focus on the two variables, tumor response and time to progression.

(i) Tumor Response

The sponsor distinguishes two types of tumor response; one is called "peer reviewed confirmed best overall tumor response", referred to as overall response, and the other is called "peer reviewed confirmed overall complete or partial tumor response", referred to as objective response.

The tumor response was evaluated "at baseline, 3 months after the start of the trial treatment and every 3 months thereafter, or when the patient discontinued treatment at or after 3 months."

Reviewers' TABLE 6.1 shows three types of response rate for each arm. The numbers in ( ) indicate response rate percentages for each type of response for each treatment group. The statistical analyses will be based on overall response confirmed by peer reviewers.

Reviewers' TABLE 6.1: Types of Response Rate in Each Treatment Arm

	Treatment Arm		
	0.5 mg	2.5 mg	AG
Total Sample Size	192	185	178
Overall Response*	32 (16.7%)	33 (17.8%)	20 (11.2%)
Objective Response**	39 (20.3%)	36 (19.5%)	26 (14.6%)
Investigator's Assessment***	32 (16.7%)	32 (17.3%)	29 (16.3%)

Note:(i) overall response rate was derived by peer reviewed confirmed best overall objective tumor response.

(ii) objective response rate was derived by peer reviewed confirmed best overall objective tumor response (whether confirmed or not)

(iii) investigator's assessment was derived by confirmed tumor response.

Note: the figures are adapted from Sponsor's Table 8.1-1.1, Table 8.1-1.3, and Table 8.1-1.4.

Logistic regression analyses were applied to compare the overall response rate between two treatment groups. Reviewer's TABLE 6.2 shows the results derived from the logistic regression analyses with and without covariates.

Reviewers' TABLE 6.2: The Results derived from a logistic regression with and without covariates

(Sponsor's Table 8.1-1.2)

	Treatment Comparison		
	0.5mg vs 2.5mg	0.5mg vs AG	2.5mg vs AG
Adjusted Odds Ratio	0.99	1.97	1.78
95% CI	(0.56, 1.75)	(1.00, 3.89)	(0.91, 3.48)
p-value	0.9748	0.0512	0.0918
Unadjusted Odds Ratio	0.92	1.58	1.72
95% CI	(0.54, 1.57)	(0.87, 2.88)	(0.94, 3.12)
p-value	0.7635	0.1354	0.0772

Note: 0.5 mg = 0.5 mg letrozole
2.5 mg = 2.5 mg letrozole
AG = 500 mg aminoglutethimide

Odds are defined as the ratio of the probability of response over the probability of nonresponse in a treatment group, and an odds ratio is defined as the ratio of the odds of two treatments. In the 0.5 mg vs. 2.5 mg comparison, the 2.5 mg treatment arm is considered as the base and in the other comparisons, the aminoglutethimide treatment group is treated as the base.

<u>0.5 mg vs 2.5 mg of letrozole</u>

There was no statistically significant effect in favor of 2.5 mg letrozole over 0.5 mg letrozole in both unadjusted and adjusted analyses. The odds ratio of overall response with 0.5 mg letrozole over 2.5 mg letrozole was 0.92 (95% CI: [0.54, 1.57], P=0.7635 in unadjusted analysis).

<u>0.5 mg of letrozole vs aminoglutethimide</u>

There was no statistically significant effect in favor of 0.5 mg letrozole over aminoglutethimide in unadjusted analysis (P=0.1354), but there was a statistically significant effect in favor of 0.5 mg letrozole over aminoglutethimide with an estimated odds ratio of 1.97 (95% CI: [1.00, 3.89], P=0.0512) in the adjusted logistic regression analysis.

2.5 mg of letrozole vs aminoglutethimide

There was no statistically significant effect in favor of 2.5 mg letrozole over aminoglutethimide in both unadjusted and adjusted analyses. The odds ratio of overall response with 2.5 mg letrozole over aminoglutethimide was 1.72 (95% CI: [0.94, 3.12], P=0.0772).

