CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

21-742

STATISTICAL REVIEW(S)

11/20/07

Statistical Review and Evaluation
(Hormone Study NEB-TX-02)

NDA Number: 21-742

Drug Name: Nebivolol Tablets

Sponsor: Mylan Bertek

Pharm/tox Reviewer: Elizabeth Hausner, D.V.M., Division of Cardiovascular and Renal Products

Project Manager: Daniel Brum, Division of Cardiovascular and Renal Products

Statistical Reviewer: Karl K. Lin, Ph.D. and Atiar Mohammad Rahman, Division of Biometrics 6, Office of Biostatistics

Document Reviewed: Effect of Subcutaneous Dihydrotestosterone (DHT) Administration on Serum Lutenizing Hormone (LH) Levels and Deydig Cell Proliferation Following Gavage Administration of Nebivolol for 28 Days in Mice, Sponsor Study Number: NEB-TX-02, February 13, 2007

Review of Sponsor's Analyses of Serum Luteinizing Hormone (LH) Data of Sponsor Study Number: NEB-TX-02, in Pharm/Tox Section of NDA 21-742

Summary

(Analysis of LH Data) When the Bonnferroni multiplicity adjustment method and the Scheffe multiplicity adjustment method (S method) were applied to multiple statistical tests, the reviewers' analysis results show that all the pairwise comparisons tested except DHT vs Nebivolol_DHT are statistically significant for all the four variables LH_DL, LH_QL, LGLH_DL, and LGLH_QL in the LH dataset submitted by the sponsor on 5/24/2007.

(Analysis of Testosterone Data) Results of two-way analysis of variance of the variables TESTOS and TESTOS (untransformed and transformed testosterone measurement) show that there is not statistically significant differences at 0.05 significant level in testosterone level between the control and the nebivolol groups, and between 4 hour and 6 hour time periods. The interaction of treatment and time period is also not statistically significant at 0.05 level of significance.

(Evaluation of Endpoints (Variables) Used by Sponsor) The results also show that the endpoints used by the sponsor (original LH measurement or logarithmically transformed LH measurement, and original or logarithmically transformed testosterone measurement) do not affect the conclusion of the results. This confirms the well known practicing statistical principle that the assumptions of normality and equal variances of the data in analysis of variance (ANOVA) are robust (i.e., the non-extreme departure of the two assumptions does not have serious effects on the results).

I. Reviewers' Analysis of LH Data

Two-way analysis of variance using treatment and hour as the factors was applied to the variables LH_DL, LH_QL, LGLH_DL, and LGLH_QL in the LH dataset submitted by the sponsor on 5/24/2007. The first two variables are different in the way that not detectable (ND) LH data points for some animals were imputed. The last two variables are the logarithmically (log_{10}) transformed variables of the first two variables.

The variables LH_DL and LH_QL were derived from another two variables MLH_DL and MLH_QL that contain missing (ND) values for some animals. The following examples illustrate the derivation of LH_DL and LH_QL from MLH_DL and MLH_QL.

Animal #	MLH_QL	MLH_DL	LH_DL	LH_QL
1001	0.79	0.79	0.79	0.79
2002	ND	0.07	0.07	0.132
4004	ND	ND	0.07	0.132
2003	ND	0.11	0.11	0.132
4006	ND	0.075	0.079	0.132
2004	ND	ND	0.07	0.132
2010	ND	0.075	0.075	0.132
2017	ND	0.2	0.2	0.2

Contrasts representing various pairwise comparisons of the treatments were tested for each of the four variables of the LH data. More specifically, the contrasts tests were: (1) Control vs DHT, (2) Control vs Nebivolol, (3) Control vs Nebivolol_DHT, (4) DHT vs Nebivolol, (5) DHT vs Nebivolol_DHT, and (6) Nebivolol vs Nebivolol_DHT .

The computer outputs of the analyses of the first two variables are presented in Attachment A, and those of the second two variables are presented in Attachment B.

Two methods of multiplicity adjustment were used in the reviewers' analysis. The two methods are the Bonferroni method and the Scheffe method (S method).

The Bopnnfeeroni adjustment method uses the adjusted level of significance adj α = (overall α) divided by the number of tests performed in the determination if a test is statistically significant. For example, 6 tests were performed for each of the variables in this review, the adjusted level of significance adj α will be 0.05/6 = 0.0083, if the overall α is 0.05. That is, a pairwise comparison of the means of two treatment group is considered as statistically significant if and only if the obtained p-value is less than 0.0083.

The p-values of the tests of contrasts are presented below. A significant test result by the Bonferroni multiplicity adjustment method is indicated by *.

For Variable LH_DL

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
Control vs DHT	1	20.63464267	20.63464267	11.01	0.0011*
Control vs Nebivolol	1	20.25585842	20.25585842	10.81	0.0012*
Control vs Nebivolol_DHT	1	18.54130909	18.54130909	9.90	0.0019*
DHT vs Nebivolol	1	80.66521929	80.66521929	43.05	<.0001*
DHT vs Nebivolol_DHT	1	0.00414397	0.00414397	0.00	0.9625
Nebivolol Nebivolol_DHT	1	73.72019864	73.72019864	39.35	<.0001*

For Variable LH_QL

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
Control vs DHT	1	19.38283320	19.38283320	10.35	0.0015*
Control vs Nebivolol	1	20.25585842	20.25585842	10.82	0.0012*
Control vs Nebivolol_DHT	1	17.53074107	17.53074107	9.36	0.0025*
DHT vs Nebivolol	1	78.20471148	78.20471148	41.76	<.0001*
DHT vs Nebivolol_DHT	1	0.00237412	0.00237412	0.00	0.9716
Nebivolol Nebivolol_DHT	1	71.71647981	71.71647981	38.30	<.0001*

For Variable LGLH_DL

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F

Control vs DHT	1	17.57211938	17.57211938	205.24	<.0001*
Control vs Nebivolol	1	0.97794570	0.97794570	11.42	0.0009*
Control vs Nebivolol_DHT	1	14.92614712	14.92614712	174.33	<.0001*
DHT vs Nebivolol	1	26.25319419	26.25319419	306.63	<.0001*
DHT vs Nebivolol_DHT	1	0.02875209	0.02875209	0.34	0.5628
Nebivolol Nebivolol_DHT	1	22.72512892	22.72512892	265.42	<.0001*

For Variable LGLH_QL

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
Control vs DHT	1	12.96888106	12.96888106	180.90	<.0001*
Control vs Nebivolol	1	0.97794570	0.97794570	13.64	0.0003*
Control vs Nebivolol_DHT	1	11.24286340	11.24286340	156.83	<.0001*
DHT vs Nebivolol	1	20.62200688	20.62200688	287.66	<.0001*
DHT vs Nebivolol_DHT	1	0.01247188	0.01247188	0.17	0.6770
Nebivolol Nebivolol_DHT	1	18.17424791	18.17424791	253.51	<.0001*

When the Bonnferroni multiplicity adjustment was applied, the reviewers' analysis results show that all the pairwise comparisons tested except DHT vs Nebivolol_DHT are statistically significant for all the four variables.

The S-method of adjustment for multiplicity is used by comparing the differences in the mean values of pairs of treatment groups with the value calculated by the following formula:

$$[(k-1)\, F^{\alpha}_{((k-1),\, v)}]^{0.5} \times S\, (2/n)^{0.5}$$

where S2 = MSerror from the analysis of variance table, k is the number of treatment groups, n is the number of observations in each treatment group, and is the 95 percentile value obtained from the F-distribution table with degrees of freedom k and v (the degree of freedom for the error term). In this review, k = 4, and v = 221, and $F^{\alpha}_{((k-1),\, v)} = 2.65$. The S^2 are 1.8736, 1.8726, 0.0856, and 0.0717, respectively, for the variables LH_DL, LH_QL, LGLH_DL, and LGLH_QL.

A pairwise comparison test between the mean values of a pair of treatment groups is considered as statistically significant if and only if the absolute difference of the sample means of the pair of the treatment groups is greater than the S-method critical value calculated by the above formula.

The S-method values for multiplicity adjustment for the above four variables are presented in the following table.

Variable of LH Data	S-Method Value for Multiplicity Adjustment
LH_DL	0.7046
LH_QL	0.7046
LGLH_DL	0.1506
LGLH_QL	0.1379

The group sample means for the four variables of LH are included in Attachment C. The following four tables contain the differences of pairs of the sample means for each of the four variables.

The differences in sample mean between pairs of treatment groups for LH_DL is presented in the following table.

	Gruop	Control	DHT	Nebivolol	Nebivolol_DHT

Group	Mean	0.9610	0.1317	1.7926	0.1425
Control	0.9610	0.0000	-0.8293*	0.8316*	-0.8185*
DHT	0.1317	0.8293*	0.0000	1.6609*	0.0108
Nebivolol	1.7926	-0.8316*	-1.6609*	0.0000	-1.6501*
Nebivolol_DHT	0.1425	0.8185*	-0.0108	1.6501*	0.0000

*: Statistically significant by S-Method by controlling the overall false positive rate at 5%.

The differences in sample mean between pairs of treatment groups for LH_QL is presented in the following table.

	Gruop	Control	DHT	Nebivolol	Nebivolol_DHT
Group	Mean	0.9610	0.1572	1.7926	0.1656
Control	0.9610	0.0000	-0.8038*	0.8316*	-0.7954*
DHT	0.1572	0.8038*	0.0000	1.6354*	0.0084
Nebivolol	1.7926	-0.8316*	-1.6354*	0.0000	-1.6270*
Nebivolol_DHT	0.1656	0.7954*	-0.0084	1.6270*	0.0000

*: Statistically significant by S-Method by controlling the overall false positive rate at 5%.

The differences in sample mean between pairs of treatment groups for LGLH_DL is presented in the following table.

	Gruop	Control	DHT	Nebivolol	Nebivolol_DHT
Group	Mean	0.1592	0.9245	0.0230	0.8964
Control	0.1592	0.0000	-0.7653*	0.1822*	-0.7372*
DHT	0.9245	0.7653*	0.0000	0.9475*	0.0281
Nebivolol	0.0230	-0.1822*	-0.9475*	0.0000	-0.9194*
Nebivolol_DHT	0.8964	0.7372*	-0.0281	0.9194*	0.0000

*: Statistically significant by S-Method by controlling the overall false positive rate at 5%.

The differences in sample mean between pairs of treatment groups for LGLH_QL is presented in the following table.

	Gruop	Control	DHT	Nebivolol	Nebivolol_DHT
Group	Mean	0.1592	0.8167	0.0230	0.7974
Control	0.1592	0.0000	-0.6575*	0.1822*	-0.6382*
DHT	0.8167	0.6575*	0.0000	0.8397*	0.0193
Nebivolol	0.0230	-0.1822*	-0.8397*	0.0000	-0.8204*
Nebivolol_DHT	0.7974	0.6382*	-0.0193	0.8204*	0.0000

*: Statistically significant by S-Method by controlling the overall false positive rate at 5%.

The analysis using the S-method of multiplicity adjustment yielded the same conclusion as that from the analysis using the Bonferronia multiplicity adjustment method, i. e., all the pairwise comparisons tested except DHT vs Nebivolol_DHT are statistically significant for all the four variables.

Main Results of LH Data Analysis

When the Bonnferroni multiplicity adjustment method and the Scheffe multiplicity adjustment method (S method) were applied to multiple statistical tests, the reviewers' analysis results show that all the pairwise comparisons tested except DHT vs Nebivolol_DHT are statistically significant for all the four variables LH_DL, LH_QL, LGLH_DL, and LGLH_QL in the LH dataset submitted by the sponsor on 5/24/2007.

II. Reviewers' Analysis of Testosterone Data

The two-way analysis of variance procedure was applied to the testosterone data also submitted by the sponsor on 5/27/2007. The detailed outputs of the analyses using untransformed and logarithmically transformed data are presented in Attachment D and Attachment E, respectively. The p-values of the tests are presented in the following two tables.

Analysis of Variance Table for Variable TESTOS

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREAT	1	41.74291215	41.74291215	1.53	0.2189
HOUR	1	15.53449991	15.53449991	0.57	0.4522
TREAT*HOUR	1	0.53467766	0.53467766	0.02	0.8889

Analysis of Variance Table for Variable LGTESTOS

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREAT	1	1.16424344	1.16424344	2.37	0.1264
HOUR	1	0.48458836	0.48458836	0.99	0.3225
TREAT*HOUR	1	1.45128633	1.45128633	2.96	0.0884

Results of two-way analysis of variance of the variables TESTOS and TESTOS (untransformed and transformed testosterone measurement) show that there is not statistically significant differences at 0.05 significant level in testosterone level between the control and the nebivolol groups, and between 4 hour and 6 hour time periods. The interaction of treatment and time period is also not statistically significant at 0.05 level of significance.

III. Reviewers' Evaluation of Endpoints (Variables) Used by Sponsor

The results also show that the endpoints used by the sponsor (original LH measurement or logarithmically transformed LH measurement, and original or logarithmically transformed testosterone measurement) do not affect the conclusion of the results. This confirms the well known practicing statistical principle that the assumptions of normality and equal variances of the data in analysis of variance (ANOVA) are robust (i.e., the non-extreme departure of the two assumptions does not have serious effects on the results).

ATTACHMENT A

Results of Analysis of LH Data

(Using Original Untransformed Data)

The GLM Procedure

Class Level Information

Class	Levels	Values
TREATMNT	4	Control DHT Nebivolol Nebivolol_DHT
HOUR	2	4 6

Number of Observations Read	229
Number of Observations Used	229

The GLM Procedure

Dependent Variable: LH_DL LH w/values<DL recoded to 0.070 (ng/ml)

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	107.2075790	15.3153684	8.17	<.0001
Error	221	414.0708494	1.8736238		
Corrected Total	228	521.2784283			

R-Square	Coeff Var	Root MSE	LH_DL Mean
0.205663	178.9654	1.368804	0.764843

Source	DF	Type I SS	Mean Square	F Value	Pr > F
TREATMNT	3	106.7167324	35.5722441	18.99	<.0001
HOUR	1	0.2601704	0.2601704	0.14	0.7098
TREATMNT*HOUR	3	0.2306761	0.0768920	0.04	0.9889

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREATMNT	3	106.5071598	35.5023866	18.95	<.0001
HOUR	1	0.2493808	0.2493808	0.13	0.7156
TREATMNT*HOUR	3	0.2306761	0.0768920	0.04	0.9889

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
Control vs DHT	1	20.63464267	20.63464267	11.01	0.0011
Control vs Nebivolol	1	20.25585842	20.25585842	10.81	0.0012
Control vs Nebivolol_DHT	1	18.54130909	18.54130909	9.90	0.0019
DHT vs Nebivolol	1	80.66521929	80.66521929	43.05	<.0001
DHT vs Nebivolol_DHT	1	0.00414397	0.00414397	0.00	0.9625
Nebivolol Nebivolol_DHT	1	73.72019864	73.72019864	39.35	<.0001

Appears This Way
On Original

The GLM Procedure

Class Level Information

Class	Levels	Values
TREATMNT	4	Control DHT Nebivolol Nebivolol_DHT
HOUR	2	4 6

Number of Observations Read	229
Number of Observations Used	229

Appears This Way
On Original

The GLM Procedure

Dependent Variable: LH_QL LH w/values<QL recoded to 0.132 (ng/ml)

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	103.7970279	14.8281468	7.92	<.0001
Error	221	413.8432378	1.8725938		
Corrected Total	228	517.6402656			

R-Square	Coeff Var	Root MSE	LH_QL Mean
0.200520	176.1673	1.368428	0.776777

Source	DF	Type I SS	Mean Square	F Value	Pr > F
TREATMNT	3	103.3159841	34.4386614	18.39	<.0001
HOUR	1	0.2369146	0.2369146	0.13	0.7224
TREATMNT*HOUR	3	0.2441292	0.0813764	0.04	0.9879

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREATMNT	3	103.1455179	34.3818393	18.36	<.0001
HOUR	1	0.2245187	0.2245187	0.12	0.7295
TREATMNT*HOUR	3	0.2441292	0.0813764	0.04	0.9879

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
Control vs DHT	1	19.38283320	19.38283320	10.35	0.0015
Control vs Nebivolol	1	20.25585842	20.25585842	10.82	0.0012
Control vs Nebivolol_DHT	1	17.53074107	17.53074107	9.36	0.0025
DHT vs Nebivolol	1	78.20471148	78.20471148	41.76	<.0001
DHT vs Nebivolol_DHT	1	0.00237412	0.00237412	0.00	0.9716
Nebivolol Nebivolol_DHT	1	71.71647981	71.71647981	38.30	<.0001

Appears This Way
On Original

ATTACHMENT B

Results of Analysis of LH Data

(Using Logarithmically (log_{10}) Transformed Data)

The GLM Procedure

Class Level Information

Class	Levels	Values
TREATMNT	4	Control DHT Nebivolol Nebivolol_DHT
HOUR	2	4 6

Number of Observations Read 229
Number of Observations Used 229

Appears This Way
On Original

The GLM Procedure

Dependent Variable: LGLH_DL log10(LH) <DL recoded

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	41.64304764	5.94900681	69.48	<.0001
Error	221	18.92173064	0.08561869		
Corrected Total	228	60.56477827			

R-Square	Coeff Var	Root MSE	LGLH_DL Mean
0.687579	-60.73409	0.292607	-0.481783

Source	DF	Type I SS	Mean Square	F Value	Pr > F
TREATMNT	3	41.46877602	13.82292534	161.45	<.0001
HOUR	1	0.10794322	0.10794322	1.26	0.2627
TREATMNT*HOUR	3	0.06632839	0.02210946	0.26	0.8554

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREATMNT	3	41.25069677	13.75023226	160.60	<.0001
HOUR	1	0.11776550	0.11776550	1.38	0.2421
TREATMNT*HOUR	3	0.06632839	0.02210946	0.26	0.8554

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
Control vs DHT	1	17.57211938	17.57211938	205.24	<.0001
Control vs Nebivolol	1	0.97794570	0.97794570	11.42	0.0009
Control vs Nebivolol_DHT	1	14.92614712	14.92614712	174.33	<.0001
DHT vs Nebivolol	1	26.25319419	26.25319419	306.63	<.0001
DHT vs Nebivolol_DHT	1	0.02875209	0.02875209	0.34	0.5628
Nebivolol Nebivolol_DHT	1	22.72512892	22.72512892	265.42	<.0001

Appears This Way
On Original

The GLM Procedure

Class Level Information

Class	Levels	Values
TREATMNT	4	Control DHT Nebivolol Nebivolol_DHT
HOUR	2	4 6

Number of Observations Read	229
Number of Observations Used	229

Appears This Way
On Original

The GLM Procedure

Dependent Variable: LGLH_QL log10(LH) <QL recoded

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	32.14813644	4.59259092	64.06	<.0001
Error	221	15.84337863	0.07168950		
Corrected Total	228	47.99151508			

R-Square	Coeff Var	Root MSE	LGLH_QL Mean
0.669871	-62.11615	0.267749	-0.431046

Source	DF	Type I SS	Mean Square	F Value	Pr > F
TREATMNT	3	32.08756910	10.69585637	149.20	<.0001
HOUR	1	0.05263571	0.05263571	0.73	0.3924
TREATMNT*HOUR	3	0.00793164	0.00264388	0.04	0.9905

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREATMNT	3	31.98518448	10.66172816	148.72	<.0001
HOUR	1	0.05492367	0.05492367	0.77	0.3824
TREATMNT*HOUR	3	0.00793164	0.00264388	0.04	0.9905

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
Control vs DHT	1	12.96888106	12.96888106	180.90	<.0001
Control vs Nebivolol	1	0.97794570	0.97794570	13.64	0.0003
Control vs Nebivolol_DHT	1	11.24286340	11.24286340	156.83	<.0001
DHT vs Nebivolol	1	20.62200688	20.62200688	287.66	<.0001
DHT vs Nebivolol_DHT	1	0.01247188	0.01247188	0.17	0.6770
Nebivolol Nebivolol_DHT	1	18.17424791	18.17424791	253.51	<.0001

Appears This Way
On Original

ATTACHMENT C

Means of Treatment Groups

(4 Hour and 6 Hour Data Combined)

The GLM Procedure

Level of TREATMNT	N	LH_DL Mean	LH_DL Std Dev
Control	60	0.96100000	1.22030796
DHT	60	0.13165000	0.06332371
Nebivolol	57	1.79263158	2.41353144
Nebivolol_DHT	52	0.14250000	0.07107949

The GLM Procedure

Level of TREATMNT	N	LH_QL Mean	LH_QL Std Dev
Control	60	0.96100000	1.22030796
DHT	60	0.15720000	0.04478832
Nebivolol	57	1.79263158	2.41353144
Nebivolol_DHT	52	0.16557692	0.05210899

The GLM Procedure

Level of TREATMNT	N	LGLH_DL Mean	LGLH_DL Std Dev
Control	60	-0.15919962	0.30369198
DHT	60	-0.92453439	0.19313941
Nebivolol	57	0.02298304	0.40529900
Nebivolol_DHT	52	-0.89643037	0.21025973

