Quantity</i>	<i>Low dose</i>	<i>Mid dose</i>	<i>High dose</i>
Rats - Female	Chi squared test statistic	2.3436	0.1632	3.5734
	p-value of comparison with control	0.1258	0.6863	0.0587
Rats - Male	Chi squared test statistic	0.7233	0.0298	2.8643
	p-value of comparison with control	0.3951	0.8630	0.0906

Table B.4

**Log-rank tests of heterogeneity of survival between control groups
NDA 204447
Animal carcinogenicity study**

<i>Species and Sex</i>	<i>Chi²</i>	<i>DF</i>	<i>P-value</i>
Rats - Female	1.5997	1	0.2059
Rats - Male	1.6104	1	0.2044

B.2 Tumor analysis

APPEARS THIS WAY ON ORIGINAL

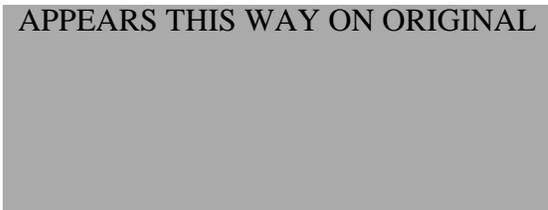


Table B.5

**Primary organs in study of female rats
NDA 204447
Animal carcinogenicity study**

<i>Organ or tissue name</i>
ADRENALS
BRAIN
CLITORAL GLANDS
H-POIETIC TUMOUR
HEART
JEJUNUM
LIVER
LN MESENTERIC
MAMMARY
OVARIES
PANCREAS
PARATHYROIDS
PITUITARY
RECTUM
SKELETAL MUSCLE
SKIN
THYMUS
THYROIDS
TONGUE
UTERINE CERVIX
UTERUS
VAGINA

Table B.6

**Primary organs in study of male rats
NDA 204447
Animal carcinogenicity study**

<i>Organ or tissue name</i>
ADRENALS
BRAIN
CAECUM
EPIDIDYMIDES
H-POIETIC TUMOUR
JEJUNUM
KIDNEYS
LIVER
LN MANDIBULAR
LN MESENTERIC
MAMMARY
PANCREAS
PARATHYROIDS
PITUITARY
PREPUTIAL GLANDS
PROSTATE
RECTUM
SKELETAL MUSCLE
SKIN
SPINAL C. CERV.
SPINAL C. LUMB.
SPLEEN
STOMACH
TESTES
THYMUS
THYROIDS

Table B.7

**Secondary organs in study of female rats
NDA 204447
Animal carcinogenicity study**

<i>Organ or tissue name</i>
ABDOMEN
ADIPOSE TISSUE
HEAD
LN LUMBAR
PINNAE

Table B.8

**Secondary organs in study of male rats
NDA 204447
Animal carcinogenicity study**

<i>Organ or tissue name</i>
ABDOMEN
LN CERVICAL
LN DEEP CERVICAL
LN RENAL
ORAL CAVITY
TAIL

Table B.9

Table of reported tumors in Rat Study
NDA 204447
Animal carcinogenicity study
Female rats

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Vehicle control Size = 55</i>	<i>Low dose Size = 55</i>	<i>Mid dose Size = 55</i>	<i>High dose Size = 55</i>
ADRENALS	CORTICAL ADENOMA	P-value of test of trend or comparison	.5987	.5330	1	.7220
		Number of animals reported with tumor	1	2	0	1
	CORTICAL CARCINOMA	P-value of test of trend or comparison	.4709		.4884	
		Number of animals reported with tumor	0	0	1	0
	MALIGNANT PHAEOCHROMOCYTOMA	P-value of test of trend or comparison	.7442	.5165		
		Number of animals reported with tumor	0	1	0	0
	PHAEOCHROMOCYTOMA	P-value of test of trend or comparison	.4100	1	1	.7286
		Number of animals reported with tumor	1	0	0	1
BRAIN	GRANULAR CELL TUMOUR	P-value of test of trend or comparison	.7442	.5165		
		Number of animals reported with tumor	0	1	0	0
	OLIGODENDROGLIOMA	P-value of test of trend or comparison	.4709		.4884	
		Number of animals reported with tumor	0	0	1	0
CLITORAL GLANDS	SQUAMOUS CELL PAPILLOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
H-POIETIC TUMOUR	MALIGNANT LYMPHOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
HEART	ENDOCARDIAL SCHWANNOMA	P-value of test of trend or comparison	.1618		.4884	.4699
		Number of animals reported with tumor	0	0	1	1
JEJUNUM	LEIOMYOMA	P-value of test of trend or comparison	.2169			.4557
		Number of animals reported with tumor	0	0	0	1
LIVER	CHOLANGIOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
	HEPATOCELLULAR ADENOMA	P-value of test of trend or comparison	.0072	.5165		.0474
		Number of animals reported with tumor	0	1	0	4
	HEPATOCELLULAR CARCINOMA	P-value of test of trend or comparison	.4709		.4884	
		Number of animals reported with tumor	0	0	1	0
LN MESENTERIC	HAEMANGIOMA	P-value of test of trend or comparison	.1828	.8454	.9361	.4318

Table B.9

Table of reported tumors in Rat Study
NDA 204447
Animal carcinogenicity study
Female rats