Note that no adjustments to the significance level were made for multiple comparisons.

(ii) Time to Progression (TTP)

Reviewers' TABLE 6.3 shows the total sample size and the number of censored subjects in each treatment arm for the time to progression (TTP) analysis.

Reviewers' TABLE 6.3: Total Sample Size and the Number of Censored Subjects in Each Treatment Arm (Adapted from Sponsor's Table 8.1 - 1.7) for TTP

	Treatment Arm		
	0.5 mg	2.5 mg	AG
Total Sample Size	192	185	178
# of Censored Patient	51 (26.5%)	61 (33.0%)	39 (21.9%)

Reviewers' TABLE 6.4 shows the results from bot both unadjusted and adjusted analyses by Cox regression analyses. Note that in adjusted analyses all prognostic factors were included in a Cox regression model.

Reviewers' TABLE 6.4: Unadjusted and Adjusted Relative Risks with Corresponding 95% CI and P-Values (Adapted from Sponsor's Table 8.1 - 1.6) for TTP

	Treatment Comparison		
	0.5 mg vs 2.5 mg	0.5 mg vs AG	2.5 mg vs AG
Adjusted Relative Risk	1.13	0.76	0.68
95% CI	(0.88, 1.46)	(0.59, 0.98)	(0.53, 0.88)
p-value	0.3373	0.0331	0.0035
Unadjusted Relative Risk	1.12	0.86	0.77
95% CI	(0.88, 1.42)	(0.68, 1.09)	(0.60, 0.98)

p-value	0.3717	0.2035	0.0366

0.5 mg vs 2.5 mg of letrozole

There was no statistically significant difference (compared to 1.0) in relative risk in either adjusted or unadjusted analyses (P=0.3373 and P=0.3717, respectively).

0.5 mg of letrozole vs aminoglutethimide

There was no statistically significant difference (compared to 1.0) in relative risk in unadjusted analysis (P=0.2035), but there was statistical significance in favor of 0.5 mg letrozole over aminoglutethimide with an estimated relative risk of 0.76 (95% CI: [0.59, 0.98], P=0.0331) in the adjusted Cox regression analysis.

2.5 mg of letrozole vs aminoglutethimide

There was a statistically significant difference (compared to 1.0) in relative risk in both unadjusted and adjusted analysis, favoring 2.5 mg of letrozole over aminoglutethimide. The estimated relative risk with 2.5 mg of letrozole over aminoglutethimide was 0.77 (95% CI: [0.60, 0.98], P=0.0366)

(II) Secondary Variables

Performance status, severity of pain and quality of life variables were measured over the study period as secondary variables. Reviewers' TABLE 6.5 presents the sample size for each treatment group over the study period. It is noted that approximately 50% of the subjects in each treatment group dropped out of the study by the end of 6 months.

Reviewers' TABLE 6.5: Missing Data Pattern Over the Study Period in Performance Status / Sample Size Changes (Adapted from Sponsor's Table 8.1-2.1)

	0.5 mg letrozole	2.5 mg letrozole	Aminoglutethimide
baseline	192	185	178
2 months	183	169	161
6 months	102	98	96
9 months	75	77	67
12 months	48	53	43

Formal statistical analyses were not performed by the sponsor.

Reviewers' Comments and Conclusions for the AR/BC 3 Study:

In this study these reviewers identified two major statistical issues, covariate adjustments in logistic regression and Cox regression analyses for the primary variables, and problematic longitudinal analyses for secondary variables. These issues were discussed in detail in section IV so that this reviewer will discuss them only briefly in this section.