The GLM Procedure

Level of TREATMNT	N	LGLH_QL Mean	LGLH_QL Std Dev
Control	60	-0.15919962	0.30369198
DHT	60	-0.81669185	0.10102486
Nebivolol	57	0.02298304	0.40529900
Nebivolol_DHT	52	-0.79742284	0.11387320

Appears This Way
On Original

ATTACHMENT D

Results of Analysis of Testosterone Data

(Using Original Untransformed Data)

The SAS System 09:12 Friday, November 16, 2007 17

The GLM Procedure

Class Level Information

Class	Levels	Values
TREAT	2	Control Nebivol
HOUR	2	4 6

Number of Observations Read 107
Number of Observations Used 107

Appears This Way
On Original

The GLM Procedure

Dependent Variable: TESTOS Testosterone (ng/ml)

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	56.289937	18.763312	0.69	0.5615
Error	103	2809.898421	27.280567		
Corrected Total	106	2866.188358			

R-Square	Coeff Var	Root MSE	TESTOS Mean
0.019639	211.4845	5.223080	2.469722

Source	DF	Type I SS	Mean Square	F Value	Pr > F
TREAT	1	40.45361624	40.45361624	1.48	0.2261
HOUR	1	15.30164303	15.30164303	0.56	0.4556
TREAT*HOUR	1	0.53467766	0.53467766	0.02	0.8889

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREAT	1	41.74291215	41.74291215	1.53	0.2189
HOUR	1	15.53449991	15.53449991	0.57	0.4522
TREAT*HOUR	1	0.53467766	0.53467766	0.02	0.8889

Appears This Way
On Original

ATTACHMENT E

Results of Analysis of Testosterone Data

(Using Logarithmically Transformed Data)

The SAS System 09:12 Friday, November 16, 2007 19

The GLM Procedure

Class Level Information

Class	Levels	Values
TREAT	2	Control Nebivol
HOUR	2	4 6

Number of Observations Read 107
Number of Observations Used 107

Appears This Way
On Original

The GLM Procedure

Dependent Variable: LGTESTOS log10(Testosterone)

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	3.00687883	1.00229294	2.04	0.1124
Error	103	50.51427124	0.49042982		
Corrected Total	106	53.52115007			

R-Square	Coeff Var	Root MSE	LGTESTOS Mean
0.056181	-274.9477	0.700307	-0.254705

Source	DF	Type I SS	Mean Square	F Value	Pr > F
TREAT	1	1.14464390	1.14464390	2.33	0.1296
HOUR	1	0.41094860	0.41094860	0.84	0.3621
TREAT*HOUR	1	1.45128633	1.45128633	2.96	0.0884

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREAT	1	1.16424344	1.16424344	2.37	0.1264
HOUR	1	0.48458836	0.48458836	0.99	0.3225
TREAT*HOUR	1	1.45128633	1.45128633	2.96	0.0884

Appears This Way
On Original

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

/s/

Karl Lin
11/20/2007 09:48:41 AM
BIOMETRICS

10/22/07

Statistical Review and Evaluation

(Hormone Studies)

NDA Number: 21-742

Drug Name: Nebivolol Tablets

Sponsor: Mylan Bertek

Pharm/tox Reviewer: Elizabeth Hausner, D.V.M., Division of Cardiovascular and Renal Products

Project Manager: Daniel Brum, Division of Cardiovascular and Renal Products

Statistical Reviewer: Karl K. Lin, Ph.D., Division of Biometrics 6, Office of Biostatistics

Document Reviewed: Interpretive Report, Serum Hormone Analysis in the Mouse and Rat, TOX 021-001, 5-17-2006, in "Nebivolol: A 13-week Endocrine Evaluation Study in Male CD-1 mice and Wistar Rats with a 2-week and 1-month Interim Sacrifice and a 1-month Recovery Period, Final Report, Study No. 04-2875, Sponsor Study No. TOX 021-001, Date: 29 June 2006

Interpretive Report, Serum Hormone Analysis in Mice Administered Nebivolol by Gavage, TOX 021-003, 3-18-2006, In "28-day Toxicity Study of Nebivolol Administered by Oral Gavage to CD-1 Mice with a 14-day Interim Sacrifice to Measure Levels of Luteinizing Hormone and Estradiol", Sponsor Study Number: TOX 021-003, Date of Study Completion: April 26, 2006

Review of Sponsor's Analyses of Serum Luteinizing Hormone (LH) Data of Studies TOX 021-001 and TOX 021-003 in Pharm/Tox Section of NDA 21-742

Summary

The sponsor's first part of analysis using multi-factor ANOVA with the LH in $\log_{10}$ scale as the response variable in both studies; and group, week and group x week interaction as factors in Study TOX 021-001; and group, day, hour(day) (i.e., hour is nested within day), group by day, group by hour(day) as factors in Study TOX 021-003 is the right statistical procedure for this type of data analysis. It is unclear to this reviewer that all analyses other than the full multi-factor ANOVAs performed by the sponsor are necessary. The reviewer gets the impression that the sponsor was doing a lot of data mining trying to find the results it desired to see.

The most serious deficiency in the sponsor data analysis is that the sponsor failed to adjust for the effect of multiple tests to control the overall false positive rate in its statistical analyses. Without adjustment for multiplicity, the overall false positive rate is expected to be much higher than the 5% level of significance used in each of the large number of individual tests performed by the sponsor. Therefore, it is very likely that the sponsor's conclusion of significant drug treatment effect on LH serum concentration is merely based on a false positive finding in its data analysis. The sponsor should use the S-method to test all the contrasts of interest in the sponsor's ANOVA analyses using all factors in each of the two studies, and to adjust the effect of multiple tests. With the adjustment for the effect of multiple tests, one will be more confident to conclude that a statistically significant treatment effect obtained in the data analysis is a true effect not a false positive effect.

In Study TOX 021-001, the sponsor excluded some data points (pulse values) from its statistical analysis. Excluding data with extreme values (or outliers) is not a statistical issue alone. Although it is true that, if extreme values can be justified as from a separate population different from the one being studied, then the exclusion of those values in the data analysis will result in more accurate results. However, the common consensus in the scientific community on this issue is that data with extreme values can not be excluded totally based on statistical tests without other valid justifications. The pharm/tox reviewer should make the determination if it is justified for the sponsor to exclude the data points from the statistical analysis.

0. Introduction

This review and evaluation report presents the results of the quick statistical review of two toxicology (hormone) studies conducted by the drug sponsor to investigate the effect of Nebivolol on Leydig cell hyperplasia in mice and rat.

Elizabeth Hausner, D.V.M., of the Division of Cardiovascular and Renal Products has requested that the Pharm/Tox Statistics Group of the Office of Biostatistics perform a statistical review and evaluation of the various statistical procedures used in the data analysis and the interpretation of results of the hormone studies by the sponsor. Although the request for the statistical review of the hormone studies was officially made by Dr. Hausner on August 9, 2007, due to the confusion caused by the recent reorganization of the Office of Biostatistics, the Pharm/Tox Statistics Group did not receive the request until September 28, 2007.

Because of the delay of the consultation request and the urgency of the medical review division to take actions based on the studies, this statistical review was done in a somewhat rush fashion. It is noted that this report is to serve only as a preliminary statistical review evaluation of the two hormone studies. There could be some other points in the sponsor's reports this reviewer has failed to address in this report. The Pharm/Tox Statistics Group can do a more detailed review including the performance of independent analyses of the hormone data to compare with the sponsor's results if it is needed.

The section below summarizes the designs used by the sponsor in the two hormone studies. Section II summarizes the various statistical methods employed by the sponsor in the data analysis of serum concentration of Luteinizing Hormone (LH) on testis and in the interpretation of study results. The last section contains the reviewer's evaluation of the appropriateness of the sponsor's statistical methods.

I. Study Designs

Study Designs of Serum Hormone Study in Mouse and Rat (TOX 021-001, 5-17-2006, a 13 week study)

Design of the Mouse Study

The following five treatment groups were used in this mouse study:

Vehicle/control
Positive control (finasteride 250/kg)
Nebivolol low dose (10 mg/kg/day)
Nebivolol mid dose (40 mg/kg/day)
Nebivolol high dose (160 mg/kg/day) (the dose was lowered to 80 mg/kg/day for 5 days (Days 15-19), and remained at the dose when it started Day 35, following a 14-day drug holiday.

Serum samples were collected at Weeks 2, 4, 13, and 17.

30 male mice were used for necropsy/hormone evaluation at each time point. Unlike a human subject, repeated LH concentration samples could not be collected over different time points from a test animal with a small body size. Therefore, different groups of animals were used to collect LH concentration samples at different individual time points in each treatment group.

It was noted that only 29 male mice were used at weeks 4, 13, and 17 for some treatment groups; and that only 20 male mice were used at Week 17 for the high dose group.

The details of the mouse study are contained in the following sponsor's experimental outline.

2.3. EXPERIMENTAL OUTLINE (MICE)

The test article (Nebivolol) was administered continuously in the diet to Groups 3 and 4 mice for up to 13 weeks, and to Group 5 mice for 13 weeks (non-continuous) due to a 2-week drug holiday. A positive control substance was administered daily by oral gavage for 13 weeks.

Group	Denomination	Daily Doses[1]			Number of Animals									
					Total		TK[2]	Necropsy/Hormone Evaluations (Toxicity Animals)[3,4]				Microscopic Pathology[5]		Sperm Analysis[6]
		Dose (mg/kg)	Route	Conc. (mg/mL)	Tox	TK	Weeks 2, 4 and 13	Week 2	Week 4	Term Week 13	Rec Week 17	PME/ MP	In situ	Weeks 13 and 17
1	Control (standard diet)	0	NA	NA	120	ND	NA	30	30	30	30	10	20	10
2	Finasteride (Positive Control)[7]	250	Gavage (5 mL/kg)	50	120	ND	NA	30	29	30	29	10	20	10
3	Nebivolol (Low Dose)	10	Dietary	variable	120	72	24	30	30	30	29	10	20	10
4	Nebivolol (Mid Dose)	40	Dietary	variable	120	72	24	30	30	29	29	10	20	10
5	Nebivolol (High Dose)	160/80[8]	Dietary	variable	120	72	24	30	29	30	20	10	20	10
6	Predose (Control for hormone assays)	0	NA	NA	30[9]	ND	NA	NA	NA	NA	NA	NA	NA	NA

[1]Doses represent active ingredient.
[2]Toxicokinetic (TK) samples were collected after 2, 4 and 13 weeks of treatment, at 6 timepoints/occasion. Blood was collected from 4 mice/timepoint as a terminal collection.
[3]Blood for hormone analysis was collected from 30 satellite control mice prior to initiation of treatment, and from 30 mice in each dose group at Weeks 2, 4 and 13, and at Recovery. On each occasion, blood was collected within a 1- or 3-hour time-window (for dietary or positive control animals, respectively). Collection was terminal.
[4]Numbers of animals for necropsy/hormone evaluations reduced due to unscheduled deaths.
[5]Post-mortem examinations and microscopic pathology (PME/MP): Limited postmortem examinations were performed on 10 mice/group at Weeks 2, 4 and 13, and at Recovery. Selected organs (testes, epididymides, prostate, seminal vesicles, and mammary gland) were weighed and preserved for histopathology. Except for one testis and one epididymis from 10 mice of each toxicity group at Week 13, the testes and epididymides of all remaining mice were removed and preserved. All carcasses were preserved whole.
[6]Sperm analyses: One testis and one epididymis from 10 mice of each main group at Weeks 13 and 17 necropsy was utilized for sperm analyses (sperm count, sperm motility and sperm morphology).
[7]Finasteride was administered by oral gavage, in 0.5% methylcellulose.
[8]The dose level for Group 5 was lowered to 80 mg/kg/day for 5 days (Days 15-19), and remained at 80 mg/kg/day when dose restarted Day 35, following a 14-day drug holiday.
[9]Group 5 animals were sacrificed at Week 8 (relative to restart date) rather than Week 4.
The first day of dosing was defined as Day 0 of the study. Conc. = Concentration; NA = not applicable; ND = not done; Tox = toxicity; Rec = Recovery

Design of the Rat Study

The number of treatment groups, the doses, and the time points of necropsy/hormone evaluation in this rat study are the same as those used in the above mouse study except that 100/kg finasteride was used for the positive control group. However, the factors described below for the rat study are not the same as the mouse study.

20 instead of 30 male rats used for necropsy/hormone evaluation at each time point in the rat study.

'The dose level for Group 5 was lowered to 80 mg/kg/day for 9 days (Days 21-29), and remained at 80 mg/kg/day when dose restarted Day 42, following a 12-day drug holiday.

The details of the mouse study are contained in the following sponsor's experimental outline.

2.4. EXPERIMENTAL OUTLINE (RATS)

The test article (Nebivolol) was administered continuously in the diet to Groups 3 and 4 rats for up to 13 weeks, and to Group 5 rats for 11 weeks (non-continuous) due to a 2-week drug holiday. A positive control substance was administered daily by oral gavage for 13 weeks.

Group	Denomination	Daily Doses[a]			Number of Animals									
					Toxic		TK[c]	Necropsy/Hormone Evaluations (Toxicity Animals)[d,e]				Microscopic Pathology[f]		Sperm Analysis[g]
		Dose (mg/kg)	Route	Conc. (mg/mL)	Tox	TK	Weeks 2, 4 and 13	Week 2	Week 4	Term Week 13[h]	Rec Week 17[i]	PME/ MP	In situ	Weeks 13 and 17
1	Control (standard diet)	0	NA	NA	80	ND	NA	20	20	20	20	10	10	10
2	Flutamide (Positive Control)[j]	100	Gavage (5 mL/kg)	20	80	ND	NA	20	20	20	20	10	10	10
3	Nebivolol (Low Dose)	10	Dietary	variable	80	24	24	20	20	20	20	10	10	10
4	Nebivolol (Mid Dose)	40	Dietary	variable	80	24	24	20	20	20	20	10	10	10
5	Nebivolol (High Dose)	160/80[k]	Dietary	variable	80	24	24	20	20	20	20	10	10	10
6	Predose (Control for hormone assays)	0	NA	NA	20[l]	ND	NA	NA	NA	NA	NA	NA	NA	NA

[a]Doses represent active ingredient.
[b]Toxicokinetic (TK) samples were collected after 2, 4 and 13 weeks of treatment, at 6 timepoints/occasion. Blood was collected from 4 rats/timepoint.
[c]Blood for hormone analysis was collected from 20 satellite control rats prior to initiation of treatment, and from 20 rats in each dose group at Weeks 2, 4 and 13, and at Recovery. On each occasion, blood was collected within a 1- or 3-hour time-window (for dietary or positive control animals, respectively). Collection was terminal.
[d]Numbers of animals for necropsy/hormone evaluations reduced due to unscheduled deaths.
[e]Post-mortem examinations and microscopic pathology (PME/MP): Limited postmortem examinations were performed on 10 rats/group at Weeks 2, 4 and 13, and at Recovery. Selected organs (testes, epididymides, prostate, seminal vesicles, and mammary gland) were weighed and preserved for histopathology. Except for one testis and one epididymis from 10 rats of each main group at Week 13, the testes and epididymides of all remaining rats were removed and preserved. All carcasses were preserved whole.
[f]Sperm analyses: One testis and one epididymis from 10 rats of each tox group at Weeks 13 and 17 necropsy were utilized for sperm analyses (sperm count, sperm motility and sperm morphology).
[g]The terminal sacrifice was at Week 7 for Group 5 relative to the restart date.
[h]The recovery sacrifice was at Week 11 for Group 5 relative to the restart date.
[i]Flutamide was administered by oral gavage, in 0.5% methylcellulose.
[j]The dose level for Group 5 was lowered to 80 mg/kg/day for 9 days (Days 21-29), and remained at 80 mg/kg/day when dose restarted Day 42, following a 12-day drug holiday.
The first day of dosing was defined as Day 0 of the study. Conc. = Concentration; NA = not applicable; ND = not done; Tox = toxicity; Rec = Recovery

Study Design of the Serum Hormone Study in Mice Administered Nibivolol by Garage Mouse TOX 021-003, 3-18-2006

Group: Vehicle/control
Nebivolol low dose (5 mg/kg/day)
Nebivolol mid dose (20 mg/kg/day)
Nebivolol high dose (80 mg/kg/day)

Serum samples were collected at 0.5, 1, 2, 3, 4, 6, and 8 hours on Days 14 and 28.

30 male mice were used for each treatment group at each time point on each day for necropsy/hormone evaluation. The total number of animals used in this study is 420 male mice.

II. Sponsor's Analysis and Interpretation Methods

The sponsor and its consultant used a similar type of methods in their analyses of the data (serum concentration of LH) and interpretations of the study results in TOX 021-

001Study (two species, mice and rats), and TOX 021-003 Study (mice only). Their methods are described in this section.

A multi-factor analysis of variance (ANOVA) was run with the LH in both original scale and $\log_{10}$ scale as the response variables in both studies; and group, week and group by week interaction as factors in Study TOX 021-001; and group, day, hour(day) (i.e., hour is nested within day), group by day, group by hour(day) as factors in Study TOX 021-003. In both studies, separate one-way and/or two-way ANOVAs were run for each or two of the factors week, day, and hour. Dunnett's test was used to determine if any of the treatment groups were significantly different from control.

In Study TOX 021-001, further analysis was carried out using Tukey's normal scores on the ranked response data (LH) to ensure that extreme values would not have undue impact on results and to stabilize within-group variances.

In Study TOX 021-001, additional analyses were performed on the data after identifying pulse values. For each group and week, an iterative program was run to identify values greater than mean+ 2 SD. A categorical analysis was run, with the larger values identified as pulse and the smaller values identified as baseline. In an additional analysis, the pulse values were deleted from analysis and a series of one-way ANOVAs (by week, with group in the model) was performed on the remaining data. The analyses were repeated using the normal scores.

The sponsor also hired a consulting statistician to analyze the LH serum concentration data. In addition to the ANOVAs on $\log_{10}$ scale data, the sponsor's consultant also ranked the samples of the treatment groups, and performed a statistical analysis on the ranks of the samples (i.e., performed a nonparametric analysis).

Nonparametric analysis was performed to compare treatments, separately for each week in Study 021-001, and each hour/day in Study Tox 021-003. Nonparametric analysis was done in SAS Proc Nparl way using the Wilcoxon Scores (the ranks) and the Van der Waerden Scores (normal scores).

In Study 021-003, the LH serum concentration data were also analyzed by the sponsor using the logistic regression method to estimate the probability that the response exceeded the threshold of 3 using the factors used in the ANOVAs as the independent variables.

Since the hypothesis being tested in these experiments was that Nebivolol increased serum concentrations of LH which in turn may lead to an increase in the incidence of Leydig cell tumors, it was determined that a one-sided statistical analysis of the data should be performed.

III. Reviewer's Evaluation of Sponsor's Analysis and Interpretation Methods

III.a Evaluation of Sponsor's Methods of Data Analysis

It is the reviewer's opinion that the sponsor's first part of analysis using multi-factor ANOVA with the LH in $\log_{10}$ scale as the response variable in both studies; and group, week and group x week interaction as factors in Study TOX 021-001; and group, day, hour(day) (i.e., hour is nested within day), group by day, group by hour(day) as factors in Study TOX 021-003 is the right statistical procedure for this type of data analysis.

However, the sponsor performed a large number of additional analyses using the ANOVA model but with subsets of factors, other parametric and non-parametric methods to support its final finding and conclusion because it felt that the LH serum concentration data in original scale and $\log_{10}$ scale were not normally distributed with unequal variances (two of the three assumptions used in ANOVA method).

It has been shown in literature that the assumptions of normality and of equal variances for ANOVA model are fairly robust (i.e., unless with extreme departures from the assumptions, the violations of the two assumptions do not have serious effects on the analysis results). It is the reviewer's opinion that the multi-factor ANOVA model including all factors in each of the studies and using the serum concentration data in $\log_{10}$ scale as the response variable is appropriate. Since the hour is nested in day in Study TOX 021-003, it may be also a reasonable alternative way, as also done by the sponsor, to perform two separate multi-factor ANOVAs for the two separate experiment days. However, this approach may lead to be a problem in interpreting the overall result if one ANOVA on the data of one day shows a significant dose-trend while the ANOVA on the data of the other day shows a non-significant dose-trend.

It is unclear to this reviewer that all analyses other than the full multi-factor ANOVAs performed by the sponsor are necessary. The reviewer gets the impression that the sponsor was doing a lot of data mining trying to find the results it desired to see.

III.b Need to Adjust for the Effect of Multiple Tests

To this reviewer, the most serious deficiency in the sponsor data analysis is that the sponsor failed to adjust for the effect of multiple tests to control the overall false positive rate in its statistical analyses. Although the sponsor did not indicate the possible level of overall false positive rate in its conclusion of a significant treatment effect (dose-trend and pairwise difference) on the LH serum level in the test animals, without adjustment for multiplicity, the overall false positive rate is expected to be much higher than the 5% level of significance used in each of the large number of individual tests performed by the sponsor. Therefore, it is very likely that the sponsor's conclusion of significant drug treatment effect on LH serum concentration is merely based on a false positive finding in its data analysis.