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Vehicle control Size = 55</i>	<i>Low dose Size = 55</i>	<i>Mid dose Size = 55</i>	<i>High dose Size = 55</i>
		Number of animals reported with tumor	3	2	1	4
	HAEMANGIOSARCOMA	P-value of test of trend or comparison	.4031	1	1	.7220
		Number of animals reported with tumor	1	0	0	1
MAMMARY	MAMMARY ADENOCARCINOMA	P-value of test of trend or comparison	.6451	.8910	.8706	.8610
		Number of animals reported with tumor	2	1	1	1
	MAMMARY ADENOMA	P-value of test of trend or comparison	.7442	.5165		
		Number of animals reported with tumor	0	1	0	0
	MAMMARY FIBROADENOMA	P-value of test of trend or comparison	.9588	.9945	.4146	.9972
		Number of animals reported with tumor	16	7	18	5
OVARIES	YOLK SAC CARCINOMA	P-value of test of trend or comparison	.4740		.4943	
		Number of animals reported with tumor	0	0	1	0
	CYSTADENOMA	P-value of test of trend or comparison	.4709		.4884	
		Number of animals reported with tumor	0	0	1	0
	GRANULOSA CELL TUMOUR	P-value of test of trend or comparison	.9357	.7690	1	1
		Number of animals reported with tumor	1	1	0	0
PANCREAS	ISLET CELL ADENOMA	P-value of test of trend or comparison	.2312			.4762
		Number of animals reported with tumor	0	0	0	1
PARATHYROIDS	CHIEF CELL ADENOMA	P-value of test of trend or comparison	.4057	1	1	.7205
		Number of animals reported with tumor	1	0	0	1
PITUITARY	ADENOMA, PARS DISTALIS	P-value of test of trend or comparison	.1914	.8703	.5289	.4148
		Number of animals reported with tumor	33	29	33	31
	ADENOMA, PARS INTERMEDIA	P-value of test of trend or comparison	.7215	1	.7412	1
		Number of animals reported with tumor	1	0	1	0
RECTUM	POLYPOID ADENOMA	P-value of test of trend or comparison	.0034		.4884	.0448
		Number of animals reported with tumor	0	0	1	4
SKELETAL MUSCLE	FIBROMA	P-value of test of trend or comparison	.4709		.4884	
		Number of animals reported with tumor	0	0	1	0

Table B.9

**Table of reported tumors in Rat Study
NDA 204447
Animal carcinogenicity study
Female rats**

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Vehicle control Size = 55</i>	<i>Low dose Size = 55</i>	<i>Mid dose Size = 55</i>	<i>High dose Size = 55</i>
SKIN	BASAL CELL CARCINOMA	P-value of test of trend or comparison	.9335	.7635	1	1
		Number of animals reported with tumor	1	1	0	0
	FIBROMA	P-value of test of trend or comparison	.7442	.5165		
		Number of animals reported with tumor	0	1	0	0
	KERATOACANTHOMA	P-value of test of trend or comparison	.2267			.4699
		Number of animals reported with tumor	0	0	0	1
	LEIOMYOSARCOMA	P-value of test of trend or comparison	.7442	.5165		
		Number of animals reported with tumor	0	1	0	0
	LIPOMA	P-value of test of trend or comparison	.2312			.4762
		Number of animals reported with tumor	0	0	0	1
	SQUAMOUS CELL PAPILOMA	P-value of test of trend or comparison	.4709		.4884	
		Number of animals reported with tumor	0	0	1	0
THYMUS	THYMOMA	P-value of test of trend or comparison	.4506	.8737	.9827	.6812
		Number of animals reported with tumor	5	3	1	4
THYROIDS	C-CELL ADENOMA	P-value of test of trend or comparison	.2055	.0088	.4782	.0495
		Number of animals reported with tumor	2	11	3	7
	C-CELL CARCINOMA	P-value of test of trend or comparison	.8737	1	.8706	1
		Number of animals reported with tumor	2	0	1	0
	FOLLICULAR CELL ADENOMA	P-value of test of trend or comparison	.9910	.8004	.9669	1
		Number of animals reported with tumor	4	3	1	0
	FOLLICULAR CELL CARCINOMA	P-value of test of trend or comparison	.1294	.5165	.1122	.2177
		Number of animals reported with tumor	0	1	3	2
TONGUE	SQUAMOUS CELL CARCINOMA	P-value of test of trend or comparison	.4709		.4884	
		Number of animals reported with tumor	0	0	1	0
UTERINE CERVIX	ADENOSQUAMOUS CARCINOMA	P-value of test of trend or comparison	.9335	.7635	1	1
		Number of animals reported with tumor	1	1	0	0
	ENDOMETRIAL STROMAL POLYP	P-value of test of trend or comparison	.8723	.1377		

Table B.9

Table of reported tumors in Rat Study
NDA 204447
Animal carcinogenicity study
Female rats

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Vehicle control Size = 55</i>	<i>Low dose Size = 55</i>	<i>Mid dose Size = 55</i>	<i>High dose Size = 55</i>
		Number of animals reported with tumor	0	3	0	0
	LEIOMYOSARCOMA	P-value of test of trend or comparison	.4709		.4884	
		Number of animals reported with tumor	0	0	1	0
	MALIGNANT SCHWANNOMA	P-value of test of trend or comparison	.7442	.5165		
		Number of animals reported with tumor	0	1	0	0
	SCHWANNOMA	P-value of test of trend or comparison	.7457	.5217		
		Number of animals reported with tumor	0	1	0	0
UTERUS	ENDOMETRIAL ADENOCARCINOMA	P-value of test of trend or comparison	.1846	.5316	.6741	.3047
		Number of animals reported with tumor	2	3	2	4
	ENDOMETRIAL ADENOMA	P-value of test of trend or comparison	.5988	.8394	1	.7889
		Number of animals reported with tumor	3	2	0	2
	ENDOMETRIAL POLYP	P-value of test of trend or comparison	.8199	.9814	.8772	.9556
		Number of animals reported with tumor	11	5	7	5
	LEIOMYOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	2	0	0	0
	LEIOMYOSARCOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0