These reviewers see two major statistical issues in covariate adjustments in both models. They are (i) misspecification of a model and (ii) a stability issue in parameter estimates as described in Section IV. For the first issue we can treat a model with covariates in both logistic and Cox models as a "working" model, since we do not know the **true** model. Therefore, parameter estimates derived from a "working" likelihood may not converge to the **true** value. To test for a treatment effect estimated from the "working" model, a "sandwich" estimate of the variance of the estimated treatment effect coefficient should be applied in order to preserve Type I error at the 0.05 level (Kent, 1982, and Lin and Wei, 1988).

For the second issue parameter estimates (regression coefficients and associated standard errors) may not be stable because the needed homogeneity assumption of odds ratios across strata for a logistic regression and proportinal hazards assumption for a Cox regression model were assumed in the applied models and many covariates were adjusted for in both models. As these reviewers discussed in Section IV, parsimonious models should be investigated to relax those assumptions and to reduce the number of covariates to be adjusted.

Reviewers' APPENDIX 6.1 shows estimated odds ratios, 95% CIs and associated p-values in each category stratified by hormone receptor status for each treatment comparison. In the comparison of 0.5 mg letrozole vs 2.5 mg letrozole treatment arms, hormone receptor status can be considered as an 'effect modification' factor and no statistically significant results were found for each hormone receptor status category. This result is consistent with the adjusted and unadjusted results reported by the sponsor. In the comparison of 0.5 mg letrozole vs aminoglutethimide treatment arms, this reviewer did not consider hormone receptor status as an 'effect modification' factor because the estimated odds ratios in RS =1 and RS =2 categories are similar even though the estimated odds ratio in RS = 3 is different. The result from the Mantel-Haenszel approach was not statistically significant. Note that no statistically significant results were found for each hormone receptor status category. In the comparison of 2.5 mg letrozole vs aminoglutethimide treatment arms, hormone receptor status can be considered as an 'effect modification' factor and no statistically significant results were found for each hormone receptor status category.

Reviewers' APPENDIX 6.2 shows estimated risk ratios and p-values by logrank tests in each category stratified by a hormone receptor status and associated p-value by a stratified logrank test in each treatment comparison. In the comparison of 0.5 mg letrozole vs 2.5 mg letrozole treatment arms and in the comparison of 0.5 mg letrozole vs aminoglutethimide treatment arms, proportinal hazard assumption across hormone receptor status categories did not hold. On the other hand, in the comparison of 2.5 mg letrozole vs aminoglutethimide treatment arms the proportional hazard assumption did hold and a statistically significant treatment effect was found by a stratified logrank test.

In this submission the sponsor reported statistical testing results by unadjusted and adjusted analyses for logistic and Cox regressions. As discussed in this review, statistical issues in covariate adjustments for logistic and Cox regressions were identified and should be addressed. The sponsor should at least validate their final models.

Masahiro Takeuchi Sc.D.
Mathematical Statistican

Roswitha Kelly M.S.
Mathematical Statistician

Concur: Dr. Gnecco 6/25/97
Dr. Chi For George Chi 6/25/97

cc:
NDA#20-726
HFD-150 / Division File
HFD-150 / Dr. John Johnson (Medical Team Leader)
HFD-150 / Dr. Genevieve Schechter
HFD-150 / Dianne Spillman, CSO
HFD-344 / Dr. B. Borton
HFD-710 / Dr. Chi
HFD-710 / Dr. Gnecco
HFD-710 / Dr. Takeuchi
HFD-710 / Ms. Kelly
HFD-710 / Chron

This review consists of 30 pages of text, 23 Reviewer's Appendices, and 1 Reviewer's figure.
Takeuchi/ 12/03/96 / WP3.1 / statistical review (femara)

Reviewer's APPENDIX 4.1.1: Estimated Odds Ratio, 95%CI and Associated P-Values Under Several Selected Covariates' Scenarios