It is reviewer's opinion that the sponsor should use the S-method (one of proposed methods for doing adjustment for effect of multiple tests for ANOVAs described in, for

example, the classical book The Analysis of Variance, by Henry Scheffe, John Wiley & Sons, 1959) to test all the contrasts of interest in the sponsor ANOVA analyses using all factors in each of the two studies, and to adjust the effect of multiple tests. The S-method and other multiplicity adjustment methods control the overall false positive rate at a desired and pre-specified level determined by the investigator. With the adjustment for the effect of multiple tests, one will be more confident to conclude that a statistically significant treatment effect obtained in the data analysis is a true effect not a false positive effect.

The purpose of the following paragraphs in this subsection is to explain the importance to adjust the effect of multiple tests in the final interpretation of results of a study. Interpreting results of a drug effect experiment using animal or human subjects is a complex process, and there are risks of both false negative and false positive results. The relatively small number of subjects used and the large variability of the study endpoint in the study population can result in the failure to detect the true effect of a drug (i.e., a false negative). Because of the large number of comparisons involved in the data analysis, a great potential exists for finding statistically significant positive trends or treatment-placebo differences due to chance alone (i.e., a false positive).

The sponsor realized the large variability in LH concentration in the animal population (a major risk of producing false negative results), and used large number of animals (30 mice or 20 rats at each test time points for each treatment group) in the two studies to reduce the false negative rate.

However, as mentioned above, the issue of controlling false positive results was not addressed in the statistical analyses using various methods performed by the sponsor and its consulting statistician. The illustrations below show the importance of the adjustment for multiple tests in the final interpretation of study results. Without the adjustment for multiplicity, the overall false positive rate is very high, and very likely a significant finding is merely a false positive and not a true effect.

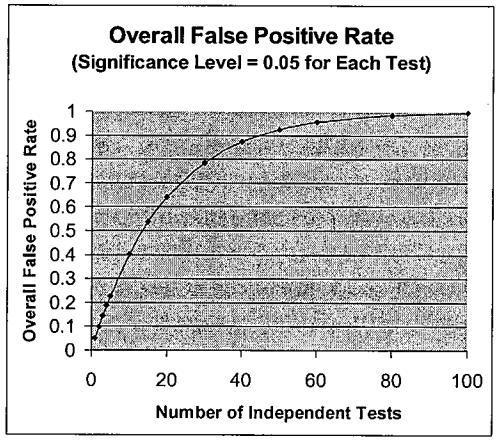
As it can be seen from the included table and graphs, the overall false positive rate increases dramatically as the number of statistical tests performed increases. For example, if one performs 10 individually independent tests each at 0.05 significance level without considering the multiplicity issues, there is a 40% chance to get at least a false positive result. The chance increases to 64% when the number of independent tests increases to 20.

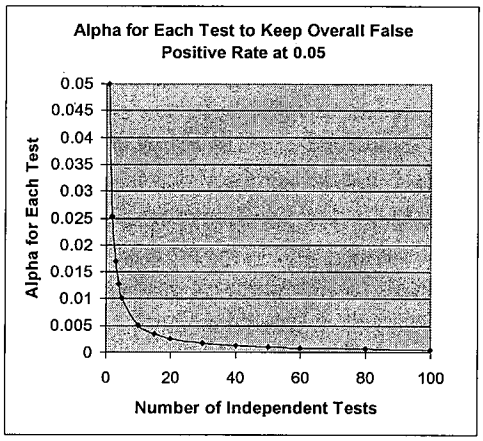
It is important to adjust for the effect of multiple tests in the final interpretation of test results if the investigator can not determine before the start of an experiment the hypotheses he or she wants to test, but just by performing various tests on the data in order to find some effects or non-effects after the experiment is completed.

Probability of Small P Value (less than 0.05 in this example) When Testing Many Null Hypotheses

Number of Independent Null Hypotheses(N)	2	3	4	5	6	7	8	9	10	20	50	100
Probability (P*) of obtaining one or more P values less than 0.05 by chance	10%	14%	19%	23%	26%	30%	34%	37%	40%	64%	92%	99%
Alpha* to keep overall risk of type I error equal to 0.05 .	0253	.0170	.0127	.0102	.0085	.0073	.0064	.0057	.0051	.0026	.0010	.0005

Note: $P^* = 100(1.00 - 0.95^N)$ and $Alpha^* = 1.00 - 0.95^{1/N}$.





(Source of the above illustrations:: Chapter 13: Multiple Comparisons, in Intuitive Biostatistics (ISBN 0-19-508607-4) by Harvey Motulsky. Copyright © 1995 by Oxford University Press Inc. http://www.graphpad.com/www/Book/mulcomp.htm)

There are various methods having proposed for multiplicity adjustment to control the overall false positive rate at a desired level. For example, in the above simple illustrations, each of the 10 and 20 statistical tests should be tested at 0.0051 and 0.0026 significance levels, respectively, instead of 0.05 if one wants to control the overall false positive rate at 0.05.

III.c. Exclusion of Data in Sponsor's Analysis

In Study TOX 021-001, the sponsor excluded some data points (pulse values) from its statistical analysis by performing the following procedure to identify them in the serum concentration data. For each group and week, an iterative program was run to identify values greater than mean+ 2 SD. A categorical analysis was run, with the larger values identified as pulse and the smaller values identified as baseline. In an additional analysis, the pulse values were deleted from analysis and a series of one-way ANOVAs (by week, with group in the model) was performed on the remaining data.

Excluding data with extreme values (or outliers) is not a statistical issue alone. There are statistical procedures to test if an extreme value can be considered as an outlier. However, the common consensus in the scientific community on this issue is that data with extreme values can not be excluded totally based on statistical tests without other valid justifications such as coding errors, subjects did not meet the selection criteria, etc. since one can keep excluding the data he or she considered as outliers in the analysis until a desired result is obtained. The possible result of excluding outliers is to make the data less variable, that, in turn, will make a test statistic calculated from the data larger, and will result in a statistically significant finding. However, it is also true that, if extreme values can be justified as from a separate population different from the one being studied, then the exclusion of those values in the data analysis will result in more accurate results.

IV. Concluding Remarks

The sponsor's first part of analysis using multi-factor ANOVA with the LH in $\log_{10}$ scale as the response variable in both studies; and group, week and group x week interaction as factors in Study TOX 021-001; and group, day, hour(day) (i.e., hour is nested within day), group by day, group by hour(day) as factors in Study TOX 021-003 is the right statistical procedure for this type of data analysis. It is unclear to this reviewer that all analyses other than the full multi-factor ANOVAs performed by the sponsor are necessary. The reviewer gets the impression that the sponsor was doing a lot of data mining trying to find the results it desired to see.

The most serious deficiency in the sponsor data analysis is that the sponsor failed to adjust for the effect of multiple tests to control the overall false positive rate in its statistical analyses. Without adjustment for multiplicity, the overall false positive rate is expected to be much higher than the 5% level of significance used in each of the large number of individual tests performed by the sponsor. Therefore, it is very likely that the sponsor's conclusion of significant drug treatment effect on LH serum concentration is merely based on a false positive finding in its data analysis. The sponsor should use the S-method to test all the contrasts of interest in the sponsor's ANOVA analyses using all factors in each of the two studies, and to adjust the effect of multiple tests. With the adjustment for the effect of multiple tests, one will be more confident to conclude that a

statistically significant treatment effect obtained in the data analysis is a true effect not a false positive effect.

In Study TOX 021-001, the sponsor excluded some data points (pulse values) from its statistical analysis. Excluding data with extreme values (or outliers) is not a statistical issue alone. Although it is true that, if extreme values can be justified as from a separate population different from the one being studied, then the exclusion of those values in the data analysis will result in more accurate results. However, the common consensus in the scientific community on this issue is that data with extreme values can not be excluded totally based on statistical tests without other valid justifications. The pharm/tox reviewer should make the determination if it is justified for the sponsor to exclude the data points from the statistical analysis.

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

/s/

Karl Lin
10/22/2007 08:23:54 AM
BIOMETRICS

Memorandum of Statistical Consult

Date: October 9, 2007

Re: NDA 21-742 (Serial 000, dated May 30, 2007)
Sponsor: Mylan Bertek Pharmaceuticals Inc.
Product: [redacted] (Nebivolol Hydrochloride)
Indication: Management of hypertension

This statistical consult is part of a DRUP clinical consult from the Division of Cardio-Renal Drug Products for nebivolol hydrochloride, a selective β1 blockade for management of hypertension, which is currently under NDA review. The DRUP clinical reviewer is Dr. Harry Handelsman.

Study NEB-PK-03 was a randomized, double-blind, placebo- and active-controlled parallel-group study in approximately 120 healthy male subjects ages 18 to 50 years comparing nebivolol to placebo in the primary pharmacodynamic endpoint of area under the curve (AUC) from time zero to 120 minutes of ACTH-stimulated serum cortisol levels on Day 56. The study consisted of one week of single-blind, matching-placebo run-in, then one week of double-blind low dose treatment (nebivolol 5 mg/day or atenolol 50 mg/day or placebo), followed by six weeks of double-blind high dose treatment (nebivolol 10 mg/day or atenolol 100 mg/day or placebo).

One of the secondary pharmacodynamic endpoints was the mean level of total testosterone (ng/dL) on Day 56. The comparison of nebivolol to placebo in this secondary endpoint was analyzed using an Analysis of Covariance (ANCOVA) model, with treatment group, metabolic status (CYP2D6 extensive metabolizers vs. CYP2D6 poor metabolizers), and study center as factors and the corresponding baseline $AUC_{0\text{-}120\ min}$ value as covariate. This analysis was the based on the ITT population, defined as all subjects who had data for Days 1, 7, and 56.

The study sample size of 30 subjects per treatment group was based on a Coefficient of Variability (CV, defined as the pooled standard deviation divided by the placebo mean) of 26% (based on the estimated CV in Study NEB-BEL-55), a true treatment difference between nebivolol and placebo of zero in the primary pharmacodynamic parameter, approximately 90% power, α-level of 0.05, one-sided t-test, and ruling out a 20% reduction in the primary pharmacodynamic parameter for nebivolol subjects relative to the placebo subjects.

The DRUP clinical reviewer requested that analyses on the difference in the change from baseline in total testosterone level between nebivolol and placebo, sample size, and power be performed. These analyses are used to determine if the study is adequate to detect a clinically meaningful difference between nebivolol and placebo in the change from baseline. This request was initiated due to the different withdrawal rates between nebivolol and placebo. The placebo group contained 52 randomized subjects, of which 4 (7.7%) withdrew from study with three due to protocol-specified cardiovascular safety criteria. The nebivolol group contained 55 randomized subjects, of which 13 (23.6%) withdrew from study with eight due to protocol-specified cardiovascular safety criteria. General results are presented and the clinical reviewer will determine the range of clinically meaningful differences in change from baseline in total testosterone level for nebivolol.

Table 1 presents the unadjusted change from baseline in total testosterone level (ng/dL) for each treatment group. The placebo group had a 2.9 ng/dL mean decrease from baseline (-0.53% change) compared to a 26.6 ng/dL mean increase from baseline (4.52% change) for the nebivolol group. Neither treatment group had a significantly different change from baseline (p-value>0.10).

Table 1
Study NEB-PK-03: Change from Baseline in Total Testosterone Level (ng/dL) for the ITT Population* by Treatment Group

	Day 56 (s.e.)	Baseline (s.e.)	Change from Baseline (s.e.)	Percent Change from Baseline	p-value
Placebo (n=48)	549.0 (18.8)	551.9 (20.0)	-2.9 (14.3)	-0.53%	0.841
Nebivolol (n=42)	588.4 (25.8)	561.7 (24.1)	26.6 (16.5)	4.52%	0.114

Source: Statistical Reviewer's calculations based on dataset D_PD.xpt, which is located in the EDR at \\Cdsesub1\N21742\N_000\2007-04-27\N21742\crt\datasets\neb-pk-03.
* ITT Population includes all subjects who had data for Days 1, 7 and 56.

Table 2 presents the adjusted change from baseline in total testosterone level (ng/dL) for each treatment group and comparison of nebivolol to placebo based on an ANCOVA model with treatment group, study center, and CYP 2D6 metabolic status as fixed factors and baseline total testosterone level as covariate. The placebo group had a 36.94 ng/dL mean increase from baseline compared to a 62.98 ng/dL mean increase from baseline for the nebivolol group. The adjusted mean nebivolol difference compared to placebo difference was -26.05 ng/dL and was not significant (p=0.223).

Table 2
Study NEB-PK-03: Adjusted† Change in Total Testosterone Level at Day 56 for the ITT Population

	Nebivolol	Placebo
n	42	48
Baseline mean (Day 7)	561.7	551.9
Day 56		
Adjusted mean change from baseline	62.98	36.94
Adjusted Mean Treatment Difference vs. Placebo (s.e.)	-26.05 (21.2)	
p-value for Adjusted Treatment Difference	0.223	

Source: Statistical Reviewer's calculations based on datasets D_PD.xpt and D_PROF.xpt, which are located in the EDR at \\Cdsesub1\N21742\N_000\2007-04-27\N21742\crt\datasets\neb-pk-03.
* ITT Population includes all subjects who had data for Days 1, 7 and 56.
† Adjusted results based on the analysis of covariance model: change from baseline = treatment group + study center + CYP 2D6 metabolic status + total testosterone level at baseline

Table 3 presents results assessing the power for varying changes from baseline in total testosterone level for nebivolol compared to placebo while holding the change from baseline in total testosterone level for placebo constant. Both treatment groups' baseline data are the observed study values and other assumptions used in these calculations are listed in the table's subtext. For example, the highlighted column presents the power for this study's nebivolol change from baseline value compared to the placebo group's change from baseline value of 29.5 ng/dL (based on Table 1 data) at 38%.

Table 3
Study NEB-PK-03: Power for Varying Changes from Baseline in Total Testosterone Level (ng/dL) for Nebivolol Compared to Placebo While Holding the Change from Baseline in Total Testosterone Level (ng/dL) for Placebo Constant*

Percent Change from baseline for Nebivolol group	-90%	-80%	-70%	-60%	-50%	-40%	-30%	-20%	-10%	-4.52%
Change from baseline for Nebivolol group	-505.53	-449.36	-393.19	-337.02	-280.85	-224.68	-168.51	-112.34	-56.17	-26.64
Power (%)	***99***	***99***	***99***	***99***	***99***	***99***	***99***	***99***	***78***	***28***

Percent Change from baseline for Nebivolol group	-4.52%	10%	20%	30%	40%	50%	60%	70%	80%	90%
Change from baseline for Nebivolol group	-26.64	56.17	112.34	168.51	224.68	280.85	337.02	393.19	449.36	505.53
Power (%)	***38***	***84***	***99***	***99***	***99***	***99***	***99***	***99***	***99***	***99***

Source: Statistical Reviewer's calculations.
* Assumptions used are as follow: Nebivolol baseline total testosterone level = 561.7, standard deviation for Nebivolol change from baseline = 106.873, Nebivolol n = 42, Change from baseline for Placebo = -2.895, standard deviation for Placebo change from baseline = 99.323, Placebo n = 48, unequal variances, 1-sided t-test, α-level 0.05

Table 4 presents results assessing the sample size needed to detect a significant difference for varying changes from baseline in total testosterone level for nebivolol compared to placebo while holding the change from baseline in total testosterone level for placebo constant. Both treatment groups' baseline data are the observed study values and other assumptions used in these calculations are listed in the table's subtext. For example, the highlighted column presents the sample size per group needed to show a significant difference for this study's nebivolol change from baseline value compared to the placebo group's change from baseline value of 29.5 ng/dL (based on Table 1 data) as 210 subjects per group.

Table 4

Study NEB-PK-03: Sample Size Needed to Detect a Significant Difference for Varying Changes from Baseline in Total Testosterone Level (ng/dL) for Nebivolol Compared to Placebo While Holding the Change from Baseline in Total Testosterone Level (ng/dL) for Placebo Constant*

Percent Change from baseline for Nebivolol group	-90%	-80%	-70%	-60%	-50%	-40%	-30%	-20%	-10%	-4.52%
Change from baseline for Nebivolol group	-505.53	-449.36	-393.19	-337.02	-280.85	-224.68	-168.51	-112.34	-56.17	-26.64
N per group	***3***	***3***	***3***	***3***	***4***	***5***	***8***	***16***	***65***	***325***
Percent Change from baseline for Nebivolol group	-4.52%	10%	20%	30%	40%	50%	60%	70%	80%	90%
Change from baseline for Nebivolol group	26.64	56.17	112.34	168.51	224.68	280.85	337.02	393.19	449.36	505.53
N per group	***210***	***53***	***15***	***8***	***5***	***4***	***3***	***3***	***3***	***3***

Source: Statistical Reviewer's calculations.

* Assumptions used are as follow: Nebivolol baseline total testosterone level = 561.7, standard deviation for Nebivolol change from baseline = 106.873, Change from baseline for Placebo = -2.895, standard deviation for Placebo change from baseline = 99.323, unequal variances, 1-sided t-test, α-level 0.05, power=90%

Sonia Castillo, Ph.D.
Mathematical Statistician
Division of Biometrics 3

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

/s/

Sonia Castillo
10/10/2007 10:17:17 AM
BIOMETRICS

12/16/04

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Pharmacoepidemiology and Statistical Science
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Serial Number:21-742

Drug Name: Nebivolol

Indication(s): Hypertension

Applicant: Bertek Pharmaceuticals, Inc.

Date(s): 4/30/2004

Review Priority: Standard

Biometrics Division: Division of Biometrics I

Statistical Reviewer: Jasmine Choi, M.S. (HFD-710)

Concurring Reviewers: H.M. James Hung, Ph.D. (HFD-710)

Medical Division: Division of Cardio-Renal Drug Product (HFD-110)

Clinical Team: Karen Hicks, M.D. (HFD-110)

Project Manager: Melissa Robb (HFD-110)

Keywords:

Analysis of covariance, closed-test procedures

Table of Contents

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

Nebivolol had a statistically significant effect on reducing sitting diastolic blood pressure (DBP) not only in non-black patients but also in black patients. The secondary analyses on other efficacy measurements confirmed that nebivolol had a statistically significant antihypertensive effect on mild to moderate hypertension population.

1.2 Brief Overview of Clinical Studies

Three randomized, double-blind, multi-center, placebo-controlled trials were conducted to study the efficacy of nebivolol monotherapy for treatment of mild to moderate hypertension. NBE-302 and NEB-305 studies were conducted for general population, and NEB-202 study was conducted for black population. Treatment was administered once daily for 12 weeks, and patients were stratified by metabolism of nebivolol, diabetes status, race (only in NEB-302 and NEB-305), age, and gender. The dose range was from 1.25mg to 40mg of nebivolol. NEB-302 included all of the doses, NEB-202 included all doses except for the lowest dose of 1.25mg, and NEB-305 included only 5mg, 10mg and 20mg doses. The primary efficacy endpoint was the change in mean sitting diastolic blood pressure (DBP) at trough at end of treatment from baseline. NEB-302 and NEB-202 were conducted in the US, and NEB-305 was conducted in U.S. and Europe.

1.3 Statistical Issues and Findings

The primary analyses were conducted using Analysis of Covariance (ANCOVA) model with treatment as a factor and baseline sitting DBP, age, gender, race, diabetes status, and metabolism of nebivolol as covariates. The step-down trend test on LS mean changes for the general population showed that the sitting DBP of all dosed groups was significantly decreased compared to the placebo groups (NEB-302, $p<0.0001$ for all doses; NEB-305, $p<0.0015$). The same analysis on NEB-202 showed a statistically significant reduction of sitting DBP in all dose groups except 2.5mg dose group. The following table shows the results of the analyses.

Table 1: Primary Analysis Results of Pivotal Studies
(Source: Reviewer's analysis)

Treatment	N	Baseline Mean	Mean at the End of Study	LS mean Change from Baseline	Step-DownTrend Test p-value
NEB-302					
Placebo	81	100.3	97.1	-2.9 (1.1)	-
1.25mg	83	98.9	90.8	-8.0 (1.1)	<0.0001
2.5mg	82	99.8	91.1	-8.5 (1.1)	<0.0001
5mg	165	99.6	91.0	-8.4 (1.0)	<0.0001
10mg	166	99.5	90.2	-9.2 (0.9)	<0.0001
20mg	166	99.4	89.5	-9.8 (0.9)	<0.0001
30/40mg	166	99.3	88.0	-11.2 (0.9)	<0.0001
NEB-305					
Placebo	75	98.7	91.4	-4.6 (1.3)	-
5mg	244	99.1	88.5	-7.8 (1.0)	0.0015
10mg	244	98.9	87.7	-8.5 (1.0)	0.0009
20mg	244	99.2	87.2	-9.1 (1.0)	0.0002
NEB-202					
Placebo	49	100.8	96.4	-2.8 (2.1)	-
2.5mg	49	99.5	92.8	-5.7 (2.1)	0.14
5mg	50	100.5	91.4	-7.7 (2.1)	0.0187
10mg	51	100.3	90.0	-8.9 (2.0)	0.0032
20mg	50	101.5	90.9	-8.9 (2.1)	0.0019
40mg	51	98.7	89.6	-8.3 (2.0)	0.0014

For the secondary analyses, change of sitting systolic blood pressure (SBP) at trough and rates of responder, which was defined as a patient whose average sitting DBP at trough was either <90mmHg at end of treatment or had decreased by ≥10mmHg from baseline, were analyzed. The results of these secondary analyses confirmed the findings from the primary analyses.