Table B.10

Table of reported tumors in Rat Study
NDA 204447
Animal carcinogenicity study
Male rats

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Vehicle control Size = 55</i>	<i>Low dose Size = 55</i>	<i>Mid dose Size = 55</i>	<i>High dose Size = 55</i>
ADRENALS	CORTICAL ADENOMA	P-value of test of trend or comparison	.1827		.5000	.4898
		Number of animals reported with tumor	0	0	1	1
	MALIGNANT PHAEOCHROMOCYTOMA	P-value of test of trend or comparison	.2437			.4898
		Number of animals reported with tumor	0	0	0	1
	PHAEOCHROMOCYTOMA	P-value of test of trend or comparison	.7846	.7475	.5000	1
		Number of animals reported with tumor	1	1	2	0
BRAIN	ASTROCYTOMA	P-value of test of trend or comparison	.5827	1	1	.8672
		Number of animals reported with tumor	2	0	0	1
	GRANULAR CELL TUMOUR	P-value of test of trend or comparison	.6555	.3199	.8787	.6760
		Number of animals reported with tumor	2	4	1	2
	MALIGNANT GRANULAR CELL TUMOUR	P-value of test of trend or comparison	.4975		.5000	
		Number of animals reported with tumor	0	0	1	0
OLIGODENDROGLIOMA	P-value of test of trend or comparison	.2437			.4898	
	Number of animals reported with tumor	0	0	0	1	
CAECUM	FIBROMA	P-value of test of trend or comparison	.2437			.4898
		Number of animals reported with tumor	0	0	0	1
	SARCOMA-NOS	P-value of test of trend or comparison	.2437			.4898
		Number of animals reported with tumor	0	0	0	1
EPIDIDYMIDES	MESOTHELIOMA	P-value of test of trend or comparison	.2899	1	.8788	.6837
		Number of animals reported with tumor	2	0	1	2
H-POIETIC TUMOUR	HISTIOCYTIC SARCOMA	P-value of test of trend or comparison	.0173		.5050	.1137
		Number of animals reported with tumor	0	0	1	3
	MALIGNANT LYMPHOMA	P-value of test of trend or comparison	.5065	.7424	.7525	.7424
Number of animals reported with tumor		1	1	1	1	
JEJUNUM	LEIOMYOMA	P-value of test of trend or comparison	.7449	.4949		
		Number of animals reported with tumor	0	1	0	0
KIDNEYS	RENAL LIPOSARCOMA	P-value of test of trend or comparison	1	1	1	1

Table B.10

Table of reported tumors in Rat Study
NDA 204447
Animal carcinogenicity study
Male rats

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Vehicle control Size = 55</i>	<i>Low dose Size = 55</i>	<i>Mid dose Size = 55</i>	<i>High dose Size = 55</i>
		Number of animals reported with tumor	1	0	0	0
LIVER	CHOLANGIOMA	P-value of test of trend or comparison	.2437			.4898
		Number of animals reported with tumor	0	0	0	1
	HEPATOCELLULAR ADENOMA	P-value of test of trend or comparison	.0364	.2424	.2475	.0539
		Number of animals reported with tumor	0	2	2	4
	HEPATOCELLULAR CARCINOMA	P-value of test of trend or comparison	.3046	.4949		.4898
		Number of animals reported with tumor	0	1	0	1
LN MANDIBULAR	HAEMANGIOMA	P-value of test of trend or comparison	.2437			.4898
		Number of animals reported with tumor	0	0	0	1
LN MESENTERIC	HAEMANGIOMA	P-value of test of trend or comparison	.0040	.6059	.0628	.0163
		Number of animals reported with tumor	6	6	13	15
MAMMARY	FIBROMA	P-value of test of trend or comparison	.7694	.5095	.5000	.8711
		Number of animals reported with tumor	2	3	3	1
	FIBROSARCOMA	P-value of test of trend or comparison	.2475			.4949
		Number of animals reported with tumor	0	0	0	1
	LIPOMA	P-value of test of trend or comparison	.7462	.4949		
		Number of animals reported with tumor	0	1	0	0
PANCREAS	ACINAR CELL ADENOMA	P-value of test of trend or comparison	.9732	.3292	.8788	1
		Number of animals reported with tumor	2	4	1	0
	ISLET CELL ADENOMA	P-value of test of trend or comparison	.7731	.7475	.3087	1
		Number of animals reported with tumor	1	1	3	0
	ISLET CELL CARCINOMA	P-value of test of trend or comparison	.0675	1	.5000	.3010
		Number of animals reported with tumor	1	0	2	3
PARATHYROIDS	CHIEF CELL ADENOMA	P-value of test of trend or comparison	.6121	.0927	.7576	.4761
		Number of animals reported with tumor	1	5	1	2
PITUITARY	ADENOMA, PARS DISTALIS	P-value of test of trend or comparison	.1459	.1511	.1247	.1040
		Number of animals reported with tumor	8	13	14	14