0.5 mg letrozole vs 2.5 mg letrozole Unadjusted vs Adjusted Odds Ratio*					
	OR	95% CI	Est.	SE	P-Value
Unadjusted	0.430	(0.244, 0.757)	-0.844	0.289	0.0035
[1]Adjusted_11	0.366	(0.197, 0.680)	-1.004	0.316	0.0015
[2]Adjusted_6	0.419	(0.234, 0.750)	-0.870	0.297	0.0034
[3]Adjusted_S	0.386	(0.214, 0.697)	-0.951	0.301	0.0016
[4]Adjusted_FDA	0.387	(0.215, 0.695)	-0.950	0.299	0.0015
[5]Adjusted_FDA*	0.384	(0.214, 0.691)	-0.956	0.299	0.0014

Odds ratio is defined as the ratio of the odds of the two treatment where odds can be defined as a probability of response over a probability of nonresponse in a treatment group

Note:(1) adjusted_11 - covariate adjustment with 11 covariates reported in this NDA submission.

(2) adjusted_6 - covariate adjustment with 6 covariates specified in the protocol dated Oct. 27, 1992.

(3) adjusted_S - covariate adjustment with covariates selected by a forward stepwise procedure without treatment effect in the procedure - **dominant site of disease** and **response to therapeutic anti-estrogen therapy** were selected.

(4) adjusted_FDA - covariate adjustment with two covariates -**hormone receptor status** and **dominant site of disease** - selected by a survey from FDA Division of Oncology

(5) adjusted_FDA* - covariates adjustment with two covariates - **hormone receptor status (ER vs Unknown)** and **dominant site of disease** - selected by FDA Division of Oncology

Reviewer's APPENDIX 4.1.2: Estimated Odds Ratio, 95%CI and Associated P-Values Under Several Selected Covariates' Scenarios

0.5 mg letrozole vs Megestrol Acetate Unadjusted vs Adjusted Odds Ratio*					
	OR	95%CI	Est	SE	P-Value
Unadjusted	0.675	(0.375, 1.216)	-0.392	0.300	0.1911
[1]Adjusted_11	0.546	(0.288, 1.035)	-0.605	0.326	0.0636
[2]Adjusted_6	0.668	(0.369, 1.209)	-0.404	0.303	0.1822
[3]Adjusted_S	0.544	(0.294, 1.006)	-0.609	0.314	0.0522
[4]Adjusted_FDA	0.615	(0.334, 1.132)	-0.486	0.311	0.1186
[5]Adjusted_FDA*	0.612	(0.332, 1.126)	-0.492	0.311	0.1141

Odds ratio is defined as the ratio of the odds of the two treatment where odds can be defined as a probability of response over a probability of nonresponse in a treatment group

Note:(1) adjusted_11 - covariate adjustment with 11 covariates reported in this NDA submission.

(2) adjusted_6 - covariate adjustment with 6 covariates specified in the protocol dated Oct. 27, 1992.

(3) adjusted_S - covariate adjustment with covariates selected by a forward stepwise procedure without treatment effect in the procedure - **performance status** and # **of sites of disease** were selected.

(4) adjusted_FDA - covariate adjustment with two covariates -**hormone receptor status** and **dominant site of disease** - selected by a survey from FDA Division of Oncology

(5) adjusted_FDA* - covariate adjustment with two covariates - **hormone receptor status (ER vs Unknown)** and **dominant site of disease** - selected by FDA Division of Oncology

Reviewer's APPENDIX 4.1.3: Estimated Odds Ratio, 95%CI and Associated P-Values Under Several Selected Covariates' Scenarios

2.5 mg letrozole vs Megestrol Acetate Unadjusted vs Adjusted Odds Ratio*					
	OR	95%CI	Est	SE	P-Value
Unadjusted	1.571	(0.934, 2.644)	0.452	0.266	0.0888
[1]Adjusted_11	1.805	(1.007, 3.235)	0.591	0.298	0.0472
[2]Adjusted_6	1.646	(0.966, 2.804)	0.499	0.272	0.0666
[3]Adjusted_S	1.655	(0.946, 2.895)	0.504	0.286	0.0777
[4]Adjusted_FDA	1.712	(0.986, 2.972)	0.538	0.281	0.0560
[5]Adjusted_FDA*	1.698	(0.979, 2.944)	0.530	0.281	0.0593

Odds ratio is defined as the ratio of the odds of the two treatment where odds can be defined as a probability of response over a probability of nonresponse in a treatment group

Note:(1) adjusted_11 - covariate adjustment with 11 covariates reported in this NDA submission.