2. INTRODUCTION

2.1 Overview

Nebivolol is a highly cardioselective β-adrenergic receptor blocker for oral administration that resulting pharmacologic profile differs from classic beta-blockers such as propranolol, atenolol, or metoprolol. The sponsor's clinical program was to show that nebivolol administered once daily for the treatment of hypertension is efficacious regardless of age, race, gender, oxidative genotype, diabetes status, or BMI. In addition, its efficacy in black hypertensive patients is as good as in non-black patients.

The efficacy of nebivolol monotherapy for treatment of hypertension was studied in three randomized, double-blind, multi-center, placebo-controlled trials (NEB-302, NEB-305, and NEB-202). Treatment was administered once daily for 12 weeks, and patients were stratified by metabolism of nebivolol, diabetes status, race (only in NEB-302 and NEB-305), age, and gender. The primary efficacy variable was the change in mean sitting DBP at trough at end of treatment compared to baseline. The dose range was evaluated ranging from 1.25mg to 40mg of nebivolol. NEB 302 included all of the doses in the range, NEB-202 included all doses except for the lowest dose of 1.25mg, and NEB-305 included only the 5mg, 10mg, and 20mg doses. The three efficacy studies were intended to demonstrate not only efficacy in the general hypertensive population in the US, but also effectiveness among the blacks, a population traditionally shown to be less responsive to beta-blocker therapy than Caucasians. Study NEB-302, conducted entirely in the US, and study NEB-305, conducted in the US and Europe, enrolled hypertensives of any race; whereas, study NEB-202 enrolled only black hypertensive patients in the US.

2.2 Data Sources

Data used for review were from the electronic submission received on 04/30/04. The network path was "\\Cdsesub1\n21742\N_000\2004-04-30\crt\datasets" in the EDR.

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

Integrated summaries of the statistical evaluation of NEB-202, NEB-302, and NEB-305 are discussed in this section.

3.1.1 Study Design

Studies NEB-302, NEB-305, and NEB-202 are all randomized, double-blind, placebo-controlled, multi-center, and parallel group studies. The studies consisted of two phases: (1) screening/washout/single-blind placebo run-in and (2) randomization/double-blind treatment. The treatment was administered once daily for 12 weeks. The target population was the patients with mild to moderate hypertension defined as a trough sitting DBP ≥95 and ≤109mmHg. Patients were stratified across all treatment arms by the following factors in order of priority: metabolism of nebivolol; diabetes status (history of diabetes mellitus vs. no history of diabetes mellitus); race (black vs. non-black, except NEB-202); age (<65 and ≥65) and gender. Clinic visits were scheduled for study Days 14, 28, 56 and 84. A 32-fold dose range was evaluated ranging from nevibolol 1.25mg to 40mg. NEB-302 included all of the doses in the range, NEB-202 included all doses except for the lowest dose of 1.25mg, and NEB-305 included only the 5mg, 10mg and 20mg doses. Patients in NEB-302 initiated therapy with nebivolol 30mg were up-titrated to 40mg after 2 weeks if the 30mg dose was tolerated (i.e. resting heart rate was >55bpm).

3.1.2 Primary and Secondary Endpoints

The primary efficacy endpoint of the studies was the change of the average sitting DBP taken at trough (24±2 hours post-previous morning's dose) at end of treatment compared to baseline. The secondary endpoints consisted of the changes at end of treatment compared to baseline in the following measurements:

- Average sitting SBP taken at trough (24±2 hours post-previous morning's dose)
- Average sitting SBP and DBP taken at peak (2-3 hours post-dose)
- Average standing SBP and DBP taken at trough
- Average standing SBP and DBP taken at peak
- Average supine SBP and DBP taken at trough
- Average supine SBP and DBP taken at peak
- Average sitting hear rate at trough

In addition to assessment of hemodynamic changes, efficacy was examined by determining response rates. A responder was defined as a patient whose average sitting DBP at end of study was either <90mmHg or had decreased by ≥10mmHg from baseline.

3.1.3 Statistical Method

An Analysis of Covariance (ANCOVA) model using treatment as a factor and baseline blood pressure, age, gender, race, diabetes status, and metabolism of nebivolol as covariates was used to test for the treatment differences in all efficacy parameters. The primary statistical method of comparison was a dose response test using a linear contrast among all treatment groups up to the 20mg dose. If this contrast was found significant, another linear contrast with all doses excluding the 20mg and 40mg doses was tested, etc.. This step-down test was stopped when no more significant differences were found. No adjustment for multiplicity was indicated because this was a closed testing. The parametric assumptions were not violated and therefore, the non-parametric analysis was not performed.

3.1.4 Sample Size Calculation

3.1.4.1 NEB-302 and NEB-305

A sample size of 59 patients would give 90% power to detect a difference of 4.4mmHg between the any nebivolol dose groups and placebo with estimated standard deviation of 7.2mmHg. To account for a 20% drop-out rate, 75 patients per group for the placebo and the nebivolol 1.25 and 2.5 mg groups and 150 patients for the nebivolol 5mg, 10 mg, 20 mg, and 30/40 mg groups were planned in NEB-302. In NEB-305, 74 patients were to be enrolled in the placebo group and 242 patients in each of the nebivolol 5, 10 and 20mg groups after considering 20% drop-out rate. It was projected that 122 patients on nebivolol 5, 10 and 20mg would give 90% power to detect a 3mmHg difference between any of these doses.

3.1.4.2 NEB-202

Assuming that the high doses of nebivolol had a least a 6.3mmHg difference from placebo, a sample size of 45 patients per group would give >90% power to detect that difference. After considering 10% dropout rate, 50 patients were planned to be enrolled in each treatment arm. The estimated standard deviation was 7.0mmHg.

Reviewer's Comments:
The significance levels used for the sample size calculations were not discussed in the study report. This reviewer assumes the sponsor used the standard significance level of 0.05.

3.1.5 Patient Disposition

A total of 909, 807 and 300 patients were included in intent-to-treat (ITT) population in the analysis of NEB-302, NEB-305, and NEB-202, respectively. The following table summarizes the disposition of the patient in the three pivotal studies.

Table 2: Patient Disposition (ITT Population)
(Source: Sponsor's table)

Study Number	Placebo	1.25mg	2.5mg	5mg	10mg	20mg	30/40mg	Total
302	81	83	82	165	166	166	166	909
305	75	N/A	N/A	244	244	244	N/A	807
202	49	N/A	49	50	51	50	51	300

3.1.6 Demographic and Baseline Characteristics

The baseline demographic characteristics including age, gender, race, diabetes status, metabolism of nebivolol (extensive metabolizer (EM) or poor metabolizer (PM)), and body mass index (BMI) were examined for the balance between the groups. Baseline diastolic blood pressure was similar across treatment groups in NEB-302, NEB-305 and NEB-202.

Table 3: Demographic Characteristics of Pivotal Studies
(Source: Sponsor's analysis confirmed by the reviewer)

NEB-302								
	Placebo	1.25	2.5	5	10	20	30/40	Total
Age								
N	81	83	82	165	166	166	166	909
Mean	56.0	55.5	53.4	54.9	55.2	54.1	54.3	54.7
(SD)	(11.6)	(11.5)	(12.3)	(11.8)	(12.5)	(11.6)	(11.6)	(11.8)
Range	24 – 80	28 – 84	24 - 81	25 - 82	23 - 83	22 - 82	26 - 78	22 – 84
Gender								
Male	46 (56.8)	46 (55.4)	53 (64.6)	96 (58.2)	93 (56.0)	92 (55.4)	92 (55.4)	518 (57.0)
Female	35 (43.2)	37 (44.6)	29 (35.4)	69 (41.8)	73 (44.0)	74 (44.6)	74 (44.6)	391 (43.0)
Race								
Black	11 (13.6)	12 (14.5)	13 (15.9)	23 (13.9)	23 (13.9)	25 (15.1)	25 (15.1)	132 (14.5)
Non-black	70 (86.4)	71 (85.5)	69 (84.1)	142 (86.1)	143 (86.1)	141 (84.9)	141 (84.9)	777 (85.5)
Caucasian	61 (75.3)	60 (72.3)	60 (73.2)	120 (72.7)	114 (68.7)	112 (67.5)	113 (681.)	640 (70.4)
Asian	0 (0.0)	1 (1.2)	0 (0.0)	1 (0.6)	1 (0.6)	2 (1.2)	1 (0.6)	6 (0.7)
Hispanic	9 (11.1)	10 (12.0)	9 (11.0)	21 (12.7)	24 (14.5)	25 (15.1)	25 (15.1)	123 (13.5)
Other	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	4 (2.4)	2 (1.2)	2 (1.2)	8 (0.9)
Diabetes								
Yes	7 (8.6)	9 (10.8)	10 (12.2)	11 (6.7)	17 (10.2)	14 (8.4)	20 (12.0)	88 (9.7)
No	74 (91.4)	74 (89.2)	72 (87.8)	154 (93.3)	149 (89.8)	152 (91.6)	146 (88.0)	821 (90.3)
EM or PM								
Poor	4 (4.9)	5 (6.0)	6 (7.3)	10 (6.1)	11 (6.6)	12 (7.2)	11 (6.6)	59 (6.5)
Extensive	77 (95.1)	78 (94.0)	76 (92.7)	155 (93.9)	155 (93.4)	154 (92.8)	155 (93.4)	850 (93.5)
BMI*								
<30	44 (54.3)	43 (51.8)	45 (54.9)	91 (55.2)	102 (61.4)	101 (60.8)	84 (50.6)	510 (56.1)
≥ 30	37 (45.7)	40 (48.2)	37 (45.1)	74 (44.8)	64 (38.6)	65 (39.2)	82 (49.4)	399 (43.9)
NEB-305								
	Placebo	N/A	N/A	5	10	20	N/A	Total
Age								

N	75	-	-	244	244	244	-	807
Mean	51.2			53.9	53.8	53.4		53.4
(SD)	(10.0)			(11.1)	(11.2)	(11.1)		(11.0)
Range	27 - 73			23 - 79	22 - 82	28 - 80		22 – 82
Gender								
Male	39 (52.0)	-	-	131 (53.7)	131 (53.7)	131 (53.7)	-	432 (53.5)
Female	36 (48.0)			113 (46.3)	113 (46.3)	113 (46.3)		375 (46.5)
Race								
Black	11 (14.7)	-	-	31 (12.7)	33 (13.5)	30 (12.3)	-	105 (13.0)
Non-black	64 (85.3)			213 (87.3)	211 (86.5)	214 (87.7)		702 (87.0)
Caucasian	60 (80.0)			190 (77.9)	191 (78.3)	192 (78.7)		633 (78.4)
Asian	0 (0.0)			4 (1.6)	2 (0.8)	3 (1.2)		9 (1.1)
Hispanic	4 (5.3)			19 (7.8)	17 (7.0)	19 (7.8)		59 (7.3)
Other	0 (0.0)			0 (0.0)	1 (0.4)	0 (0.0)		1 (0.1)
Diabetes								
Yes	4 (5.3)	-	-	9 (3.7)	12 (4.9)	12 (4.9)	-	37 (4.6)
No	71 (94.7)			235 (96.3)	232 (95.1)	232 (95.1)		770 (95.4)
EM or PM								
Poor	4 (5.3)	-	-	15 (6.1)	15 (6.1)	16 (6.6)	-	50 (6.2)
Extensive	71 (94.7)			229 (93.9)	229 (93.9)	228 (93.4)		757 (93.8)
BMI*								
<30	48 (64.0)	-	-	152 (62.6)	145 (59.4)	137 (56.4)	-	482 (59.9)
≥ 30	27 (36.0)			91 (37.4)	99 (40.6)	106 (43.6)		323 (40.1)
Missing	0			1	0	1		2
NEB-202								
	Placebo	N/A	2.5	5	10	20	40	Total
Age								
N	49	-	49	50	51	50	51	300
Mean	49.7		49.9	51.6	50.5	51.3	52.3	50.9
(SD)	(9.1)		(9.6)	(10.5)	(10.5)	(10.8)	(12.0)	(10.4)
Range	34 – 70		33 - 75	26 - 77	29 - 79	28 – 74	28 - 79	26 – 79
Gender								
Male	23 (46.9)	-	26 (53.1)	22 (44.0)	22 (43.1)	21 (42.0)	22 (43.1)	136 (45.3)
Female	26 (53.1)		23 (46.9)	28 (56.0)	29 (56.9)	29 (58.0)	29 (56.9)	164 (54.7)
Diabetes								
Yes	6 (12.2)	-	7 (14.3)	8 (16.0)	6 (11.8)	7 (14.0)	9 (17.6)	43 (14.3)
No	43 (87.8)		42 (85.7)	42 (84.0)	45 (88.2)	43 (86.0)	42 (82.4)	257 (85.7)
EM or PM								
Poor	0 (0.0)	-	1 (2.0)	1 (2.0)	2 (3.9)	1 (2.0)	2 (3.9)	7 (2.3)
Extensive	49 (100.0)		48 (98.0)	49 (98.0)	49 (96.1)	49 (89.0)	49 (96.1)	293 (97.7)
BMI*								
<30	21 (42.9)	-	26 (53.1)	26 (52.0)	26 (51.0)	25 (50.0)	20 (39.2)	144 (48.0)
≥ 30	28 (57.1)		23 (46.9)	24 (48.0)	25 (49.0)	25 (50.0)	31 (60.8)	156 (52.0)

*: BMI is the baseline weight in kilograms divided by the square of the baseline height in meters

The balance of sitting DBP and sitting SBP among the groups were also analyzed. The mean baselines were similar across the groups. The following table shows the results.

Table 4: Baseline Sitting DBP and SBP

(Source: Sponsor's analysis confirmed by the reviewer)

NEB-302								
	Placebo	1.25	2.5	5	10	20	30/40	Total
Sitting Diastolic Blood Pressure (mmHg)								
N	81	83	82	165	166	166	166	909
Mean	100.3	98.9	99.8	99.6	99.5	99.4	99.3	99.5
(SD)	(4.3)	(4.5)	(3.5)	(3.9)	(4.1)	(3.5)	(3.6)	(3.9)
Range	85 – 113	77 – 110	95 - 109	83 – 108	84 - 109	90 - 109	90 - 109	77 – 113
Sitting Systolic Blood Pressure (mmHg)								
N	81	83	82	165	166	166	166	909
Mean	154.9	152.2	150.1	152.6	155.8	151.9	153.1	153.1
(SD)	(15.8)	(14.4)	(13.4)	(13.3)	(14.7)	(15.4)	(14.5)	(14.6)
Range	126 - 197	129 – 195	123 - 185	127 – 189	127 - 195	116 - 195	123 - 196	116 – 197
NEB-305								
	Placebo	N/A	N/A	5mg	10mg	20mg	N/A	Total
Sitting Diastolic Blood Pressure (mmHg)								
N	75	-	-	244	244	244	-	807
Mean	98.7			99.1	98.9	99.2		99.0
(SD)	(3.3)			(3.8)	(4.4)	(3.7)		(3.9)
Range	89 - 108			89 – 111	80 - 119	90 - 112		80 - 119
Sitting Systolic Blood Pressure (mmHg)								
N	75	-	-	244	244	244	-	807
Mean	149.9			151.8	150.5	151.9		151.3
(SD)	(12.5)			(13.2)	(13.1)	(14.8)		(13.6)
Range	126 - 192			119 – 195	121 - 187	117 - 191		117 - 195
NEB-202								
	Placebo	N/A	2.5mg	5mg	10mg	20mg	40mg	Total
Sitting Diastolic Blood Pressure (mmHg)								
N	49	-	49	50	51	50	51	300
Mean	100.8		99.5	100.5	100.3	101.5	98.7	100.2
(SD)	(4.0)		(4.3)	(4.4)	(4.6)	(4.7)	(3.9)	(4.4)
Range	95 - 111		83 - 107	91 - 109	86 - 111	90 - 115	89 - 107	83 - 115
Sitting Systolic Blood Pressure (mmHg)								
N	49	-	49	50	51	50	51	300
Mean	151.4		148.6	151.7	154.2	156.4	150.9	152.2
(SD)	(13.9)		(13.6)	(13.6)	(13.6)	(12.7)	(15.3)	(13.9)
Range	121 - 180		113 - 179	121 - 181	128 - 187	131 - 186	126 - 188	113 - 188

3.1.7 Primary and Secondary Efficacy Results

3.1.7.1 General Hypertensive Population (NEB-302 and NEB-305)

The step-down trend test on LS mean change of sitting DBP from baseline was performed to examine the dose-response relationship of nebivolol and to identify effective antihypertensive doses relative to placebo. The primary analyses showed that the reduction of sitting DBP in all dosed groups of nebivolol was significantly greater than the ones of the placebo group (NEB-302, $p<0.0001$ for all doses; NEB-305, $p=0.0015$ for 5mg, $p=0.0009$ for 10mg, and $p=0.0002$ for 20mg). The results are summarized in the table below.

Table 5: Primary Analysis Results of Pivotal Studies (NEB-302 and NEB-305)
(Source: Sponsor's analysis confirmed by the reviewer)

Treatment	N	Baseline Mean	Mean at the End of Study	LS mean Change from Baseline	Step-DownTrend Test p-value
NEB-302					
Placebo	81	100.3	97.1	-2.9 (1.1)	-
1.25mg	83	98.9	90.8	-8.0 (1.1)	<0.0001
2.5mg	82	99.8	91.1	-8.5 (1.1)	<0.0001
5mg	165	99.6	91.0	-8.4 (1.0)	<0.0001
10mg	166	99.5	90.2	-9.2 (0.9)	<0.0001
20mg	166	99.4	89.5	-9.8 (0.9)	<0.0001
30/40mg	166	99.3	88.0	-11.2 (0.9)	<0.0001
NEB-305					
Placebo	75	98.7	91.4	-4.6 (1.3)	
5mg	244	99.1	88.5	-7.8 (1.0)	0.0015
10mg	244	98.9	87.7	-8.5 (1.0)	0.0009
20mg	244	99.2	87.2	-9.1 (1.0)	0.0002

For the secondary analyses, change of sitting SBP at trough was analyzed by the same statistical method as used for the primary analysis. In addition, difference in response rates between the treatment groups was analyzed.

In NEB-302, the step-down trend test on sitting SBP showed significant trends ($p \leq 0.002$) for all dose ranges tested. In NEB-305, the step-down trend test was statistically significant only for the 20mg dose.

Table 6: Analysis Results of the Trough Sitting SBP in NEB-302 and NEB-305
(Source: Sponsor's analysis confirmed by the reviewer)

	NEB-302			NEB-305		
Treatment	N	LS mean Change	p-value*	N	LS mean change	p-value*
Trough Sitting SBP						
Placebo	81	2.2	-	75	-0.4	-
1.25mg	83	-4.4	0.002	N/A	N/A	N/A
2.5mg	82	-6.3	<0.001	N/A	N/A	N/A
5mg	165	-5.9	<0.001	244	-4.2	0.035
10mg	166	-7.0	<0.001	244	-3.5	0.086
20mg	166	-6.5	<0.001	244	-6.7	<0.001
30/40mg	166	-9.5	N/A	N/A	N/A	N/A

*step-down trend test p-value from an ANCOVA with factor treatment and covariates baseline blood pressure, EM or PM classification, diabetes status, gender, race and age group.

A responder was defined as a patient whose average sitting DBP at trough was either <90mmHg at the end of treatment or had decreased by ≥10mmHg from baseline. The percentage of responders in NEB-302 increased as the dose increased from 45.8% in the nebivolol 1.25mg group to 64.5% in the nebivolol 30/40mg group. The response rates in NEB-305 were 66.0%, 66.8%, and 68.9% in the nebivolol 5mg, 10mg and 20mg groups, respectively. Wald Chi-Square Test was used for trend from logistic regression with factor treatment and covariates baseline blood pressure, EM or PM classification, diabetes status, gender, race, and age group. As done for the blood pressure measurements, step-down testing scheme began with placebo through 20mg and proceeds to step-down until the trend test contains only placebo and the lowest dose group. The results for each treatment group in each study were statistically significant ($p \leq 0.009$). The following table shows the results of the analyses.

Table 7: Responder Rates in NEB-302 and NEB-305
(Source: Sponsor's analysis confirmed by the reviewer)

	NEB-302			NEB-305		
Treatment	Total N	Responder N (%)	p-value	Total N	Responder N (%)	p-value
Placebo	81	20 (24.7)		75	37 (49.3)	
1.25mg	83	38 (45.8)	0.008	-	-	-
2.5mg	82	41 (50.0)	0.001	-	-	-
5mg	165	83 (50.3)	<0.001	244	161 (66.0)	0.009
10mg	166	89 (53.6)	<0.001	244	163 (66.8)	0.005
20mg	166	99 (59.6)	<0.001	244	168 (68.9)	0.002
30/40mg	166	107 (64.5)	NA	-	-	-

Other than trough sitting SBP, standing DBP and SBP at trough and supine DBP and SBP at trough were also analyzed with the same statistical method. The results of the analyses on these

secondary endpoints were generally same as the ones of the trough sitting DBP and SBP. The table showing the results of the analyses can be found in Appendix.