Table B.10

Table of reported tumors in Rat Study
NDA 204447
Animal carcinogenicity study
Male rats

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Vehicle control Size = 55</i>	<i>Low dose Size = 55</i>	<i>Mid dose Size = 55</i>	<i>High dose Size = 55</i>
	ADENOMA, PARS INTERMEDIA	P-value of test of trend or comparison	.4249	1	1	.7369
		Number of animals reported with tumor	1	0	0	1
PREPUTIAL GLANDS	SQUAMOUS CELL CARCINOMA	P-value of test of trend or comparison	.7286	.2424	.5000	
		Number of animals reported with tumor	0	2	1	0
	SQUAMOUS CELL PAPILOMA	P-value of test of trend or comparison	.3046	.4949		.4898
		Number of animals reported with tumor	0	1	0	1
PROSTATE	ADENOCARCINOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
	ADENOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
RECTUM	POLYPOID ADENOMA	P-value of test of trend or comparison	.4975		.5000	
		Number of animals reported with tumor	0	0	1	0
SKELETAL MUSCLE	HAEMANGIOSARCOMA	P-value of test of trend or comparison	.9365	.7475	1	1
		Number of animals reported with tumor	1	1	0	0
SKIN	BASAL CELL TUMOUR	P-value of test of trend or comparison	.8792	.8750	.6913	1
		Number of animals reported with tumor	2	1	2	0
	FIBROMA	P-value of test of trend or comparison	.1600	1	.7525	.4845
		Number of animals reported with tumor	1	0	1	2
	HAEMANGIOSARCOMA	P-value of test of trend or comparison	.2437			.4898
		Number of animals reported with tumor	0	0	0	1
	KERATOACANTHOMA	P-value of test of trend or comparison	.8338	.6611	.3575	.9388
		Number of animals reported with tumor	3	3	5	1
	LIPOSARCOMA	P-value of test of trend or comparison	.4975		.5000	
		Number of animals reported with tumor	0	0	1	0
	SARCOMA-NOS	P-value of test of trend or comparison	.2475			.4949
		Number of animals reported with tumor	0	0	0	1
SPINAL C. LUMB.	MENINGIOMA	P-value of test of trend or comparison	1	1	1	1

Table B.10

Table of reported tumors in Rat Study
NDA 204447
Animal carcinogenicity study
Male rats

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Vehicle control Size = 55</i>	<i>Low dose Size = 55</i>	<i>Mid dose Size = 55</i>	<i>High dose Size = 55</i>
SPLEEN	HAEMANGIOMA	Number of animals reported with tumor	1	0	0	0
		P-value of test of trend or comparison	1	1	1	1
	HAEMANGIOSARCOMA	Number of animals reported with tumor	1	0	0	0
P-value of test of trend or comparison		.8096	.2424			
Number of animals reported with tumor		0	2	0	0	
STOMACH	LEIOMYOMA	P-value of test of trend or comparison	.4897		.4845	
		Number of animals reported with tumor	0	0	1	0
	MALIGNANT SCHWANNOMA	P-value of test of trend or comparison	.7423	.4949		
Number of animals reported with tumor		0	1	0	0	
TESTES		INTERSTITIAL (LEYDIG) CELL ADENOMA	P-value of test of trend or comparison	.1737	.8788	1
	Number of animals reported with tumor		2	1	0	3
	MESOTHELIOMA	P-value of test of trend or comparison	.3046	.4949		.4898
Number of animals reported with tumor		0	1	0	1	
THYMUS		MALIGNANT THYMOMA	P-value of test of trend or comparison	.4923		.5000
	Number of animals reported with tumor		0	0	1	0
	THYMOMA	P-value of test of trend or comparison	.7367	.7475	.1811	1
Number of animals reported with tumor		1	1	4	0	
THYROIDS		C-CELL ADENOMA	P-value of test of trend or comparison	.3569	.5813	.3763
	Number of animals reported with tumor		7	7	9	8
	C-CELL CARCINOMA	P-value of test of trend or comparison	.2247	.7424	1	.4770
Number of animals reported with tumor		1	1	0	2	
FOLLICULAR CELL ADENOMA		P-value of test of trend or comparison	.0553	.8932	.7900	.2505
	Number of animals reported with tumor	4	2	3	7	
	FOLLICULAR CELL CARCINOMA	P-value of test of trend or comparison	.7492	.6760	1	.8672
Number of animals reported with tumor		2	2	0	1	

Table B.11

Table of reported tumors in Rat Study
NDA 204447
Animal carcinogenicity study
Female rats
Composite endpoints