(2) adjusted_6 - covariate adjustment with 6 covariates specified in the protocol dated Oct. 27, 1992.

(3) adjusted_S - covariate adjustment with covariates selected by a forward stepwise procedure without treatment effect in the procedure - **dominant site of disease** and **# of sites of disease** were selected.

(4) adjusted_FDA - covariate adjustment with two covariates -**hormone receptor status** and **dominant site of disease** - selected by a survey from FDA Division of Oncology

(5) adjusted_FDA* - covariate adjustment with two covariates - **hormone receptor status (ER vs Unknown)** and **dominant site of disease** - selected by FDA Division of Oncology

Reviewer's Appendix 4.1.4:

Estimated Odds Ratios, 95% CIs and Associated P-Values by Exact and Mantel-Haenszel Methods in Each Category Stratified by Hormone Receptor Status

0.5 mg letrozole vs 2.5 mg letrozole				
		RS=1*	RS=2*	RS=3*
	Est. OR	0.200	0.151	0.732
Exact	95%CI	(0.054, 0.636)	(0.003, 1..303)	(0.318, 1.681)
	P-value	P=0.0037	P=0.109	P=0.540
M-H*	95%CI	(0.068, 0.589)	(0.02, 1.295)	(0.342, 1.567)
	p-value	P=0.0039	P=0.0842	p=0.422
Homogeneity: BD*: p=0.086 Zelen:p=0.09 Note: Hormone receptor status can be considered as an 'effect modification' factor.				
	Est. OR	0.205		Stratified by dominant site of disease: ORsoft=0.545 ORbone=0.40 ORvis=4.632 Note: dominant site of disease also can be considered as an 'effect modification' factor.
Exact	95%CI	(0.065, 0.555)		
	p-value	P=0.0007		
M-H*	95%CI	(0.080, 0.528)		
	p-value	P=0.0014		
Stratified by dominant sites of disease				
	Est. OR	0.1706		
Exact	95%CI	(0.05, 0.472)		
	p-value	P=0.0002		
Note: estimated ORs are considered to be uniform across dominant sites of disease				

Note that M-H stands for Mantel-Haenszel approach.

Reviewer's Appendix 4.1.5:

Estimated Odds Ratios, 95% CIs and Associated P-Values by Exact and Mantel-Haenszel Methods in Each Category Stratified by Hormone Receptor Status

0.5 mg letrozole vs Megestrol Acetate					
		RS=1*	RS=2*	RS=3*	
	Est. OR	0.469	0.105	1.294	
Exact	95%CI	(0.119, 1.619)	(0.002, 0.848)	(0.528, 3.241)	
	P-value	P=0.287	P=0.028	P=0.685	
M-H*	95%CI	(0.151, 1.451)	(0.01, 0.873)	(0.569, 2.947)	
	p-value	P=0.189	P=0.0367	p=0.538	
Homogeneity: BD*: p=0.037 Zelen:p=0.047 Note: Hormone receptor status can be considered as an 'effect modification' factor.					
	Est. OR	0.297		Stratified by dominant site of disease: ORsoft=1.052 ORbone=4.211 ORvis=0.778 Note: dominant site of disease also can be considered as an 'effect modification' factor.	
Exact	95%CI	(0.09, 0.820)			
	p-value	P=0.016			
M-H*	95%CI	(0.113, 0.775)			
	p-value	P=0.0134			
Stratified by dominant sites of disease					
	Est. OR	0.268			
Exact	95%CI	(0.08, 0.762)			
	p-value	P=0.010			
Note: estimated ORs are considered to be uniform across dominant sites of disease					

Note that M-H stands for Mantel-Haenszel Approach.