3.1.7.2 Black Patients (NEB-202)

The analysis on the LS mean change of sitting DBP from the baseline was performed. The results of reviewer's analysis showed somewhat different p-values, but the conclusion was the same. The results of the analysis showed that the nebivolol contrast ranging from placebo to 40mg was significant at end of study (p=0.0014). In addition, the linear contrasts for the 5, 10, and 20mg doses of nebivolol were statistically significant (p=0.0187, p=0.0032, and p=0.0019, respectively), demonstrating that the dose range of 5 to 40mg was effective in reducing sitting DBP. The 2.5mg group also showed the decrease of sitting DBP compared to the placebo group, but the test did not show the statistical significance (p=0.14).

Table 8: Primary Analysis Results (NEB-202)
(Source: Reviewer's Analysis)

Treatment	N	Baseline Mean	Mean at the End of Study	LS mean Change from Baseline	Step-Down Trend Test p-value
Placebo	49	100.8	96.4	-2.8 (2.1)	-
2.5mg	49	99.5	92.8	-5.7 (2.1)	0.14
5mg	50	100.5	91.4	-7.7 (2.1)	0.0187
10mg	51	100.3	90.0	-8.9 (2.0)	0.0032
20mg	50	101.5	90.9	-8.9 (2.1)	0.0019
40mg	51	98.7	89.6	-8.3 (2.0)	0.0014

To corroborate the results of NEB-202, the sponsor pooled efficacy data of the black patients in NEB-302 and NEB-305 as well as the blacks from NEB-302 and NEB-305 with NEB-202. However, this reviewer did not perform the pooled analysis because NEB-202 alone shows the efficacy of nebivolol among the black patients, and pooling of the studies may inflate type I error.

For the secondary efficacy analyses, sitting SBP at trough and response rates were analyzed. The analysis on sitting SBP showed that nebivolol is effective over the dose range of 10 to 40mg ($p \leq 0.044$). The step-down test was not significant for the 2.5mg or 5mg nebivolol doses, but these doses showed numerical improvement over placebo. The following table shows the results of the analysis.

Table 9: Mean Change from Baseline to End of Study in Sitting Systolic Blood Pressure (NEB-202) *(Source: Sponsor's analysis confirmed by the reviewer)*

Treatment	N	Baseline Mean	Mean at the End of Study	LS mean Change from Baseline	Step-Down Trend Test p-value
Placebo	49	151.4	147.8	-0.4 (3.8)	-
2.5mg	49	148.6	144.0	-1.9 (3.7)	0.611
5mg	50	151.7	145.8	-3.0 (3.7)	0.383
10mg	51	154.2	144.0	-6.4 (3.6)	0.044
20mg	50	156.4	144.4	-7.6 (3.7)	0.005
40mg	51	150.9	141.4	-7.2 (3.5)	0.002

The rates of responders (patients whose average sitting DBP at end of study were either <90mmHg or had decreased by ≥10mmHg from baseline) were analyzed. At the end of treatment, there were significantly more responders for nebivolol doses of 5 mg and above compared to placebo. There was more responders in nebivolol 2.5mg group compared to the placebo group, but the difference was not statistically significant. The following table shows the results of the analysis.

Table 10: Responder Rates (NEB-202)
(Source: Sponsor's analysis confirmed by the reviewer)

Treatment	Total N	Responder n (%)	p-value
Placebo	49	13 (26.5)	
2.5mg	49	18 (36.7)	0.287
5mg	50	29 (58.0)	0.002
10mg	51	30 (58.8)	<0.001
20mg	50	32 (64.0)	<0.001
40mg	51	29 (56.9)	<0.001

Other than sitting SBP and responder rates, the LS mean change of standing DBP and SBP, supine DBP and SBP were also analyzed. The results of these endpoints were generally same as those of sitting DBP and SBP. A table showing the results of this analysis can be found in Appendix.

3.2 Evaluation of Safety

To document the effects of nebivolol on the electrocardiographic intervals of normal healthy volunteers administered nebivolol, 20 and 40 mg, a randomized, placebo- and active-controlled, parallel-group study in healthy patients (NEB-122) was conducted. Patients were randomized in a 1:1:1:1 ratio to one of four treatments: nebivolol 20mg/40mg, placebo, atenolol 100mg/200mg (active control), or moxifloxacin 400mg (positive control). The subjects were confined to the clinic and received study medication once daily for 7 days. Electrocardiographic intervals were recorded continuously for 24 hours on days 0, 1, 4, and 7 and analyzed at 15 time points on each day (0, 0.5, 1, 1.5, 2, 2.5, 3, 4, 6, 10, 12, 14, 16, 18, and 24 hours post dosing). A sample size of

59 evaluable subjects per group was planned to test with 90% power that the upper bound for the 95% two-sided confidence interval for the difference in mean change from baseline QTc between nebivolol and placebo was no greater than 6 msec. The primary ECG analysis at 2 hours after dosing on Day 7 was performed on 71, 60, 67, and 69 subjects in the nebivolol, atenolol, moxifloxacin, and placebo groups, respectively. For the primary analysis, an analysis of covariance (ANCOVA) with treatment as the main factor and average baseline QTc and gender as the covariates were used. A population correction factor was used to correct QT interval measurements for heart rate. This population correction factor was derived using the mean of all Day 0 data for all subjects (15 time points per subject). The correction factor was defined as the exponent (X) used in the calculation of QTc (where $QTc=QT/(RR)^X$) such that the slope of the resulting QTc versus RR was essentially zero, and was determined through iteration to be 0.329. QT intervals were also corrected using Bazett's and Fridericia's formula.

The primary endpoint was an evaluation of the QTc interval change from baseline (Day 0; average of 15 time points) to 2 hours after dosing on Day 7. If the upper bound of the confidence interval was <6 msec, it would confirm that nebivolol did not cause a clinically significant prolongation in QTc as compared to placebo. Application of Bazett's formula resulted in much greater mean QTc interval decreases, and the results were different from those of using population correction factor or Fridericia's formula because Bazett's formula overcorrected when heart rates were low. Therefore, Bezett's formula was not used for the QTc evaluation for this QT study. The following table shows the interval changes in QTc by treatment comparisons.

Table 11: QTc Interval Change from Day 0 to 2 Hours after Dosing on Day 7
(Source: Sponsor's analysis)

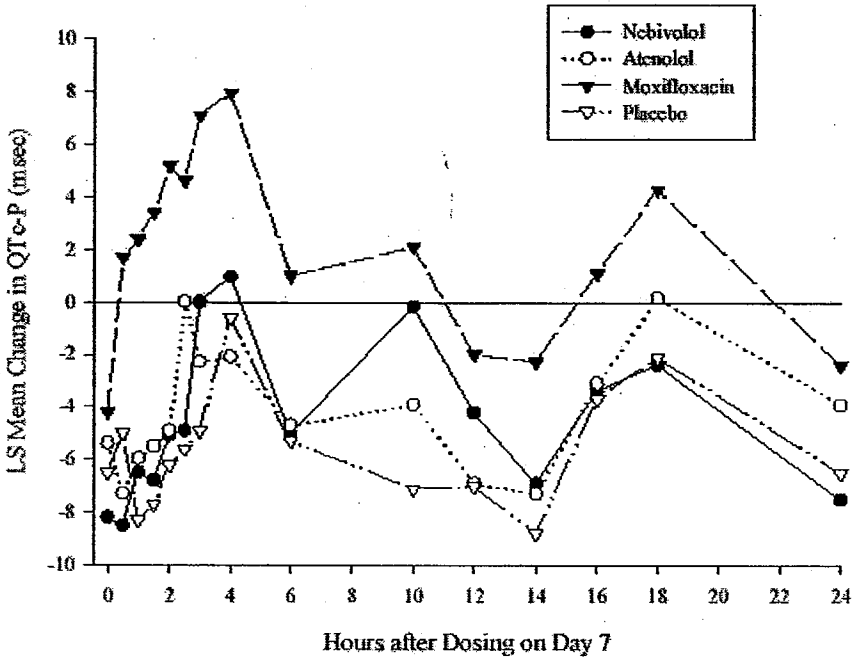
QTc parameter	Comparison	LS Mean of Test	LS Mean of Reference	Difference	95% C.I.	p-value
Population correction factor	Neb. vs. Placebo	-5.0628	-6.2074	1.1446	-4.0907, 6.3779	0.6672
	Moxi vs. Placebo	5.2175	-6.2074	11.4249	6.0936, 16.7563	<0.0001
	Neb. vs. Moxi.	-5.0628	5.2175	-10.280	-15.583, -4.9774	0.0002
	Neb. vs. Atenolol	-5.0628	-4.9142	-0.1486	-5.5762, 5.2790	0.9570
Fridericia's formula	Neb. vs. Placebo	-5.7041	-6.3817	0.6776	-4.5716, 5.9268	0.7996
	Moxi vs. Placebo	5.2976	-6.3817	11.6792	6.336, 17.0248	<0.0001
	Neb. vs. Moxi.	-5.7041	5.2976	-11.002	-16.319, -5.6847	0.0001
	Neb. vs. Atenolol	-5.7041	-5.3953	-0.3088	-5.7508, 5.1332	0.9111

The table showed moxifloxacin resulted in a statistically significant increase in mean QTc interval at 2 hours post dose on Day 7. This indicated the assay sensitivity of NEB-122 study.

In the comparison of nebivolol versus placebo, the mean difference in QTc interval (95% C.I.) at 2 hours after dosing on Day 7 was 1.14 msec (-4.09, 6.38) using the population correction factor and 0.68 msec (-4.57, 5.93) using Fridericia's formula. The sponsor stated that the small differences in the change from mean baseline QTc intervals between nebivolol and placebo clearly demonstrate that nebivolol had no statistically or clinically significant effect on QT interval.

As a secondary analysis, pairwise comparisons of the changes from baseline QTc interval to all time points were evaluated. The following graph showed the LS mean change in QTc-P (population correction) on Day 7

Figure 1: Graph of Changes of QTc on Day 7 *(Source: Sponsor's graph)*



The results of the sponsor's analysis at most other time points on Day 7 were consistent with those of the primary time point. The difference between the nebivolol and placebo was statistically significant (p=0.0146 by population correction factor, and p=0.0255 by Fridericia's formula) at 10 hours. Nevertheless, the least squares mean for nebivolol had little decrease from the baseline (–0.17 using population correction factor, –0.41 using Fridericia's formula), and the small p-values were driven by the greater magnitude of decrease of QTc observed in the placebo group (-7.1 using population correction factor, -6.7 using Fridericia's formula). Therefore, the sponsor claimed no clinically significant effect on QTc interval.

The reviewer's analysis showed different p-values. Although the sponsor's analysis did not show, this reviewer's analysis using the population correction factor showed the statistically

significant difference between the nebivolol group and the placebo group at 3 hours on Day 7. The mean change of QTc of the nebivolol group was 0.5771 and the one of the placebo group was –4.5224, and the p-value was 0.0408. The same analysis using the Fridericia's formula did not meet the significance level of 0.05, but the p-value was close to it (p=0.0643). The following table shows the results of the reviewer's analysis on 15 time points on Day 7.

Table 12: QTc Changes from Baseline on Day 7 *(Source: Reviewer's analysis)*

	LS mean change of Nebivolol	LS mean change of placebo	Difference	95% C.I.	p-value
Fridericia's Formula					
0 hour	-7.9268	-5.8803	-2.0465	-7.5435, 3.4505	0.4669
0.5 hour	-8.8581	-4.8994	-3.9588	-9.2241, 1.3065	0.1429
1 hour	-7.1297	-8.1466	1.0170	-3.7217, 5.7556	0.6747
1.5 hour	-6.9671	-7.4315	0.4644	-5.1234, 6.0523	0.8708
2 hour	-5.8088	-6.2377	0.4290	-4.8881, 5.7461	0.8746
2.5 hour	-5.6669	-5.3043	-0.3627	-5.3847, 4.6593	0.8877
3 hour	0.1767	-4.4557	4.6324	-0.2341,9.4989	0.0643
4 hour	0.4766	-0.4531	0.9297	-4.3343, 6.1937	0.7298
6 hour	-4.8953	-4.5991	-0.2962	-5.7448, 5.1524	0.9153
10 hour	-0.0053	-6.3171	6.3117	0.6085, 12.0149	0.0319
12 hour	-4.2940	-6.7354	2.4414	-2.5734, 7.4563	0.3417
14 hour	-7.3824	-8.2778	6.7669	-0.5302, 14.0640	0.7368
16 hour	-3.2928	-3.4307	0.1379	-4.8405, 5.1143	0.9568
18 hour	-2.3628	-1.5783	-0.7844	-6.2866, 4.7177	0.7804
24 hour	-7.9335	-5.9069	-2.0266	-7.7041, 3.6509	0.4854
Population Correction Factor					
0 hour	-7.9672	-6.0379	-1.9293	-3.5603, 7.4189	0.4921
0.5 hour	-8.5128	-4.9225	-3.5902	-8.8377, 1.6573	0.1822
1 hour	-6.4415	-8.2220	1.7805	-2.9121, 6.4731	0.4584
1.5 hour	-6.6070	-7.4476	0.8405	-4.6971, 6.3781	0.7665
2 hour	-5.2825	-6.2003	0.9189	-4.4307, 6.2685	0.7372
2.5 hour	-5.0964	-5.2357	0.1393	-4.9392, 5.2179	0.9572
3 hour	0.5771	-4.5224	5.0995	0.2614, 9.9376	0.0408
4 hour	0.6323	-0.5466	1.1789	-4.1068, 6.4646	0.6627
6 hour	-4.6468	-4.9100	0.2632	-5.1683, 5.6948	0.9245
10 hour	0.2164	-6.5050	6.7214	1.0155, 12.4274	0.0225
12 hour	-4.1912	-6.7322	2.5409	-2.4455, 7.5273	0.3197
14 hour	-7.1203	-8.5689	1.4486	-3.7401, 6.6373	0.5852
16 hour	-3.5118	-3.4542	-0.0575	-5.0465, 4.9315	0.9820
18 hour	-2.0696	-1.7131	-0.3564	-5.8754, 5.1626	0.8995
24 hour	-7.5915	-6.1133	-1.4782	-7.1416, 4.1852	0.6098

Since there was a statistically significant gender effect (p=0.0017), this reviewer analyzed each gender separately. Females generally showed greater QTc interval increase then males. The females in the nebivolol group had a statistically significant increase of QTc interval by Fridericia's formula at 3 and 10 hours on Day 7 (p=0.0167 and 0.0338) compared to the placebo group, but males did not show the statistically significant difference at those time points. As shown in the table below, the mean changes of QTc interval at 3 hours were 6.5569 for the nebivolol group and –2.5777 for the placebo group, and the ones at 10 hours were 0.5661 for the nebivolol group and –8.4570 for the placebo group. The analysis results by population correction factor were similar. As the sponsor stated, the QTc changes at 10 hours among females may not be clinically significant since the QTc of the nebivolol group increased only 0.5661 msec from the baseline. However, the QTc change of the nebivolol group at 3 hours was 6.9322 and 6.5569 by the population correction factor and the Fridericia's formula, respectively, and this increase was statistically significant from baseline and from the placebo group.

Table 13: QTc Changes by Gender at 3 Hours and 10 Hours on Day 7 (Fridericia's Formula) *(Source: Reviewer's analysis)*

	LS mean change of Nebivolol	LS mean change of Placebo	Difference	95% C.I.	p-value
3 hours on Day 7					
Male	-5.6326	-6.7841	1.1516	-5.3517, 7.6549	0.7295
Female	6.5569	-2.5777	9.1346	1.8760, 16.3932	0.0167
Male + Female	0.1767	-4.4557	4.6324	-0.2341, 9.4989	0.0643
10 hours on Day 7					
Male	-0.7559	-5.4496	4.6937	-3.1398, 12.5272	0.2440
Female	0.56613	-8.4570	9.0231	0.8936, 17.1526	0.0338
Male + Female	-0.0053	-6.3171	6.3117	0.6185, 12.0149	0.0319

Table 14: QTc Changes by Gender at 3 Hours and 10 Hours on Day 7 (Population Correction Factor) *(Source: Reviewer's analysis)*

	LS mean change of Nebivolol	LS mean change of Placebo	Difference	95% C.I.	p-value
3 hours on Day 7					
Male	-5.2347	-6.9539	1.7192	-4.7449, 8.1833	0.6037
Female	6.9322	-2.5293	9.4615	2.2391, 16.6839	0.0129
Male + Female	0.5771	-4.5224	5.0995	0.2614, 9.9376	0.0408
10 hours on Day 7					
Male	-0.5078	-5.6206	5.1128	-2.7396, 12.9651	0.2059
Female	0.7567	-8.6672	9.4239	1.3086, 17.5394	0.0267
Male + Female	0.2164	-6.5050	6.7214	1.0155, 12.4274	0.0225

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race and Age

The LS mean change in sitting DBP from baseline to end of study by age, gender, race was analyzed. In general, the results showed a decrease in sitting DBP over placebo for all subgroups.

Table 15: Subgroup Analysis on Age, Gender, and Race (NEB-302)
(Source: Sponsor's analysis)

	Placebo	1.25mg	2.5mg	5mg	10mg	20mg	30/40mg
AGE							
<65							
N	64	65	68	132	125	134	128
LS mean	-2.3	-8.1	-8.3	-8.3	-9.2	-9.6	-11.5
≥ 65							
N	17	18	14	33	41	32	38
LS mean	-6.0	-7.9	-9.5	-9.4	-9.6	-10.8	-10.6
GENDER							
Male							
N	46	46	53	96	93	92	92
LS mean	-2.2	-7.1	-7.9	-8.1	-8.4	-9.3	-11.9
Female							
N	35	37	29	69	73	74	74
LS mean	-4.2	-9.2	-9.2	-8.9	-10.5	-10.6	-6.5
Race							
Black							
N	11	12	13	23	23	25	25
LS mean	-0.5	-10.5	-6.2	-6.7	-8.9	-4.3	-10.6
Non-black							
N	70	71	69	142	143	141	141
LS mean	-5.1	-9.3	-10.5	-10.4	-11.0	-12.5	-13.1

Appears This Way
On Original

Table 16: Subgroup Analysis on Age, Gender and Race (NEB-305)

(Source: Sponsor's analysis)

	Placebo	5mg	10mg	20mg
AGE				
<65				
N	67	198	197	196
LS mean	-3.9	-7.6	-8.1	-8.8
≥ 65				
N	8	45	47	47
LS mean	-9.6	-9.9	-11.1	-11.8
GENDER				
Male				
N	39	130	131	131
LS mean	-5.7	-8.5	-9.5	-9.4
Female				
N	36	113	113	112
LS mean	-3.9	-7.5	-7.6	-9.3
Race				
Black				
N	11	31	33	30
LS mean	-5.3	-10.7	-8.2	-8.7
Non-black				
N	64	212	211	213
LS mean	-5.9	-8.9	-10.0	-10.8

Table 17: Subgroup Analysis on Age, and Gender (NEB-202)

(Source: Sponsor's analysis)

	Placebo	2.5mg	5mg	10mg	20mg	40mg
AGE						
<65						
N	44	45	44	45	45	42
LS mean	-4.0	-6.7	-7.8	-8.8	-9.3	-9.0
≥ 65						
N	5	4	6	6	5	9
LS mean	1.9	3.5	-9.8	-13.0	-8.1	-8.9
GENDER						
Male						
N	23	26	22	22	21	22
LS mean	-1.7	-7.8	-10.9	-9.1	-9.7	-9.0
Female						
N	26	23	28	29	29	29
LS mean	-1.1	-1.1	-2.4	-6.2	-5.6	-5.4

4.1 Other Special/Subgroup Populations

Subgroup of BMI, diabetes status, and EM or PM classification were analyzed for the LS mean change in sitting DBP from baseline to end of study. In general, the results showed a decrease in sitting DBP over placebo for all subgroup. However, the diabetes group in NEB-305 trial showed greater reduction of sitting DBP in placebo group compared to the nebivolol treated groups. The following tables summarize the subgroups analyses results of NEB-302, NEB-305, and NEB-202.