<i>Composite endpoint</i>	<i>Quantity</i>	<i>Vehicle control Size = 55</i>	<i>Low dose Size = 55</i>	<i>Mid dose Size = 55</i>	<i>High dose Size = 55</i>
C-cell tumors	P-value of test of trend or comparison	.3406	.0533	.6173	.1833
	Number of animals reported with tumor	4	11	4	7
Cortical cell tumors	P-value of test of trend or comparison	.5627	.5330	.7412	.7220
	Number of animals reported with tumor	1	2	1	1
Cytadenomas, granulosa cell tumors and luteomas	P-value of test of trend or comparison	.8065	.7690	.7412	1
	Number of animals reported with tumor	1	1	1	0
Endometrial polyps and stromal cell polyps, and stromal cell sarcomas	P-value of test of trend or comparison	.8946	.8825	.8772	.9556
	Number of animals reported with tumor	11	8	7	5
Fibromas, fibrosarcomas, and sarcoms (NOS) of the skin	P-value of test of trend or comparison	.7442	.5165		
	Number of animals reported with tumor	0	1	0	0
Follicular cell tumors	P-value of test of trend or comparison	.7547	.6673	.6036	.8627
	Number of animals reported with tumor	4	4	4	2
Glial cell tumors	P-value of test of trend or comparison	.4709		.4884	
	Number of animals reported with tumor	0	0	1	0
Granular cell tumors	P-value of test of trend or comparison	.7442	.5165		
	Number of animals reported with tumor	0	1	0	0
Haemangiomas and Haemangiosarcomas	P-value of test of trend or comparison	.2319	.9569	.9850	.5680
	Number of animals reported with tumor	5	2	1	5
Hepatocellular tumors	P-value of test of trend or comparison	.0103	.5165	.4884	.0474
	Number of animals reported with tumor	0	1	1	4
Islet cell tumors	P-value of test of trend or comparison	.2312			.4762
	Number of animals reported with tumor	0	0	0	1
Mammary adenocarcinomas and adenoacanthomas	P-value of test of trend or comparison	.6451	.8910	.8706	.8610
	Number of animals reported with tumor	2	1	1	1
Mammary adenocarcinomas and adenomas	P-value of test of trend or comparison	.7217	.7163	.8706	.8610
	Number of animals reported with tumor	2	2	1	1

Table B.11

Table of reported tumors in Rat Study
NDA 204447
Animal carcinogenicity study
Female rats
Composite endpoints

<i>Composite endpoint</i>	<i>Quantity</i>	<i>Vehicle control Size = 55</i>	<i>Low dose Size = 55</i>	<i>Mid dose Size = 55</i>	<i>High dose Size = 55</i>
Mammary adenomas and fibroadenomas	P-value of test of trend or comparison	.9666	.9883	.4146	.9972
	Number of animals reported with tumor	16	8	18	5
Mesotheliomas	P-value of test of trend or comparison	.1653		.4884	.4762
	Number of animals reported with tumor	0	0	1	1
Pars distalis and pars intermedia tumors	P-value of test of trend or comparison	.2282	.9095	.5277	.4952
	Number of animals reported with tumor	34	29	34	31
Pheochromocytomas	P-value of test of trend or comparison	.5169	.7690	1	.7286
	Number of animals reported with tumor	1	1	0	1
Squamous cell tumors of the skin	P-value of test of trend or comparison	.6756	.7157	.6653	.8511
	Number of animals reported with tumor	2	2	2	1
Thymomas	P-value of test of trend or comparison	.4606	.8802	.9839	.6955
	Number of animals reported with tumor	5	3	1	4
Uterine and cervical adenomas, adenocarcinomas, and adenosquamous carcinomas	P-value of test of trend or comparison	.4213	.6517	.9637	.5530
	Number of animals reported with tumor	6	6	2	6

Table B.12

**Table of reported tumors in Rat Study
NDA 204447
Animal carcinogenicity study
Male rats
Composite endpoints**

<i>Composite endpoint</i>	<i>Quantity</i>	<i>Vehicle control Size = 55</i>	<i>Low dose Size = 55</i>	<i>Mid dose Size = 55</i>	<i>High dose Size = 55</i>
C-cell tumors	P-value of test of trend or comparison	.2015	.4661	.3763	.2669
	Number of animals reported with tumor	7	8	9	10
Cortical cell tumors	P-value of test of trend or comparison	.1827		.5000	.4898
	Number of animals reported with tumor	0	0	1	1
Fibromas, fibrosarcomas, and sarcoms (NOS) of the skin	P-value of test of trend or comparison	.0611	1	.7525	.3010
	Number of animals reported with tumor	1	0	1	3
Follicular cell tumors	P-value of test of trend or comparison	.1331	.8242	.9200	.3562
	Number of animals reported with tumor	6	4	3	8
Glial cell tumors	P-value of test of trend or comparison	.2630	1	1	.6681
	Number of animals reported with tumor	2	0	0	2
Granular cell tumors	P-value of test of trend or comparison	.6512	.3199	.6913	.6760
	Number of animals reported with tumor	2	4	2	2
Haemangiomas and Haemangiosarcomas	P-value of test of trend or comparison	.0180	.3763	.1059	.0313
	Number of animals reported with tumor	7	9	13	15
Hepatocellular tumors	P-value of test of trend or comparison	.0277	.1175	.2475	.0252
	Number of animals reported with tumor	0	3	2	5
Islet cell tumors	P-value of test of trend or comparison	.2461	.8750	.2180	.4903
	Number of animals reported with tumor	2	1	5	3
Mammary fibroadenomas and fibrosarcomas	P-value of test of trend or comparison	.5699	.5095	.5000	.6837
	Number of animals reported with tumor	2	3	3	2
Mesotheliomas	P-value of test of trend or comparison	.2019	.8750	.8788	.4903
	Number of animals reported with tumor	2	1	1	3
Pars distalis and pars intermedia tumors	P-value of test of trend or comparison	.1242	.2192	.1859	.1102
	Number of animals reported with tumor	9	13	14	15
Pheochromocytomas	P-value of test of trend or comparison	.4888	.7475	.5000	.7423
	Number of animals reported with tumor	1	1	2	1