Reviewer's Appendix 4.1.6:

Estimated Odds Ratios, 95% CIs and Associated P-Values by Exact and Mantel-Haenszel Methods in Each Category Stratified by Hormone Receptor Status

2.5 mg letrozole vs Megestrol Acetate				
		RS=1*	RS=2*	RS=3*
	Est. OR	2.341	0.691	1.768
Exact	95%CI	(0.890, 6.352)	(0.195, 2.373)	(0.730, 4.384)
	P-value	P=0.091	P=0.701	P=0.238
M-H*	95%CI	(0.967, 5.670)	(0.231, 2.070)	(0.784, 3.985)
	p-value	P=0.0589	P=0.509	p=0.170
Homogeneity: BD*: p=0.086 Zelen:p=0.151 Note: Hormone receptor status can be considered as an 'effect modification' factor.				
		Stratified by dominant site of disease: ORsoft=3.000 ORbone=3.167 ORvis=3.480 Homogeneity: BD*: p=0.992 Zelen : p=1.000 OR=3.105 95%CI : (1.093, 9.627) P=0.0313	Stratified by dominant site of disease: ORsoft=0.357 ORbone=2.100 ORvis=0.684 Homogeneity: BD*: p=0.705 Zelen : p=1.000 OR=0.444 95%CI : (0.070, 2.456) P=0.469 Note: To the reviewer dominant site of disease can be considered as an 'effect modification' factor.	Stratified by dominant site of disease: ORsoft=1.929 ORbone=0.909 ORvis=1.944 Homogeneity: BD*: p=0.880 Zelen : p=1.000 OR=1.790 95%CI : (0.706, 4.666) P=0.254

Note that M-H stands for Mantel-Haenszel Approach.

Reviewer's Appendix 4.2.1: Estimated Relative Risk, 95%CI, and Associated P-values along with Parameter Estimates and Estimated Standard Errors under Several Sets of Covariates

0.5 mg letrozole vs 2.5 mg letrozole					
	RR	95%CI	Est.	SE	P-Value
Unadjusted	1.26	(0.97, 1.64)	0.234	0.134	0.0813
[1]Adjusted_11	1.37	(1.05, 1.80)	0.316	0.138	0.0219
[2]Adjusted_8	1.41	(1.08, 1.85)	0.343	0.138	0.0130
[3]Adjusted_S	1.26	(0.97, 1.64)	0.232	0.135	0.0849
[4]Adjusted_FDA	1.29	(0.99, 1.68)	0.255	0.135	0.0599
[5]Adjusted_FDA*	1.31	(1.00, 1.70)	0.267	0.136	0.0492

Note:(1) adjusted_11 - covariate adjustment with 11 covariates reported in this NDA submission.

(2) adjusted_8 - covariate adjustment with 8 covariates specified in the protocol dated on Oct. 27, 1992.

(3) adjusted_S - covariate adjustment with covariates selected by a forward stepwise procedure without treatment effect in the procedure - **Age category, number of sites involved, disease free interval, performance status, and response to prior anti-estogen therapy** were selected.