Table 18: Subgroup Analysis on BMI, Diabetes Status, and EM or PM Classification (NEB-302)
(Source: Sponsor's analysis)

	Placebo	1.25mg	2.5mg	5mg	10mg	20mg	30/40mg
BMI							
<30							
N	44	43	45	91	102	101	84
LS mean	-5.0	-9.8	-9.6	-10.7	-11.0	-11.4	-13.8
≥ 30							
N	37	40	37	74	64	65	82
LS mean	-0.4	-5.8	-7.0	-5.6	-6.7	-7.7	-8.1
Diabetes							
Yes							
N	7	9	10	11	17	14	20
LS mean	-4.9	-9.3	-14.7	-6.9	-10.7	-13.2	-10.9
No							
N	74	74	72	154	149	152	146
LS mean	-2.1	-7.1	-7.0	-7.8	-8.4	-8.8	-10.7
EM/PM							
PM							
N	4	5	6	10	11	12	11
LS mean	-2.7	-8.9	-12.8	-10.8	-13.6	-10.0	-11.0
EM							
N	77	78	76	155	155	154	155
LS mean	-2.1	-7.1	-7.2	-7.4	-7.9	-8.8	-10.3

Table 19: Subgroup Analysis on BMI, Diabetes Status, and EM or PM Classification (NEB-305)
(Source: Sponsor's analysis)

	Placebo	5mg	10mg	20mg
BMI				
<30				
N	48	152	145	137
LS mean	-4.4	-7.6	-8.3	-9.5
≥ 30				
N	27	91	99	106
LS mean	-5.1	-9.0	-9.4	-9.4
Diabetes				
Yes				
N	4	9	12	12
LS mean	-11.8	-9.9	-7.7	-8.4
No				
N	71	234	232	231
LS mean	-4.9	-8.5	-9.3	-9.9
EM/PM				
PM				
N	4	15	15	16
LS mean	-5.7	-9.8	-9.3	-10.8
EM				
N	71	228	229	227
LS mean	-5.2	-8.4	-9.2	-9.7

Table 20: Subgroup Analysis on BMI, and Diabetes Status (NEB-202)
(Source: Sponsor's analysis)

	Placebo	2.5mg	5mg	10mg	20mg	40mg
BMI						
<30						
N	21	26	26	26	25	20
LS mean	-2.1	-8.3	-8.5	-8.0	-12.5	-9.0
≥ 30						
N	28	23	24	25	25	31
LS mean	-4.4	-4.4	-8.2	-10.9	-6.2	-8.7
Diabetes						
Yes						
N	6	7	8	6	7	9
LS mean	-2.2	-7.1	-3.3	-6.6	-11.4	-9.9
No						
N	43	42	42	45	43	42
LS mean	-4.3	-6.7	-9.8	-10.5	-9.5	-9.4

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

The primary analyses on general population (NEB-302 and NEB-305) showed that the sitting DBP of all dosed groups was significantly decreased compared to the placebo groups (NEB-302, $p<0.0001$ for all doses; NEB-305, $p<0.0015$). The same analysis on black population (NEB-202) showed a statistically significant reduction of sitting DBP in all dose groups except 2.5mg dose group. The following table shows the results of the analyses.

Table 21: Primary Analysis Results of Pivotal Studies
(Source: Reviewer's analysis)

Treatment	N	Baseline Mean	Mean at the End of Study	LS mean Change from Baseline	Step-DownTrend Test p-value
NEB-302					
Placebo	81	100.3	97.1	-2.9 (1.1)	-
1.25mg	83	98.9	90.8	-8.0 (1.1)	<0.0001
2.5mg	82	99.8	91.1	-8.5 (1.1)	<0.0001
5mg	165	99.6	91.0	-8.4 (1.0)	<0.0001
10mg	166	99.5	90.2	-9.2 (0.9)	<0.0001
20mg	166	99.4	89.5	-9.8 (0.9)	<0.0001
30/40mg	166	99.3	88.0	-11.2 (0.9)	<0.0001
NEB-305					
Placebo	75	98.7	91.4	-4.6 (1.3)	-
5mg	244	99.1	88.5	-7.8 (1.0)	0.0015
10mg	244	98.9	87.7	-8.5 (1.0)	0.0009
20mg	244	99.2	87.2	-9.1 (1.0)	0.0002
NEB-202 (Black Population)					
Placebo	49	100.8	96.4	-2.8 (2.1)	-
2.5mg	49	99.5	92.8	-5.7 (2.1)	0.14
5mg	50	100.5	91.4	-7.7 (2.1)	0.0187
10mg	51	100.3	90.0	-8.9 (2.0)	0.0032
20mg	50	101.5	90.9	-8.9 (2.1)	0.0019
40mg	51	98.7	89.6	-8.3 (2.0)	0.0014

For the secondary analyses, change of sitting SBP at trough and rates of responder, which was defined as a patient whose average sitting DBP at trough was either <90mmHg at the end of treatment or had decreased by ≥10mmHg from baseline, were analyzed. The results of these secondary analyses confirmed the findings from the primary analyses.

5.2 Conclusions and Recommendations

Nebivolol had a statistically significant effect on reducing sitting DBP in both non-black and black populations. Secondary analyses on other efficacy measurements confirmed that nebivolol had antihypertensive effects in mild to moderate hypertensive population.

APPENDICES

NEB-302 Summary of Results of Step-Down Trend Test, LS mean, and Difference from Placebo in LS Mean Change in Blood Pressure from Baseline to End of Study at Trough (ITT) *(Source: Sponsor's analysis)*

	Sitting			Standing			Supine		
	p-value[a,b]	LS Mean[c]	LS Mean Diff	p-value[a,b]	LS Mean[c]	LS Mean Diff	p-value[a,b]	LS Mean[c]	LS Mean Diff
Placebo									
DBP	—	-2.9	—	—	0	—	—	-2.5	—
SBP	—	2.2	—	—	3.8	—	—	0.6	—
Nebivolol 1.25mg									
DBP	<0.001	-8.0	-5.1	<0.001	-4.6	-4.5	0.018	-5.5	-3.0
SBP	0.002	-4.4	-6.6	0.002	-3.0	-6.8	0.011	-4.7	-5.3
Nebivolol 2.5mg									
DBP	<0.001	-8.5	-5.6	<0.001	-6.5	-6.4	<0.001	-7.6	-5.1
SBP	<0.001	-6.3	-8.4	<0.001	-6.3	-10.1	<0.001	-8.3	-8.9
Nebivolol 5mg									
DBP	<0.001	-8.4	-5.5	<0.001	-5.2	-5.2	<0.001	-7.4	-4.9
SBP	<0.001	-5.9	-8.1	<0.001	-4.1	-8.0	<0.001	-7.6	-8.2
Nebivolol 10mg									
DBP	<0.001	-9.2	-6.3	<0.001	-6.6	-6.6	<0.001	-7.9	-5.4
SBP	<0.001	-7.0	-9.2	<0.001	-5.3	-9.1	<0.001	-7.1	-7.7
Nebivolol 20mg									
DBP	<0.001	-9.8	-6.9	<0.001	-7.4	-7.3	<0.001	-8.4	-5.9
SBP	<0.001	-6.5	-8.6	<0.001	-5.1	-8.9	<0.001	-7.1	-7.8
Nebivolol 30/40mg									
DBP	—	-11.2	-8.3	—	-9.1	-9.1	—	-10.1	-7.6
SBP	—	-9.5	-11.7	—	-8.5	-12.4	—	-10.9	-11.5

NEB-305 Summary of Results of the Step-Down Trend Test, LS Mean, and Difference from Placebo in LS Mean Change in Blood Pressure from Baseline to End of Study at Trough (ITT) *(Source: Sponsor's analysis)*

	Sitting			Standing			Supine		
	p-value[a,b]	LS Mean[c]	LS Mean Diff	p-value[a,b]	LS Mean[c]	LS Mean Diff	p-value[a,b]	LS Mean[c]	LS Mean Diff
Placebo									
DBP	—	-4.6	—	—	-3.7	—	—	-3.4	—
SBP	—	-0.4	—	—	-0.9	—	—	1.0	—
Nebivolol 5mg									
DBP	0.002	-7.8	-3.2	0.002	-6.9	-3.2	<0.001	-7.8	-4.4
SBP	0.035*	-4.2	-3.8	0.016*	-5.3	-4.4	0.012*	-3.6	-4.6
Nebivolol 10mg									
DBP	<0.001	-8.5	-3.9	<0.001	-7.2	-3.5	<0.001	-7.7	-4.3
SBP	0.086	-3.5	-3.1	0.107	-3.8	-3.0	0.082	-2.2	-3.2
Nebivolol 20mg									
DBP	<0.001	-9.1	-4.5	<0.001	-8.1	-4.4	<0.001	-8.4	-5.0
SBP	<0.001	-6.7	-6.3	0.002	-7.2	-6.4	<0.001	-5.9	-7.0

Appears This Way
On Original

NEB-202 Summary of Results of the Step-Down Trend Test, LS Mean, and Difference from Placebo in LS Mean Change in Blood Pressure from Baseline to End of Study at Trough (ITT) *(Source: Sponsor's analysis)*

Blood Pressure Parameter	Placebo	Nebivolol 2.5mg			Nebivolol 5mg			Nebivolol 10mg			Nebivolol 20mg			Nebivolol 40mg		
	LS[c] Mean	p-value[a,b]	LS[c] Mean	LS Mean Diff[b,c]	p-value[a,b]	LS[c] Mean	LS Mean Diff[b,c]	p-value[a,b]	LS[c] Mean	LS Mean Diff[b,c]	p-value[a,b]	LS[c] Mean	LS Mean Diff[b,c]	p-value[a,b]	LS[c] Mean	LS Mean Diff[b,c]
Sitting																
DBP	-2.8	0.084	-5.7	-2.9	0.004	-7.7	-4.9	<0.001	-8.9	-6.1	<0.001	-8.9	-6.0	<0.001	-8.3	-5.5
SBP	-0.4	0.611[*]	-1.9	-1.5	0.383	-3.0	-2.6	0.044	-6.4	-6.0	0.005	-7.6	-7.3	0.002	-7.2	-6.8
Standing																
DBP	-5.1	0.651	-5.9	-0.8	0.044	-8.7	-3.6	0.003	-9.7	-4.6	0.002	-9.4	-4.3	<0.001	-10.1	-5.0
SBP	-4.0	>0.999[*]	-4.0	0.0	0.292[*]	-7.2	-3.2	0.175[*]	-7.2	-3.2	0.093	-8.1	-4.1	0.016	-10.2	-6.2
Supine																
DBP	-4.4	0.056	-7.8	-3.3	0.028	-8.2	-3.8	0.001	-10.1	-5.7	0.001	-9.6	-5.2	0.001	-9.5	-5.1
SBP	-5.4	0.943[*]	-5.1	0.2	0.963[*]	-5.5	-0.1	0.142[*]	-9.6	-4.3	0.175[*]	-7.4	-2.1	0.054	-9.6	-4.3

Data Source: Table 2.1.1, Table 2.2.1, Table 2.5.1, Table 2.6.1, Table 2.9.1, Table 2.10.1

[a]P-value from step-down trend test. Step-down testing began with placebo to nebivolol 40mg and proceeded to step down until the test contained only placebo and nebivolol 2.5mg.

[b]From an ANCOVA with factor treatment and covariates baseline blood pressure, age group, gender, diabetes status, and metabolism rate

[c]LS mean change in DBP or SBP from baseline to end of study; difference from placebo in LS mean change in DBP or SBP from baseline to end of study

[*]: P-values associated with lower doses are not applicable in the context of step-down trend testing due to the non-significant result at the higher dose.

Note: P-value and LS mean difference are not applicable for placebo; therefore, these columns are not displayed.

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

/s/

Jasmine Choi
12/16/04 04:00:06 PM
BIOMETRICS

James Hung
12/16/04 04:02:02 PM
BIOMETRICS

Kooros Mahjoob
12/17/04 02:09:39 PM
BIOMETRICS

Statistical Review and Evaluation

Review of Carcinogenicity Studies

NDA#:	21,742
APPLICANT:	Berteck Pharmaceuticals, Inc.
NAME OF DRUG:	Nebivolol
STUDIES REVIEWED:	Two-Year Study in Rats and 18 Months Study in Mice
PHARMACOLOGY REVIEWER:	Elizabeth Hausner, Ph.D. (HFD-110)
STATISTICAL REVIEWER:	Jasmine Choi, M.S. (HFD-710)

This review consists of 6 pages of text and another 17 pages of graphs and tables.

Table of Contents

Table of Tables

Appendix

1. Introduction

This NDA was submitted for assessing the influence of Nebivolol on tumor formation. A statistical review was done for two carcinogenicity studies: a two year rat study (Study 1968) and an 18 months mouse study (Study 1967).

2. Two-Year Study in Rat (Study 1968)

2.1 Study Design

A total of 500 SPF Wistar rats (50 rats/sex/group) were assigned to the untreated control, the vehicle control, the 2.5 mg/kg/day, the 10 mg/kg/day, or the 40 mg/kg/day dose groups. The drug was administered orally through the powdered diet for 24 (females) to 26 (males) months. The drug was delivered in a vehicle, beta cyclodextrin, to enhance oral bioavailability in rodents. At the time of the study, little was known of the pharmacological activity of beta cyclodextrins and therefore a vehicle control group was included into the design. Nebivolol for humans does not include the vehicle. At the end of the treatment period all surviving animals were sacrificed.

2.2 Sponsor's Analysis Methods and Results

The study was extended beyond the original length till approximately 50% mortality was reached in the control and the low dose groups. Male rats were sacrificed terminally after approximately 111 weeks and female rats were sacrificed terminally after approximately 105 weeks. For survival analysis, the trend test and the pair-wise comparisons against the control groups were performed using Chi-square tests. No positive trend in mortality rates was observed. The Chi-square tests revealed no differences in mortality rates between groups.

Before analyzing each tumor type, age-adjusted positive trends in total numbers of animals bearing fatal, incidental, and fatal and/or incidental tumors were performed. The death rate method was applied to the fatal tumor type, and the prevalence method was applied to the incidental tumor type. Equidistant dose levels 0, 1, 2, 3 were used for the control, low, medium and high dose groups. For each tumor type and for each context of observation (fatal or incidental tumor), the asymptotic probability for a dose related trend was computed. Where appropriate, the fatal and incidental trends statistics and their variances were summed, and the overall asymptotic one-tailed probability for that tumor type was computed. When the asymptotic p-value was at least marginally statistically significant ($p<0.10$) and when the total number of tumor bearing animals in all treatment groups was 8 or less, the "exact" age-adjusted Cochran-Armitage trend test was considered the preferred results. The trend test was performed against both the untreated control group and the vehicle control group. The significance level used for the trend test was not mentioned in the study report.

The analysis showed a positive dose-related trend for the incidence of vascular neoplasia in the spleen of males in comparison with the untreated control and the vehicle control

groups (p=0.0364 against untreated control, p=0.0250 against vehicle control). This finding comprised 2 hemangioendotheliomas and 2 hemangioendothelial sarcomas in the 40 mg/kg dosed male group and zero tumors in either control group. However, the sponsor stated that this finding was considered coincidental and of no relevance since the spleen is part of the lymphoid system and therefore the incidence of vascular neoplasia should be evaluated for the overall lymphoid system. The incidences for vascular neoplasia in the lymphoid system of all control and dosed groups fell within historical ranges. Other than this finding, the tumor analysis did not reveal any other positive tumor trend in either the male or female rats.

2.3 Reviewer's Analysis Methods and Results

This reviewer performed dose-mortality trend tests and homogeneity tests as survival analysis. The survival of both male and female rats was similar across all groups (Appendices 1-4). The test results are summarized in the following table.

Table 1: Survival Analysis of Rats

Gender	Tests	P-values	
		Cox	Kruskal-Wallis
Male	Dose-Mortality Trend	0.8654	0.9371
	Homogeneity	0.8530	0.7206
Female	Dose-Mortality Trend	0.2212	0.7416
	Homogeneity	0.3684	0.4479

Tumor findings were analyzed by an exact permutation trend test using actual dose levels, 0, 2.5, 10, and 40 mg/kg as weights. The following standard approach was used in the analyses: for rare tumors (defined as an incidence rate ≤1% usually based on concurrent controls) the significance level was 0.025 while for common tumors, the significance level was 0.005. Fatal and incidental tumors were analyzed separately using the death rate and the prevalence method, respectively. When fatal and incidental tumors occurred in the same time interval, the asymptotic test was used since the exact test is not accurate in those circumstances, and the asymptotic test may give a better approximation unless the number of tumors was small. This rule was different from the sponsor's. As noted above, the use of exact test depended on whether there were at most eight animals with tumors and whether the p-value of the asymptotic test was less than 0.1. Also, this reviewer used actual dose levels when the sponsor used 0, 1, 2, and 3 for weights. A pairwise comparison between the untreated control group and the vehicle control group was also performed. These comparisons showed no statistically significant difference in mortality or in tumor incidences. In the reviewer's analysis, mortality and tumor trends were analyzed using the vehicle control group. At the request of Elizabeth Hausner, Ph.D., the reviewing pharmacologist, the analysis was performed against the untreated control group as well. The results of these analyses were not different from the ones obtained with the vehicle control group. Only results with the vehicle controls are presented in the Appendix 5 and 6.

Using the analysis method as stated above, no statistically significant positive tumor trend was found in either male or female rats.

2.4 Validity of Rat Study

Since there was no statistically significant dose related tumor trend in rats, the validity of the study was evaluated for each gender. The following two questions need to be answered:

1) Were enough animals exposed for a sufficient length of time to allow for late developing tumors?
2) Were the dose levels high enough to pose a reasonable tumor challenge in the animals?

To answer the first question, usually the proportions of animals surviving at weeks 80-90 would be examined. This particular study, however, was extended till 50% survival was achieved. Therefore, it can be concluded that sufficient numbers of animals were at risk for a sufficient length of time.

In determining the appropriateness of the chosen dose levels, it is generally accepted that the high dose should be close to the MTD (Maximum Tolerated Dose). One of the following criteria should be met:

1) A dose is considered adequate if there is a detectable reduction in average body weights of up to 10% in the dosed group relative to the controls.
2) The administered dose is also considered an MTD if dosed animals exhibit severe toxic effects attributed to the chemical.
3) A dose is considered adequate if the dosed animals show a slightly increased mortality compared to the controls.

The sponsor's body weight graphs for the male and female rats (Appendices 7 and 8) showed more than 10% reduction of body weights of the high dose groups compared to the untreated control and the vehicle control groups, giving sufficient support for criterion 1. It was concluded that enough animals were exposed to the drug for a sufficient length of time and that the high dose was close to the MTD.

3. 18 Month Study in Mouse (Study 1967)

3.1 Study Design

A total of 500 SPF Albino Swiss mice (50 mice/sex/group) were assigned to the untreated control, the vehicle control, the 2.5 mg/kg/day, the 10 mg/kg/day, or the 40 mg/kg/day dose groups. The drug was administered orally as a coprecipitate with beta-cyclodextrin by admixture with the diet. At the time of the study, little was known of the pharmacological activity of beta cyclodextrins and therefore a vehicle control group was

included into the design. Nebivolol for humans does not include the vehicle. At the end of the treatment period all surviving animals were sacrificed.

3.2 Sponsor's Analysis Method and Results

The duration of the study was extended to approximately 19 to 20 months from originally planned 18 months till survival of the control and low dose groups dropped to approximately 50%. As survival analysis, one-sided trend tests and Chi-square tests were performed. No positive dose-related trend in mortality rates was found. The Chi-square tests showed no significant differences in mortality between groups.

Before analyzing the incidence of the various tumor types, an age-adjusted analysis was carried out for all tumor-bearing animals, for all fatal tumor-bearing animals, and for all incidental tumor-bearing animals. These analyses did not reveal a single positive dose related trend in either males or females. For each individual tumor type, the overall asymptotic one-sided probability for that tumor type was computed. If the asymptotic p-value was less than 0.10 and the total number of tumor bearing animals was 8 or less, then the exact age-adjusted Cochran-Armitage trend test was performed. Equidistant dose levels 0, 1, 2, and 3 were used. The significance levels were not pre-specified. A significant trend was found in testicular Leydig cell tumors among male mice when using either untreated controls or vehicle controls ($p<0.001$). This effect was confirmed by the pairwise comparisons (Chi-square test) between the 40 mg/kg dose group and each of the control groups. It is considered to be test article related.

3.3 Reviewer's Analysis Methods and Results

Dose-mortality trend tests and homogeneity tests were performed as survival analyses. As shown in Appendices 9-12, the survival of male and female mice was similar across all groups. The following table summarizes the test results.

Table 2: Survival Analysis of Mice

Gender	Tests	P-values	
		Cox	Kruskal-Wallis
Male	Dose-Mortality Trend	0.7162	0.9082
	Homogeneity	0.6936	0.8026
Female	Dose-Mortality Trend	0.9039	0.9417
	Homogeneity	0.9214	0.9119

An exact permutation trend test was performed to test individual tumor incidences for positive linear trend. The asymptotic test was used when the same tumor occurred in the fatal and incidental context in the same time interval and if the number of tumors was not small. This analysis method is different from the sponsor's, but did not lead to different conclusions. Levels of significance for rare and common tumors were used as described in section 2.3. The mortality and the tumor trend tests were performed using the vehicle control group. No statistically significant differences in mortality or in tumor incidences were observed between the untreated control and the vehicle control groups. At the

request of reviewing pharmacologist, the tumor trend tests were also performed using the untreated control group. These tests did not show any different results from the ones with the vehicle control group.

Leydig cell tumor in the testis occurred in 1, 2, 0, 1, and 21 animals of the untreated control, the vehicle control, the low, the med., and the high dose male groups, respectively. The asymptotic test was performed for this common tumor and the p-value reached the significance level ($p < 0.001$). This finding is summarized in the table below. No other tumor finding reached statistical significance among the male or female mice (Appendix 13&14).