Table B.12

**Table of reported tumors in Rat Study
NDA 204447
Animal carcinogenicity study
Male rats
Composite endpoints**

<i>Composite endpoint</i>	<i>Quantity</i>	<i>Vehicle control Size = 55</i>	<i>Low dose Size = 55</i>	<i>Mid dose Size = 55</i>	<i>High dose Size = 55</i>
Prostate tumors	P-value of test of trend or comparison	1	1	1	1
	Number of animals reported with tumor	2	0	0	0
Squamous cell tumors of the skin	P-value of test of trend or comparison	.8682	.3578	.3578	.8883
	Number of animals reported with tumor	4	6	6	2
Testicular tumors	P-value of test of trend or comparison	.2491	.6913	1	.4903
	Number of animals reported with tumor	2	2	0	3
Thymomas	P-value of test of trend or comparison	.7269	.7475	.1069	1
	Number of animals reported with tumor	1	1	5	0

Table B.13

**Table of tumors reported significant ($\alpha < 0.05$) in at least one arm - Rat Study
NDA 204447
Animal carcinogenicity study
Female rats**

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Vehicle control</i>	<i>Low dose</i>	<i>Mid dose</i>	<i>High dose</i>
LIVER	HEPATOCELLULAR ADENOMA	P-value of test of trend or comparison	.0072	.5165		.0474
		Number of animals reported with tumor	0	1	0	4
		Poly-3 adjusted incidence rate	0.0%	2.1%	0.0%	9.8%
		95% CI for poly-3 adjusted incidence rate (%)	(0,8.0)	(0.05,11.3)	(0,8.4)	(2.72,23.7)
		Poly-3 adjusted number of animals at risk	44.6	47.6	42.1	40.6
RECTUM	POLYPOID ADENOMA	P-value of test of trend or comparison	.0034		.4884	.0448
		Number of animals reported with tumor	0	0	1	4
		Poly-3 adjusted incidence rate	0.0%	0.0%	2.4%	10%
		95% CI for poly-3 adjusted incidence rate (%)	(0,8.0)	(0,7.5)	(0.06,12.6)	(2.79,24.2)
		Poly-3 adjusted number of animals at risk	44.6	47.6	42.1	39.5
THYROIDS	C-CELL ADENOMA	P-value of test of trend or comparison	.2055	.0088	.4782	.0495
		Number of animals reported with tumor	2	11	3	7
		Poly-3 adjusted incidence rate	4.4%	23%	6.9%	18%
		95% CI for poly-3 adjusted incidence rate (%)	(0.53,15.1)	(12,38.0)	(1.43,19.1)	(7.34,33.5)
		Poly-3 adjusted number of animals at risk	45.0	47.9	43.6	39.9

Table B.14

Table of tumors reported significant ($\alpha < 0.05$) in at least one arm - Rat Study
NDA 204447
Animal carcinogenicity study
Male rats

<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Vehicle control</i>	<i>Low dose</i>	<i>Mid dose</i>	<i>High dose</i>
H-POIETIC TUMOUR	HISTIOCYTIC SARCOMA	P-value of test of trend or comparison	.0173		.5050	.1137
		Number of animals reported with tumor	0	0	1	3
		Poly-3 adjusted incidence rate	0.0%	0.0%	1.9%	6.1%
		95% CI for poly-3 adjusted incidence rate (%)	(0,7.1)	(0,7.3)	(0.05,10.4)	(1.28,17.2)
		Poly-3 adjusted number of animals at risk	50.8	49.8	51.3	48.9
LIVER	HEPATOCELLULAR ADENOMA	P-value of test of trend or comparison	.0364	.2424	.2475	.0539
		Number of animals reported with tumor	0	2	2	4
		Poly-3 adjusted incidence rate	0.0%	4.0%	3.9%	8.2%
		95% CI for poly-3 adjusted incidence rate (%)	(0,7.1)	(0.49,14.0)	(0.48,13.7)	(2.27,20.0)
		Poly-3 adjusted number of animals at risk	50.8	49.8	50.8	48.8
LN MESENTERIC	HAEMANGIOMA	P-value of test of trend or comparison	.0040	.6059	.0628	.0163
		Number of animals reported with tumor	6	6	13	15
		Poly-3 adjusted incidence rate	12%	12%	25%	31%
		95% CI for poly-3 adjusted incidence rate (%)	(4.35,23.9)	(4.44,24.3)	(14,39.6)	(18.3,46.3)
		Poly-3 adjusted number of animals at risk	51.5	50.8	51.2	48.9

Table B.15

**Table of tumors reported significant ($\alpha < 0.05$) in at least one arm - Rat Study
NDA 204447
Animal carcinogenicity study
Female rats
Composite endpoints**

<i>Composite endpoint</i>	<i>Quantity</i>	<i>Vehicle control</i>	<i>Low dose</i>	<i>Mid dose</i>	<i>High dose</i>
Hepatocellular tumors	P-value of test of trend or comparison	.0103	.5165	.4884	.0474
	Number of animals reported with tumor	0	1	1	4
	Poly-3 adjusted incidence rate	0.0%	2.1%	2.4%	9.8%
	95% CI for poly-3 adjusted incidence rate (%)	(0,8.0)	(0.05,11.3)	(0.06,12.6)	(2.72,23.7)
	Poly-3 adjusted number of animals at risk	44.6	47.6	42.1	40.6

Table B.16

Table of tumors reported significant ($\alpha < 0.05$) in at least one arm - Rat Study
NDA 204447
Animal carcinogenicity study
Male rats
Composite endpoints