(4) adjusted_FDA - covariate adjustment with two covariates -**hormone receptor status** and **dominant site of disease** - selected by a survey from FDA Division of Oncology

(5) adjusted_FDA* - covariate adjustment with two covariates - **hormone receptor status (ER vs Unknown)** and **dominant site of disease** - selected by FDA Division of Oncology

Reviewer's Appendix 4.2.2: Estimated Relative Risk, 95%CI, and Associated P-values along with Parameter Estimates and Estimated Standard Errors under Several Sets of Covariates

0.5 mg letrozole vs Megestrol Acetate					
	RR	95%CI	Est.	SE	P-Value
Unadjusted	0.98	(0.77, 1.26)	-0.0164	0.127	0.8973
[1]Adjusted_11	1.10	(0.85, 1.42)	0.0937	0.131	0.4741
[2]Adjusted_8	1.08	(0.84, 1.40)	0.0801	0.130	0.5376
[3]Adjusted_S	1.01	(0.79, 1.30)	0.0143	0.128	0.9106
[4]Adjusted_FDA	1.01	(0.79, 1.30)	0.0137	0.128	0.9145
[5]Adjusted_FDA*	1.02	(0.79, 1.31)	0.0186	0.128	0.8847

Note:(1) adjusted_11 - covariate adjustment with 11 covariates reported in this NDA submission.

(2) adjusted_8 - covariate adjustment with 8 covariates specified in the protocol dated on Oct. 27, 1992.

(3) adjusted_S - covariate adjustment with covariates selected by a forward stepwise procedure without treatment effect in the procedure - **viscreal, age category, and performance status** were selected.

(4) adjusted_FDA - covariate adjustment with two covariates **-hormone receptor status** and **dominant site of disease** - selected by a survey from FDA Division of Oncology

(5) adjusted_FDA* - covariate adjustment with two covariates - **hormone receptor status (ER vs Unknown)** and **dominant site of disease** - selected by FDA Division of Oncology

Reviewer's Appendix 4.2.3: Estimated Relative Risk, 95%CI, and Associated P-values along with Parameter Estimates and Estimated Standard Errors under Several Sets of Covariates

2.5 mg letrozole vs Megestrol Acetate					
	RR	95%CI	Est.	SE	P-Value
Unadjusted	0.77	(0.60, 1.00)	- 0.258	0.131	0.0488
[1]Adjusted_11	0.84	(0.65, 1.09)	- 0.175	0.134	0.1927
[2]Adjusted_8	0.80	(0.61, 1.03)	- 0.230	0.133	0.0850
[3]Adjusted_S	0.82	(0.64, 1.07)	- 0.195	0.132	0.1397
[4]Adjusted_FDA	0.79	(0.61, 1.03)	- 0.232	0.131	0.0777
[5]Adjusted_FDA*	0.78	(0.61, 1.02)	- 0.237	0.131	0.0706

Note:(1) adjusted_11 - covariate adjustment with 11 covariates reported in this NDA submission.

(2) adjusted_8 - covariate adjustment with 8 covariates specified in the protocol dated on Oct. 27, 1992.

(3) adjusted_S - covariate adjustment with covariates selected by a forward stepwise procedure without treatment effect in the procedure - **number of sites involved, performance status, age category, disease free interval** were selected.

(4) adjusted_FDA - covariate adjustment with two covariates -**hormone receptor status** and **dominant site of disease** - selected by a survey from FDA Division of Oncology

(5) adjusted_FDA* - covariate adjustment with two covariates - **hormone receptor status (ER vs Unknown)** and **dominant site of disease** - selected by FDA Division of Oncology

Reviewer's Appendix 4.2.4: Total Sample Size, Event Occured, and Censored Sample Size among Dominant Site of Disease within Each Hormone Receptor Status for the Time to Progression

		0.5 mg vs 2.5 mg			0.5 mg vs MA			2.5 mg vs MA		
HRS	DSD	Total	Event	censor	Total	Event	Censor	Total	Event	Censor
RS=1	soft	30	18	12	38	21	17	28	9	19
	vis*	56	42	14	56	43	13	62	47	15
	bone	40	21	19	45	27	18	37	24	13
RS=2	soft	19	11	8	20	12	8	21	11	10
	vis*	35	26	9	29	24	5	34	28	6
	bone	24	15	9	27	17	10	29	20	9
RS=3	soft	59	29	30	56	32	24	53	29	24
	vis*	53	35	18	62	49	13	55	38	17
	bone	46	30	16	44	27	17	44	35	9

Note: - HRS stands for hormone receptor status (RS=1 stands for ER/PR+, RS=2 stands for ER or PR+, and RS=3 stands for Unknown.)
- DSD stands for dominant site of disease.