Table 3: Significant Tumor Findings and Test Results for Male Mice

Organ	Tumor	Vehicle	2.5 mg/kg	10 mg/kg	40 mg/kg	Asymptotic trend test
Testis	Leydig cell tumor	2	0	1	21	<0.001

3.4 Validity of the Female Mouse Study

The validity of the female mouse study was evaluated because no statistically significant tumor trend was found. The criteria used were as stated in section 2.4. Though the survival rate at weeks 80-90 fell just below 50% (Appendix 12) for some treatment groups, one can conclude that sufficient numbers of animals were exposed long enough. This argument is also supported by the fact that the study was extended beyond 78 weeks till 50% mortality was reached. The mean body weight graph and the mortality across the groups were examined to determine whether the high dose was close to MTD. Based on the sponsor's graph for female body weights (Appendix 16), at weeks 72 and 80, the high dose group had average body weights of about 10% less than the vehicle control group. However, this difference was not observed through out the study. During the first year with some exemptions, differences of only 5% or less were observed. The mortality rate of the high dose female mice was similar to the mortality rate of the vehicle controls. Based on these evaluations, the high dose did not reach the MTD. The evaluation of severe toxic effects will be left to the expertise of the reviewing pharmacologist.

4. Summary

4.1 Rat Study

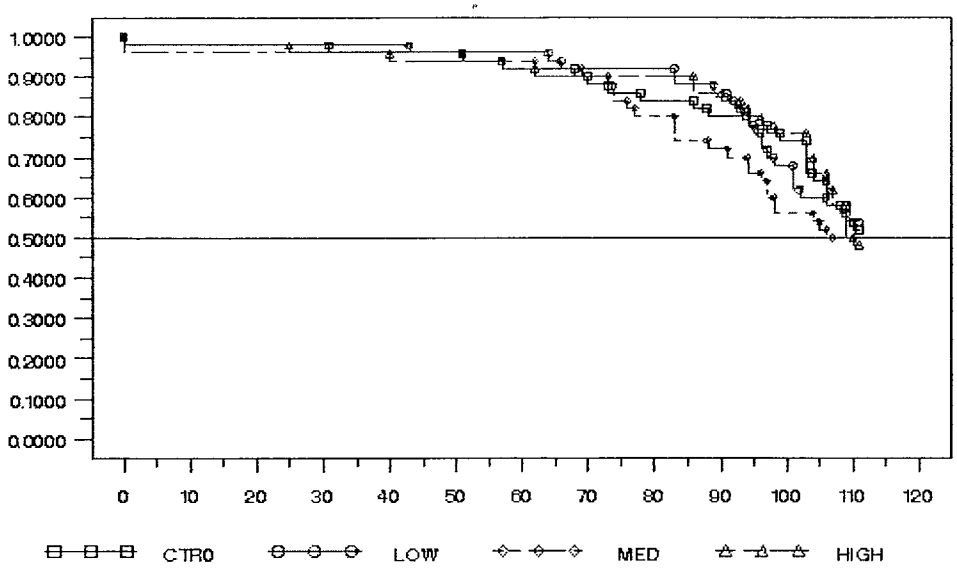
Fifty rats per group/sex received the drug at levels of 0, 0, 2.5, 10, and 40 mg/kg/day in the diet for up to 111 weeks. The second control group received the vehicle in the diet. The dose-mortality trend tests and homogeneity test for both genders showed no statistically significant treatment effect on survival. In fact the study had been extended till 50% mortality was reached. The reviewer's analysis showed that none of the tumors reached statistical significance by trend test in either gender. In the sponsor's analysis, hemangioendothelial sarcoma and hemangioendothelioma in the spleen was combined, and this showed a significant trend. However, those two tumor types were analyzed

separately in this reviewer's analysis, and the trends were not statistically significant. This reviewer evaluated the validity of the study. Based on the statistical criteria, there were a sufficient number of rats living long enough to present late developing tumors, and the high dose reached the MTD.

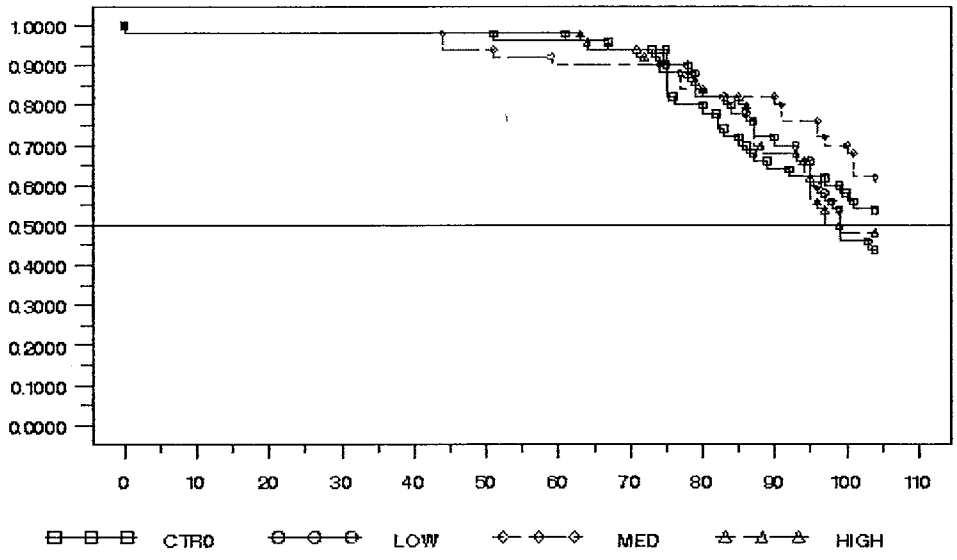
4.2 Mouse Study

Fifty mice per group/sex received nebivolol at levels of 0, 0, 2.5, 10, and 40 mg/kg/day in diet for up to 20 months. The second control group received the vehicle in the diet. The survival analyses showed that the treatment did not affect the survival of either gender. Leydig cell tumor in the testis showed a statistically significant trend in males. There was no other tumor with a positive dose-related trend. Since there was no statistically significant tumor finding in the females, the validity of the study was evaluated. The evaluation suggested that enough numbers of animals were at risk for a sufficient length of time, but that the high dose did not reach the MTD.

Appendix 1: Survival Graph of Male Rats



Appendix 2: Survival Graph of Female Rats



Appendix 3: Mortality of Male Rats

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTR0	0-52	50	2	48	96	4
	53-78	48	5	43	86	14
	79-91	43	2	41	82	18
	92-110	41	14	27	54	46
	FINALKILL111 -112	27	27	0		
LOW	0-52	50	1	49	98	2
	53-78	49	2	47	94	6
	79-91	47	4	43	86	14
	92-110	43	15	28	56	44
	FINALKILL111 -112	28	28	0		
MED	0-52	50	2	48	96	4
	53-78	48	7	41	82	18
	79-91	41	5	36	72	28
	92-110	36	11	25	50	50
	FINALKILL111 -112	25	25	0		
HIGH	0-52	50	2	48	96	4
	53-78	48	2	46	92	8
	79-91	46	3	43	86	14
	92-110	43	18	25	50	50
	FINALKILL111 -112	25	25	0		

Appendix 4: Mortality of Female Rats

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTR0	0-52	50	1	49	98	2
	53-78	49	8	41	82	18
	79-91	41	8	33	66	34
	92-104	33	6	27	54	46
	FINALKILL105 -105	27	27	0		
LOW	53-78	50	5	45	90	10
	79-91	45	9	36	72	28
	92-104	36	14	22	44	56
	FINALKILL105 -105	22	22	0		
MED	0-52	50	3	47	94	6
	53-78	47	3	44	88	12
	79-91	44	4	40	80	20
	92-104	40	9	31	62	38
	FINALKILL105 -105	31	31	0		
HIGH	53-78	50	6	44	88	12
	79-91	44	9	35	70	30
	92-104	35	11	24	48	52
	FINALKILL105 -105	24	24	0		

Appendix 5: Tumor Trend Test of Male Rats

Organ Code	Organ Name	Tumor Code	Tumor Name	CTR0	LOW	MED	HIGH	P-Value (Exact Method)	P-Value (Asymptotic Method)
C12	Heart	M61	Sarcoma	0	1	1	0	0.6319	0.7477
D12	Jaw	871	Carcinoma, squamous cell	0	0	0	1	0.2721	0.0586
D41	Small intestine, duodenum	6	Adenocarcinoma	0	1	0	0	0.7351	0.7573
D43	Small intestine, ileum	6	Adenocarcinoma	1	0	0	0	1	0.7966
E1	Pituitary gland	4	Adenoma	23	26	27	22	0.8051	0.8045
E1	Pituitary gland	Z53	Neurofibrosarcoma	1	0	0	0	1	0.8026
E3	Adrenal gland	Z91	Phaeochromocytoma, benign	14	8	11	11	0.5519	0.554
E4	Thyroid gland	4	Adenoma	2	2	0	3	0.2439	0.2266
E4	Thyroid gland	491	Adenoma, papillary cystic	2	1	3	2	0.3895	0.4139
E4	Thyroid gland	6	Adenocarcinoma	0	2	1	1	0.3872	0.446
E4	Thyroid gland	E4	Adenoma, "light cell" solid	1	4	6	4	0.3474	0.345
E5	Parathyroid gland	4	Adenoma	0	1	0	0	0.8462	0.749
G11	Testis	ML1	Leydig cell tumor, benign	12	8	13	9	0.68	0.6818
G11	Testis	MM1	Mesothelioma	1	0	0	0	1	0.7966
G12	Epididymis	MM1	Mesothelioma	0	0	0	1	0.3103	0.076
G21	Prostate	4	Adenoma	0	1	0	0	0.7429	0.7488
G21	Prostate	415	Adenoma, papillary	0	0	1	0	0.4762	0.5773
G21	Prostate	892	Carcinosarcoma	0	0	2	0	0.4909	0.6117
H1	Spleen	M21	Fibroma	0	1	1	0	0.6044	0.7233
H1	Spleen	MV1	Hemangioendothelioma	0	1	0	2	0.0971	0.0507
H1	Spleen	MV2	Hemangioendothelial sarcoma	0	0	0	2	0.0665	0.0103
H39	Lymph node(s), mesenteric	MV1	Hemangioendothelioma	1	3	4	1	0.7554	0.7623
H4	Hematopoietic system	H12	Lymphoid leukemia	0	2	1	0	0.7136	0.8041
H4	Hematopoietic system	H121	Lymphoid leukemia, lymphocytic	0	1	2	0	0.6497	0.7413
H4	Hematopoietic system	H129	Lymphoid leukemia, monocytic	0	1	0	0	0.7429	0.7488
H4	Hematopoietic system	H161	Lymphoma, lymphocytic	0	1	0	0	0.7586	0.7753
H4	Hematopoietic system	H432	Reticulosis, malignant	0	0	0	1	0.2487	0.0485
H4	Hematopoietic system	H62	Histiocytic sarcoma	0	1	0	0	0.7324	0.7604
I1	Skin	21	Papilloma	3	2	9	4	0.423	0.4322

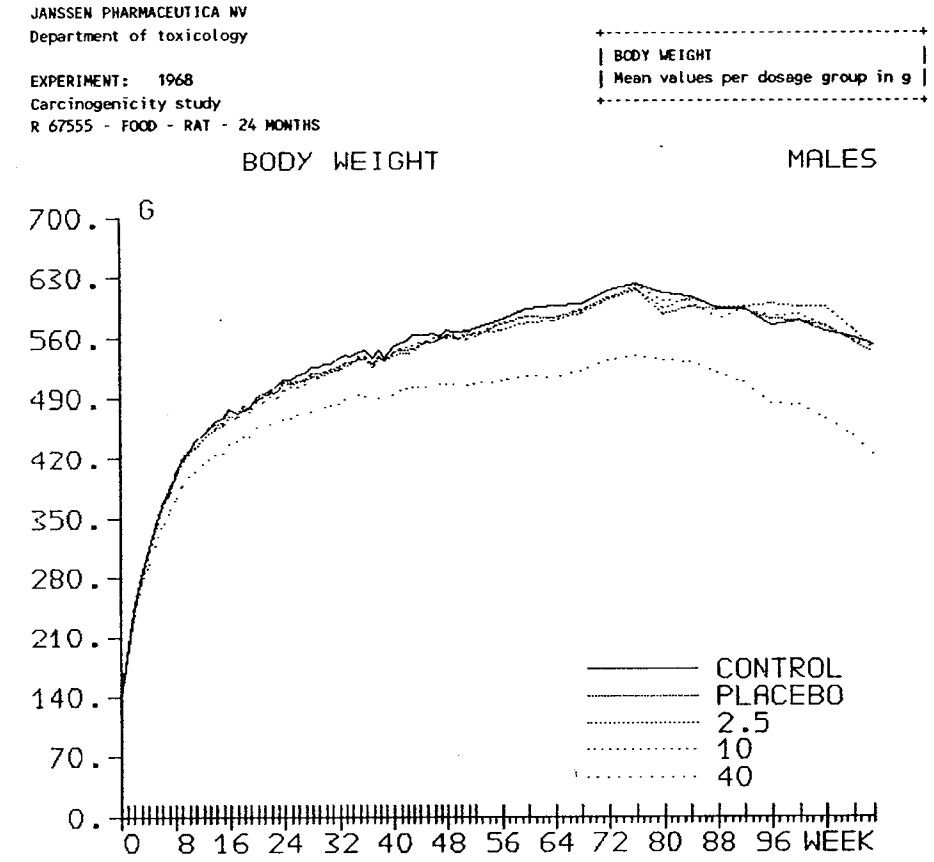
I1	Skin	853	Carcinoma, sebaceous squamous	1	0	0	0	1	0.8034
I1	Skin	871	Carcinoma, squamous cell	0	2	1	0	0.7426	0.827
I1	Skin	M21	Fibroma	0	1	0	0	0.7429	0.7488
I2	Mammary gland	411	Adenoma, acinar	0	0	1	0	0.4762	0.5773
I2	Mammary gland	441	Fibroadenoma	1	1	1	0	0.822	0.842
I2	Mammary gland	442	Adenofibroma	1	0	0	0	1	0.7966
I2	Mammary gland	6	Adenocarcinoma	0	1	0	0	0.7586	0.7753
L1	Liver	L1	Hepatoma	10	13	17	2	0.9993	0.9987
L1	Liver	L2	Hepatocytic carcinoma	1	4	1	0	0.9469	0.9398
L1	Liver	MV8	Hemangioma	0	0	0	1	0.2381	0.0438
M8	Soft tissue	M11	Lipoma	0	1	1	0	0.6319	0.7477
M8	Soft tissue	M21	Fibroma	0	1	2	1	0.3221	0.3834
M8	Soft tissue	M24	Fibrosarcoma	0	1	0	0	0.7574	0.7597
M8	Soft tissue	M241	Fibrosarcoma, histiocytic type	0	1	0	1	0.3338	0.2479
M8	Soft tissue	M61	Sarcoma	0	3	1	2	0.3114	0.3338
M8	Soft tissue	M613	Sarcoma, cystic	0	0	2	0	0.5028	0.6213
M8	Soft tissue	MV1	Hemangioendothelioma	1	1	1	0	0.8132	0.8292
M8	Soft tissue	MV2	Hemangioendothelial sarcoma	0	0	1	0	0.4762	0.5773
M8	Soft tissue	MV8	Hemangioma	1	0	0	0	1	0.7966
N1	Brain	Z312	Astrocytoma, malignant	0	1	0	0	0.7513	0.7551
N1	Brain	Z321	Oligodendroglioma, benign	0	0	1	0	0.4916	0.5927
N1	Brain	Z323	Mixed astrocytoma-oligodendrog	0	0	0	1	0.2778	0.0611
N1	Brain	Z41	Granular cell tumor, benign	1	0	0	0	1	0.8011
N1	Brain	Z812	Meningioma, malignant	0	0	0	1	0.2487	0.0485
O2	Ear	853	Carcinoma, sebaceous squamous	0	0	1	0	0.4898	0.5999
P	Pancreas	492	Adenoma, exocrine	16	16	10	7	0.9916	0.99
P	Pancreas	493	Adenoma, endocrine	5	10	4	5	0.7758	0.7769
P	Pancreas	661	Adenocarcinoma, exocrine	3	3	1	0	0.9814	0.9619
P	Pancreas	662	Adenocarcinoma, endocrine	1	0	0	0	1	0.7966
P	Pancreas	893	Carcinoma, exocrine	0	1	0	0	0.7429	0.7488
S11	Salivary gland(s), parotid gla	4	Adenoma	0	0	1	0	0.7188	0.8225
T7	Tail	21	Papilloma	0	1	0	0	0.7429	0.7488
U1	Kidney	418	Adenoma, tubular	0	0	1	0	0.4762	0.5773

Appendix 6: Tumor Trend Test of Female Rats

Organ Code	Organ Name	Tumor Code	Tumor Name	CTR0	LOW	MED	HIGH	P-Value (Exact Method)	P-Value (Asymptotic Method)
C12	Heart	M61	Sarcoma	0	1	0	0	0.777	0.7595
D11	Mouth	M24	Fibrosarcoma	0	0	0	1	0.2539	0.0506
D12	Jaw	871	Carcinoma, squamous cell	0	1	0	0	0.7459	0.7522
D41	Small intestine, duodenum	M21	Fibroma	0	1	0	0	0.7404	0.7568
D41	Small intestine, duodenum	M74	Fibroleiomyosarcoma	0	1	0	0	0.7647	0.7566
E1	Pituitary gland	4	Adenoma	36	35	27	22	0.9974	0.9968
E1	Pituitary gland	6	Adenocarcinoma	0	0	0	1	0.2273	0.0401
E3	Adrenal gland	4	Adenoma	0	1	2	2	0.1469	0.1584
E3	Adrenal gland	Z91	Phaeochromocytoma, benign	2	2	2	2	0.5131	0.5144
E4	Thyroid gland	4	Adenoma	1	1	0	0	0.9345	0.8602
E4	Thyroid gland	491	Adenoma, papillary cystic	0	0	0	2	0.0515	0.0064
E4	Thyroid gland	6	Adenocarcinoma	0	1	0	0	0.7404	0.7568
E4	Thyroid gland	E4	Adenoma, "light cell" solid	3	4	5	1	0.908	0.9035
E4	Thyroid gland	E8	Carcinoma, "light cell" solid	0	1	0	0	0.7404	0.7568
G31	Ovary	G44	Granulosa-theca cell tumor, be	1	0	0	0	1	0.7847
G33	Uterus	422	Polyp	11	3	6	2	0.9849	0.9788
G33	Uterus	6	Adenocarcinoma	0	1	3	0	0.6781	0.7288
G33	Uterus	M74	Fibroleiomyosarcoma	0	1	0	0	0.7404	0.7568
G34	Cervix	422	Polyp	0	0	0	1	0.2308	0.0414
G34	Cervix	891	Carcinoma, poorly differentiat	0	1	0	0	0.7644	0.7603
G34	Cervix	M61	Sarcoma	1	0	0	0	1	0.8054
G35	Vagina	422	Polyp	1	0	0	1	0.4423	0.2866
G35	Vagina	M21	Fibroma	1	0	0	0	1	0.7847
G37	Clitoral gland	871	Carcinoma, squamous cell	0	1	0	0	0.85	0.7717
H4	Hematopoietic system	H111	Lymphosarcoma, lymphocytic	0	2	1	0	0.7728	0.8325
H4	Hematopoietic system	H121	Lymphoid leukemia, lymphocytic	2	1	0	0	0.9854	0.9138
H4	Hematopoietic system	H152	Thymoma, predominantly lymphoc	2	0	7	2	0.4494	0.4582
I2	Mammary gland	4	Adenoma	0	1	1	0	0.6242	0.7327
I2	Mammary gland	411	Adenoma, acinar	0	0	0	1	0.2308	0.0414