<i>Composite endpoint</i>	<i>Quantity</i>	<i>Vehicle control</i>	<i>Low dose</i>	<i>Mid dose</i>	<i>High dose</i>
Haemangiomas and Haemangiosarcomas	P-value of test of trend or comparison	.0180	.3763	.1059	.0313
	Number of animals reported with tumor	7	9	13	15
	Poly-3 adjusted incidence rate	14%	18%	25%	31%
	95% CI for poly-3 adjusted incidence rate (%)	(5.59,26.3)	(8.4,31.4)	(14,39.6)	(18.3,46.3)
	Poly-3 adjusted number of animals at risk	51.5	50.8	51.2	48.9
Hepatocellular tumors	P-value of test of trend or comparison	.0277	.1175	.2475	.0252
	Number of animals reported with tumor	0	3	2	5
	Poly-3 adjusted incidence rate	0.0%	6.0%	3.9%	10%
	95% CI for poly-3 adjusted incidence rate (%)	(0,7.1)	(1.25,16.9)	(0.48,13.7)	(3.4,22.7)
	Poly-3 adjusted number of animals at risk	50.8	49.8	50.8	48.8

**Tumor counts for organs reported widely under
NDA 204447
Animal carcinogenicity study
Rats**

<i>Species and Sex</i>	<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Vehicle control</i>	<i>Low dose</i>	<i>Mid dose</i>	<i>High dose</i>
Rats - Female	ABDOMEN	HAEMANGIOSARCOMA	Number of tumors found	1	0	0	0
			Number of animals examined	3	0	2	4
	ADIPOSE TISSUE	MESOTHELIOMA	Number of tumors found	0	0	1	1
			Number of animals examined	6	5	5	8
	HEAD	SQUAMOUS CELL CARCINOMA, ZYMBAL'S GLAND	Number of tumors found	0	1	0	0
			Number of animals examined	0	1	1	0
	LN LUMBAR	NEUROENDOCRINE TUMOUR	Number of tumors found	1	0	0	0
			Number of animals examined	2	0	0	0
	PINNAE	HAIR FOLLICLE TUMOUR	Number of tumors found	0	0	1	0
			Number of animals examined	0	0	1	1
		SCHWANNOMA	Number of tumors found	0	0	0	1
			Number of animals examined	0	0	1	1
Rats - Male	ABDOMEN	LEIOMYOSARCOMA	Number of tumors found	0	0	1	0
			Number of animals examined	3	1	3	2
		MALIGNANT SCHWANNOMA	Number of tumors found	1	0	0	0
			Number of animals examined	3	1	3	2
	LN CERVICAL	SCHWANNOMA	Number of tumors found	1	0	0	0
			Number of animals examined	1	0	0	1
	LN DEEP CERVICAL	LIPOSARCOMA	Number of tumors found	0	0	1	0
			Number of animals examined	6	4	3	6
	ORAL CAVITY	SQUAMOUS CELL CARCINOMA	Number of tumors found	1	0	0	0
			Number of animals examined	1	0	0	1

Table B.17

Tumor counts for organs reported widely analyzed
NDA 204447
Animal carcinogenicity study
Rats

<i>Species and Sex</i>	<i>Organ or tissue name</i>	<i>Tumor name</i>	<i>Quantity</i>	<i>Vehicle control</i>	<i>Low dose</i>	<i>Mid dose</i>	<i>High dose</i>
	TAIL	LEIOMYOMA	Number of tumors found	0	1	0	0
			Number of animals examined	0	5	4	4

Table B.17

B.3 Unexamined and autolytic organs

APPEARS THIS WAY ON ORIGINAL



Table B.18

**Organs reported as autolytic
NDA 204447
Animal carcinogenicity study
Female Rats**

<i>Organ or tissue name</i>	<i>Vehicle control(count)</i>	<i>Vehicle control(%)</i>	<i>Low dose(count)</i>	<i>Low dose(%)</i>	<i>Mid dose(count)</i>	<i>Mid dose(%)</i>	<i>High dose(count)</i>	<i>High dose(%)</i>	<i>Total(count)</i>	<i>Total(%)</i>
JEJUNUM	1	1.8%	.	.	2	3.6%	5	9.1%	8	3.6%
RECTUM	3	5.5%	3	1.4%
THYMUS	1	1.8%	1	0.5%

Table B.19

**Organs reported as autolytic
NDA 204447
Animal carcinogenicity study
Male Rats**

<i>Organ or tissue name</i>	<i>Vehicle control(count)</i>	<i>Vehicle control(%)</i>	<i>Low dose(count)</i>	<i>Low dose(%)</i>	<i>Mid dose(count)</i>	<i>Mid dose(%)</i>	<i>High dose(count)</i>	<i>High dose(%)</i>	<i>Total(count)</i>	<i>Total(%)</i>
CAECUM	.	.	1	1.8%	.	.	1	1.8%	2	0.9%
JEJUNUM	1	1.8%	2	3.6%	1	1.8%	2	3.6%	6	2.7%