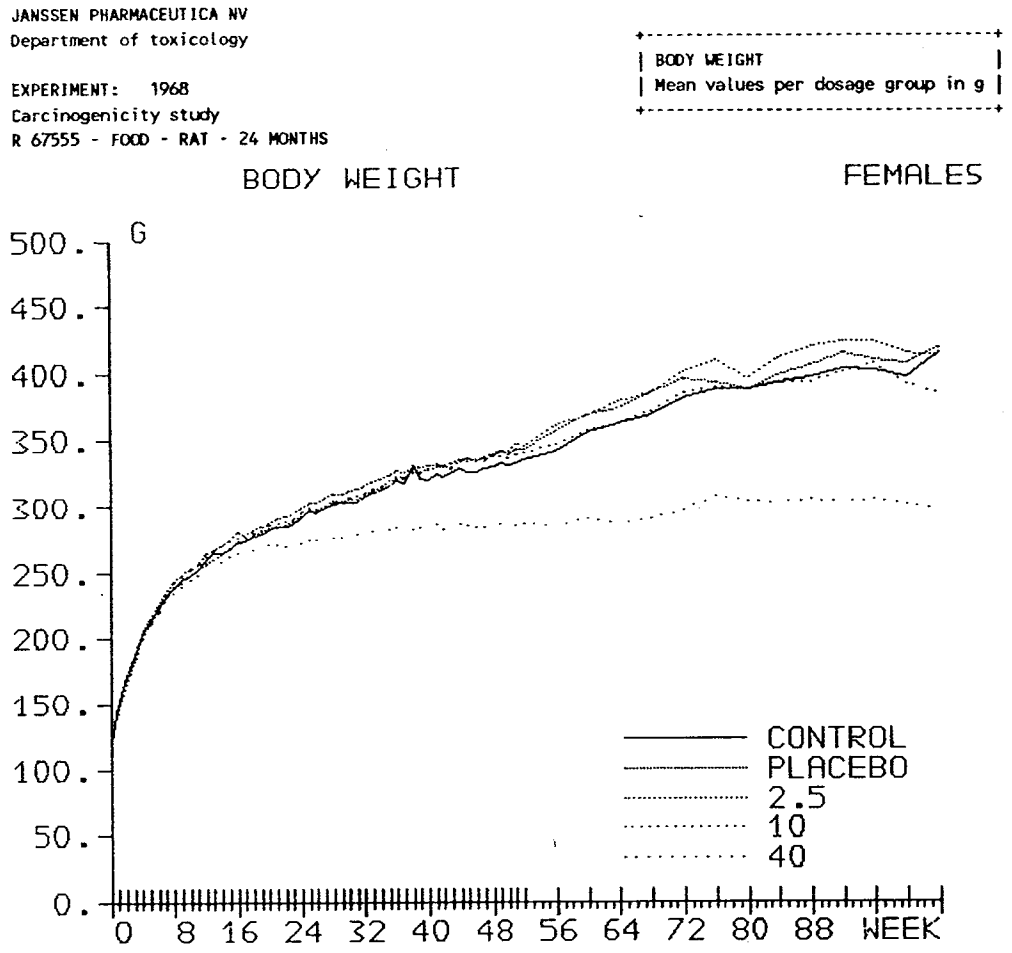
I2	Mammary gland	415	Adenoma, papillary	0	1	0	0	0.85	0.7717
I2	Mammary gland	441	Fibroadenoma	20	21	22	10	0.9974	0.9967
I2	Mammary gland	442	Adenofibroma	6	7	13	7	0.5087	0.5144
I2	Mammary gland	621	Adenocarcinoma, acinar	0	0	0	1	0.275	0.0596
I2	Mammary gland	625	Adenocarcinoma, papillary	4	2	1	1	0.8852	0.867
I2	Mammary gland	892	Carcinosarcoma	1	0	0	0	1	0.802
I2	Mammary gland	M21	Fibroma	0	1	3	0	0.6997	0.7437
L1	Liver	4L22	Cholangioma, cystic	0	0	1	0	0.5	0.6118
L1	Liver	L1	Hepatoma	10	11	9	1	0.9997	0.9991
L1	Liver	L2	Hepatocytic carcinoma	1	3	2	0	0.9019	0.9101
M8	Soft tissue	M241	Fibrosarcoma, histiocytic type	1	1	0	0	0.9385	0.8572
M8	Soft tissue	MV2	Hemangioendothelial sarcoma	0	1	0	0	0.7684	0.7604
O2	Ear	871	Carcinoma, squamous cell	1	0	0	0	1	0.8046
P	Pancreas	492	Adenoma, exocrine	2	2	1	0	0.9553	0.9288
P	Pancreas	493	Adenoma, endocrine	0	1	3	2	0.1711	0.1813
T7	Tail	21	Papilloma	1	0	0	0	1	0.8154
U1	Kidney	418	Adenoma, tubular	0	1	0	0	0.7333	0.7584
U1	Kidney	626	Adenocarcinoma, tubular	0	0	1	0	0.5288	0.5832
U1	Kidney	M11	Lipoma	0	1	1	0	0.6613	0.7462

Appendix 7: Graph of Mean Body Weight of Male Rats



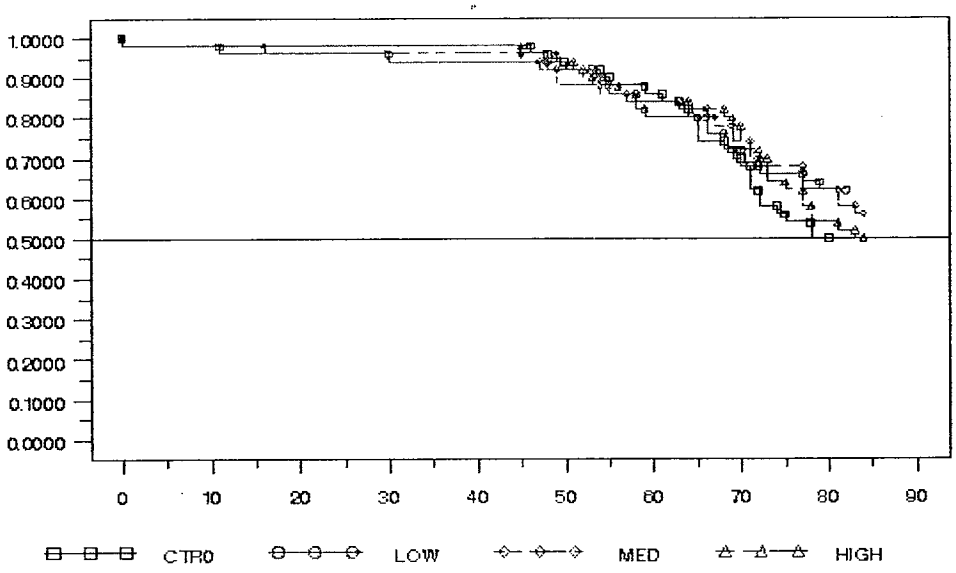
Appears This Way
On Original

Appendix 8: Graph of Mean Body Weight of Female Rats

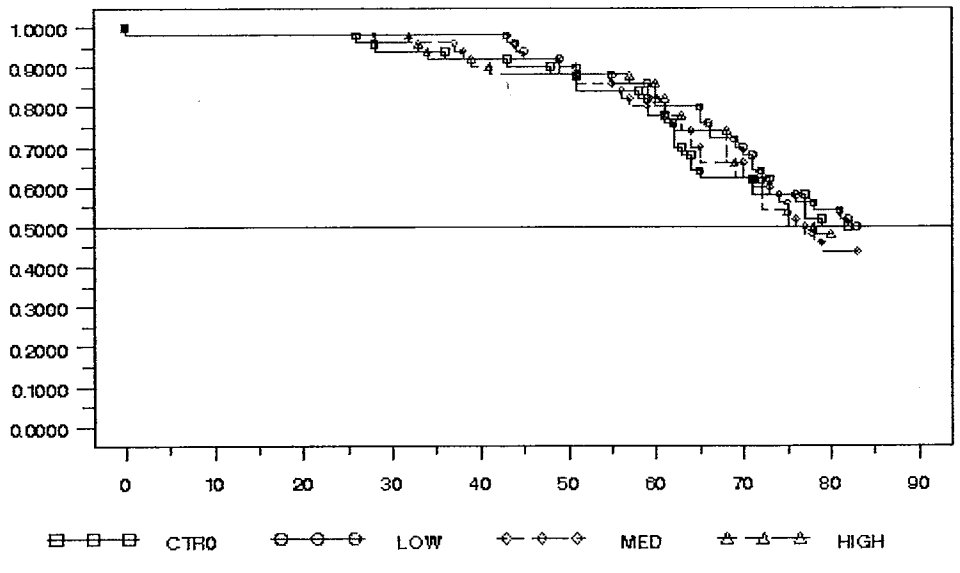


Appears This Way
On Original

Appendix 9: Survival Graph of Male Mice



Appendix 10: Survival Graph of Female Mice



Appendix 11: Mortality of Male Mice

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTR0	0-50	50	3	47	94	6
	51-78	47	20	27	54	46
	79-84	27	2	25	50	50
	FINALKILL 85-85	25	25	0		
LOW	0-50	50	3	47	94	6
	51-78	47	14	33	66	34
	79-84	33	2	31	62	38
	FINALKILL 85-85	31	31	0		
MED	0-50	50	4	46	92	8
	51-78	46	12	34	68	32
	79-84	34	6	28	56	44
	FINALKILL 85-85	28	28	0		
HIGH	0-50	50	2	48	96	4
	51-78	48	19	29	58	42
	79-84	29	4	25	50	50
	FINALKILL 85-85	25	25	0		

Appendix 12: Mortality of Female Mice

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTR0	0-50	50	5	45	90	10
	51-78	45	16	29	58	42
	79-83	29	4	25	50	50
	FINALKILL 84-85	25	25	0		
LOW	0-50	50	4	46	92	8
	51-78	46	18	28	56	44
	79-83	28	3	25	50	50
	FINALKILL 84-85	25	25	0		
MED	0-50	50	4	46	92	8
	51-78	46	22	24	48	52
	79-83	24	2	22	44	56
	FINALKILL 84-85	22	22	0		
HIGH	0-50	50	5	45	90	10
	51-78	45	20	25	50	50
	79-83	25	1	24	48	52
	FINALKILL 84-85	24	24	0		

Appendix 13: Tumor Trend of Male Mice

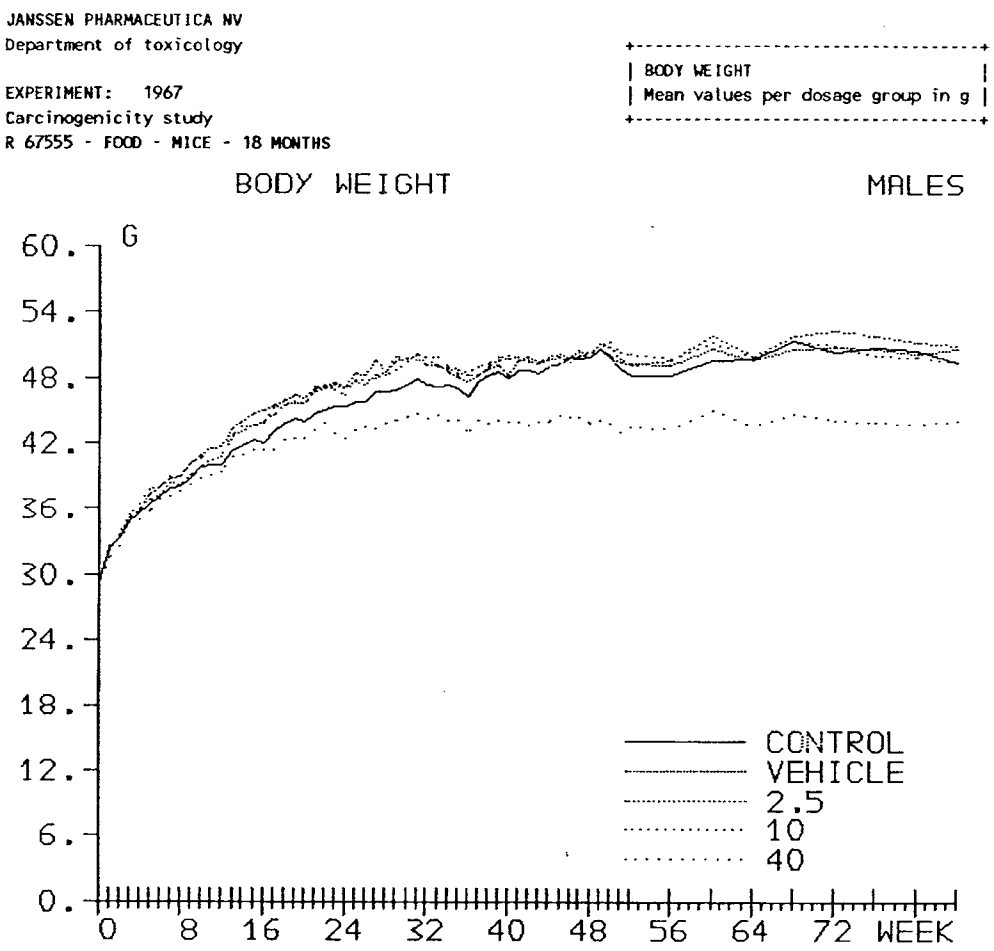
Organ Code	Organ Name	Tumor Code	Tumor Name	CTR0	LOW	MED	HIGH	P-Value (Exact Method)	P-Value (Asymptotic Method)
D51	Large intestine, cecum	636	Adenocarcinoma, signet-ring ce	0	0	1	0	0.525	0.5929
G11	Testis	ML1	Leydig cell tumor, benign	2	0	1	21	0	0
G12	Epididymis	ML1	Leydig cell tumor, benign	0	0	1	0	0.4862	0.5757
G17	Scrotum	M241	Fibrosarcoma, histiocytic type	1	0	0	0	1	0.7988
H4	Hematopoietic system	H11	Lymphosarcoma	1	1	1	0	0.8279	0.8303
H4	Hematopoietic system	H12	Lymphoid leukemia	0	0	0	1	0.25	0.049
H4	Hematopoietic system	H122	Lymphoid leukemia, lymphoblast	1	0	0	0	1	0.8097
H4	Hematopoietic system	H15	Thymoma	0	0	1	0	0.5025	0.5925
L1	Liver	L1	Hepatic neoplastic nodule	5	8	6	2	0.9622	0.9557
L1	Liver	L2	Hepatocytic carcinoma	4	4	4	0	0.9838	0.9746
L1	Liver	L3	Hepatoblastoma	0	1	0	0	0.7706	0.7504
L1	Liver	MV1	Hemangioendothelioma	1	1	2	0	0.823	0.8513
M1	Bone	M93	Osteosarcoma	0	1	0	0	0.7475	0.7578
M8	Soft tissue	MV2	Hemangioendothelial sarcoma	0	0	1	0	0.525	0.5929
N3	Peripheral nervous system	Z52	Neurofibroma	0	1	0	0	0.7461	0.7576
P	Pancreas	493	Adenoma, endocrine	0	1	0	0	0.7706	0.7504
R2	Lung	R1	Primary lung tumor, benign	10	18	12	6	0.985	0.9822
R2	Lung	R2	Primary lung tumor, malignant	5	9	3	3	0.9018	0.8963
U1	Kidney	626	Adenocarcinoma, tubular	1	0	0	0	1	0.7988

Appendix 14: Tumor Trend of Female Mice

Organ Code	Organ Name	Tumor Code	Tumor Name	CTR0	LOW	MED	HIGH	P-Value (Exact Method)	P-Value (Asymptotic Method)
D11	Mouth	871	Carcinoma, squamous cell	0	0	0	1	0.25	0.0488
D3	Stomach	422	Polyp, benign	0	0	0	1	0.25	0.0488
D3	Stomach	871	Carcinoma, squamous cell	1	0	0	0	1	0.7991
E1	Pituitary gland	4	Adenoma	0	0	0	1	0.2679	0.057
G31	Ovary	491	Adenoma, papillary cystic	0	0	0	1	0.2533	0.0509
G31	Ovary	G43	Granulosa-theca cell tumor, ma	0	1	0	0	0.7867	0.77
G31	Ovary	G51	Luteal cell tumor, benign	0	1	0	0	0.7396	0.7526
G31	Ovary	M71	Leiomyoma	1	0	0	0	1	0.7991
G31	Ovary	MV1	Hemangioendothelioma	0	1	1	0	0.6444	0.7515
G31	Ovary	MV8	Hemangioma	1	2	1	1	0.532	0.5532
G33	Uterus	422	Polyp, benign	3	4	2	3	0.497	0.503
G33	Uterus	622	Polyp, malignant	2	0	2	0	0.8361	0.8637
G33	Uterus	8	Carcinoma	0	0	1	0	0.4959	0.5902
G33	Uterus	M61	Sarcoma	0	0	1	0	0.4792	0.585
G33	Uterus	M71	Leiomyoma	0	1	0	0	0.7396	0.7526
G33	Uterus	M74	Fibroleiomyosarcoma	0	0	1	0	0.4701	0.5708
G33	Uterus	MV8	Hemangioma	2	0	0	0	1	0.881
G34	Cervix	M61	Sarcoma	0	0	1	0	0.4792	0.585
G34	Cervix	MV2	Hemangioendothelial sarcoma	0	0	1	0	0.4623	0.5719
G35	Vagina	422	Polyp, benign	0	1	0	0	0.7474	0.7548
H4	Hematopoietic system	H11	Lymphosarcoma	1	2	1	2	0.3511	0.3464
H4	Hematopoietic system	H12	Lymphoid leukemia	2	3	2	1	0.7929	0.804
H4	Hematopoietic system	H123	Lymphoid leukemia, lymphocytic	0	1	0	0	0.756	0.7611
H4	Hematopoietic system	H15	Thymoma	0	3	2	1	0.542	0.6098
H4	Hematopoietic system	H21	Myeloid leukemia	0	0	1	2	0.0603	0.034
H4	Hematopoietic system	H62	Histiocytic sarcoma	1	1	2	0	0.7824	0.8267
I2	Mammary gland	6	Adenocarcinoma	1	1	2	0	0.7687	0.8188
L1	Liver	L1	Hepatic neoplastic nodule	4	1	2	0	0.9513	0.9408
L1	Liver	L2	Hepatocytic carcinoma	1	1	0	0	0.9342	0.8552
L1	Liver	MV1	Hemangioendothelioma	1	2	1	1	0.5724	0.5849
M8	Soft tissue	M24	Fibrosarcoma	0	0	1	0	0.4701	0.5708
M8	Soft tissue	MV2	Hemangioendothelial sarcoma	0	0	1	0	0.3	0.4069

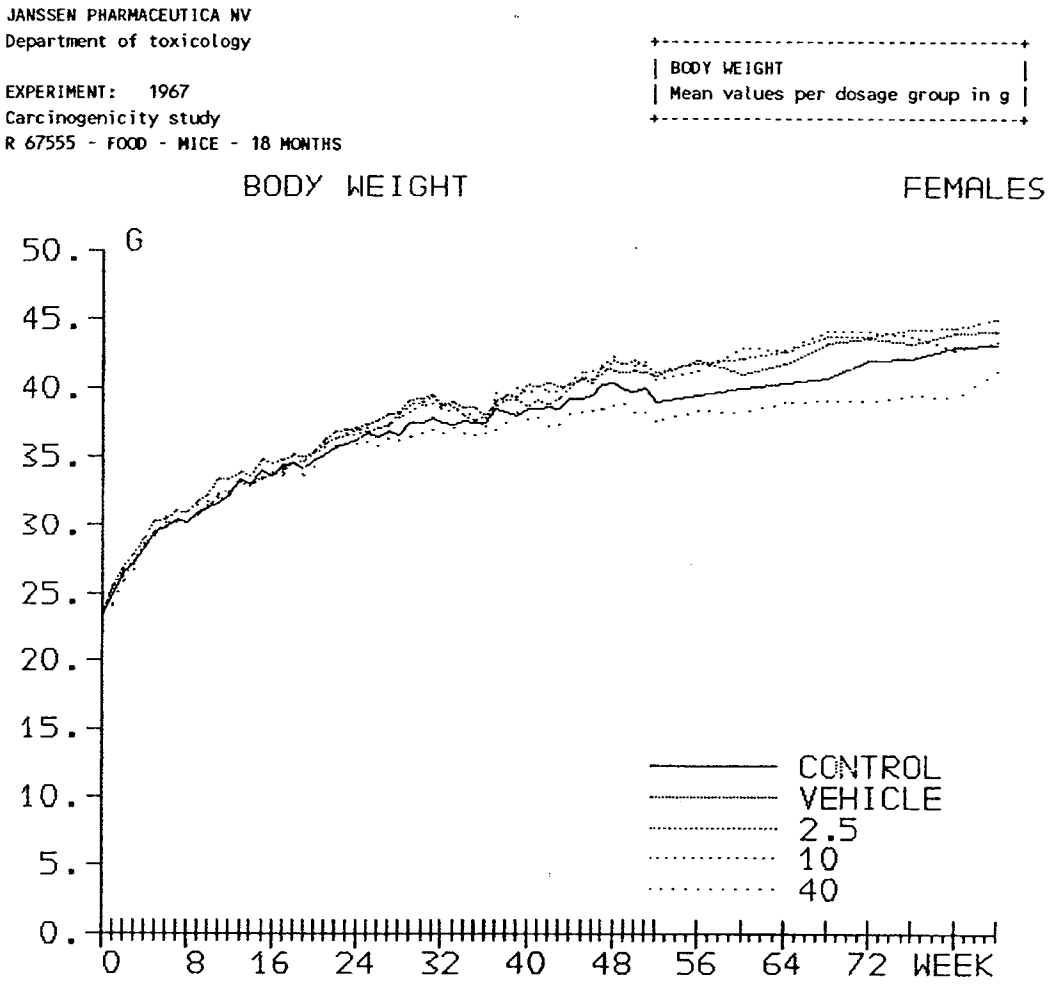
N1	Brain	Z36	Tumor of glia - glioma, malign	0	0	0	1	0.2558	0.0515
R2	Lung	R1	Primary lung tumor, benign	12	5	5	6	0.8146	0.8149
R2	Lung	R2	Primary lung tumor, malignant	2	2	2	3	0.1924	0.1872

Appendix 15: Graph of Mean Body Weight of Male Mice



Appears This Way
On Original

Appendix 16: Graph of Mean Body Weight of Female Mice



Appears This Way
On Original

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

/s/

Jasmine Choi
7/21/04 11:33:37 AM
BIOMETRICS

Roswitha Kelly
7/21/04 02:45:30 PM
BIOMETRICS

Kooros Mahjoob
7/22/04 07:25:50 PM
BIOMETRICS