Table B.20

**Organs reported as unexamined
NDA 204447
Animal carcinogenicity study
Female Rats**

<i>Organ or tissue name</i>	<i>Vehicle control(count)</i>	<i>Vehicle control(%)</i>	<i>Low dose(count)</i>	<i>Low dose(%)</i>	<i>Mid dose(count)</i>	<i>Mid dose(%)</i>	<i>High dose(count)</i>	<i>High dose(%)</i>	<i>Total(count)</i>	<i>Total(%)</i>
ABDOMEN	52	95%	55	100%	53	96%	51	93%	211	96%
ADIPOSE TISSUE	49	89%	50	91%	50	91%	47	85%	196	89%
ADRENALS	1	1.8%	1	0.5%
CLITORAL GLANDS	1	1.8%	.	.	1	1.8%	1	1.8%	3	1.4%
HEAD	55	100%	54	98%	54	98%	55	100%	218	99%
HEART	1	1.8%	1	0.5%
LN LUMBAR	53	96%	55	100%	55	100%	55	100%	218	99%
LN MESENTERIC	1	1.8%	1	0.5%
PARATHYROIDS	3	5.5%	3	5.5%	4	7.3%	3	5.5%	13	5.9%
PINNAE	55	100%	55	100%	54	98%	54	98%	218	99%
THYMUS	.	.	1	1.8%	1	1.8%	1	1.8%	3	1.4%
THYROIDS	1	1.8%	1	0.5%
UTERINE CERVIX	1	1.8%	1	0.5%
VAGINA	1	1.8%	1	0.5%

Table B.21

**Organs reported as unexamined
NDA 204447
Animal carcinogenicity study
Male Rats**

<i>Organ or tissue name</i>	<i>Vehicle control(count)</i>	<i>Vehicle control(%)</i>	<i>Low dose(count)</i>	<i>Low dose(%)</i>	<i>Mid dose(count)</i>	<i>Mid dose(%)</i>	<i>High dose(count)</i>	<i>High dose(%)</i>	<i>Total(count)</i>	<i>Total(%)</i>
ABDOMEN	52	95%	54	98%	52	95%	53	96%	211	96%
LN CERVICAL	54	98%	55	100%	55	100%	54	98%	218	99%
LN DEEP CERVICAL	49	89%	51	93%	52	95%	49	89%	201	91%
LN RENAL	52	95%	55	100%	53	96%	53	96%	213	97%
ORAL CAVITY	54	98%	55	100%	55	100%	54	98%	218	99%
PARATHYROIDS	1	1.8%	3	5.5%	.	.	3	5.5%	7	3.2%
PITUITARY	.	.	1	1.8%	.	.	1	1.8%	2	0.9%
PREPUTIAL GLANDS	1	1.8%	1	0.5%
STOMACH	3	5.5%	.	.	3	1.4%
TAIL	55	100%	50	91%	51	93%	51	93%	207	94%
THYMUS	1	1.8%	3	5.5%	4	1.8%

B.4 Weight changes

Table B.22: Weight changes by group (rats)

Sex	Vehicle control	Lu AA 21004					
		Δ_L	$\frac{\Delta_L}{\Delta_{CP}} - 1$	Δ_M	$\frac{\Delta_M}{\Delta_{CP}} - 1$	Δ_H	$\frac{\Delta_H}{\Delta_{CP}} - 1$
Female	259	282	9%	296	14%	280	8%
Male	456	459	0.7%	504	11%	471	3%

Bibliography

- [1] A J Bailer and C J Portier. Effects of treatment-induced mortality and tumor-induced mortality on tests for carcinogenicity in small samples. *Biometrics*, 44(2):417–31, 1988.
- [2] G S Bieler and R L Williams. Ratio estimates, the delta method, and quantal response tests for increased carcinogenicity. *Biometrics*, 49(3):793–801, 1993.
- [3] KC Chu, C Cueto, and JM Ward. Factors in the evaluation of 200 National Cancer Institute carcinogen bioassays. *Journal of toxicology and environmental health*, 8(1-2):251–280, 1981.
- [4] J K Haseman. A reexamination of false-positive rates carcinogenesis studies. *Fundamental and applied toxicology*, 3(4):334–339, 1983.
- [5] K K Lin and M A Rahman. Overall false positive rates in tests for linear trend in tumor incidence in animal carcinogenicity studies of new drugs. *J Biopharm Stat*, 8(1):1–15; discussion 17–22, 1998.
- [6] R Peto, M C Pike, N E Day, R G Gray, P N Lee, S Parish, J Peto, S Richards, and J Wahrendorf. Guidelines for simple, sensitive significance tests for carcinogenic effects in long-term animal experiments. *IARC Monogr Eval Carcinog Risk Chem Hum Suppl*, NIL(2 Suppl):311–426, 1980.
- [7] M A Rahman and K K Lin. A comparison of false positive rates of peto and poly-3 methods for long-term carcinogenicity data analysis using multiple comparison adjustment method suggested by Lin and Rahman. *J Biopharm Stat*, 18(5):949–58, 2008.

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

/s/

MATTHEW T JACKSON
05/01/2013

KARL K LIN
05/01/2013
Concur with review

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

NDA Number: 204447

Applicant: Takeda GRD

Stamp Date: October 2, 2012

**Drug Name: Vortioxetine
(Lu AA21004)**

NDA/BLA Type: Original NDA

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

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

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? Yes

If the NDA/BLA is not fileable from the statistical perspective, state the reasons and provide comments to be sent to the Applicant.

Please identify and list any potential review issues to be forwarded to the Applicant for the 74-day letter.

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

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

George Kordzakhia	November 13, 2012
Reviewing Statistician	Date
Peiling Yang	November 13, 2012
Supervisor/Team Leader	Date

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

/s/

GEORGE KORDZAKHIA
11/13/2012

PEILING YANG
11/13/2012