Reviewer's Appendix 4.2.5: Estimated Relative Risk among Dominant Site of Disease within Each Dominant Site of Disease for the Time to Progression

HRS	DSD	0.5 mg vs 2.5 mg	0.5 mg vs MA	2.5 mg vs MA
RS=1	soft	4.190	2.715	0.540
	visceral	1.408	0.968	0.602
	bone	1.663	0.987	0.653
RS=2	soft	3.128	5.050	1.171
	visceral	0.955	1.158	1.181
	bone	1.853	1.995	1.013
RS=3	soft	1.347	0.667	0.501
	visceral	1.561	1.077	0.667
	bone	0.347	0.451	1.203

Note: - HRS stands for hormone receptor status so that RS=1 stands for ER/PR+, RS=2 stands for ER or PR+, and RS=3 stands for Unknown.
- DSD stands for dominant site of disease.

Reviewer's APPENDIX 4.3.1 A Summary of Longitudinal Linear Models

We briefly outline longitudinal linear models, which can be applied under an ignorable missing assumption or within a homogeneity group under a nonignorable missing assumption.

In a general longitudinal analysis Zeger *et al* (1988) make a distinction between two types of longitudinal analyses: a "subject-specific (SS) model" (a type of mixed effects model) and a "population-averaged (PA) model". In the SS model, we are mainly concerned with individuals' response over time, and the heterogeneity of the data from each individual can be explicitly modeled. On the other hand, the PA model focuses on the average response and the heterogeneity of individuals is not considered in the model.

The SS model focuses on the between-subject variability in a data set. The variance can be modeled explicitly, and will contribute to the marginal covariance structure and/or the marginal mean functions in the SS models. This is the approach used in the linear mixed effects model. However, if the analysis is not focused on accounting for between subject variability, a PA model approach, with relaxed assumptions, can be applied. This is a Generalized Estimating Equations (GEE) approach. As noted by Zeger *et al* (1988), a marginal covariance structure, which is one of challenges in a repeated measurement setting, can be explained by the two approaches in a different fashion. On the other hand, a marginal means, in our case, intercept and a slope, will not be affected by the two approaches.

1. Subject-Specific Linear Models

The linear mixed effects models have been investigated by a number of researchers (Harville, 1976 and 1977, and Rao, 1965, 1967, and 1975). As described above, by introducing distributional assumptions for each individual's random variability, a marginal covariance structures can be explained explicitly. Of particular interest in the regulatory context, Laird and Ware (1982) have described the application of these models to unbalanced (in general we have a balanced design in a clinical trial setting) and incomplete data based on the assumption that a missing mechanism is defined as "missing at random" (MAR), belonging to an ignorable missing mechanism. The model can be defined as

$$\mathbf{y}_i = \mathbf{X}_i\beta + \mathbf{Z}_i\mathbf{b}_i + \varepsilon_i$$

where Z_i is a known design matrix of random effects, b_i, and b_i and ε_i are N (0, Ω) and N (0, $\sigma^2 I_i$) respectively. Note that we assume that b_i and ε_i are independent of each other. To estimate the fixed effects parameters (population parameters), we need to know the marginal means and marginal covariance matrix. Applying the independence assumption of b_i and ε_i with the corresponding expectation equal to 0, we will obtain

$$E(\mathbf{y}_i) = \mathbf{X}_i\beta \text{ and } \operatorname{cov}(\mathbf{y}_i) = \mathbf{Z}_i\Omega\,\mathbf{Z}_i^T + \sigma^2\mathbf{I}_i = \mathbf{V}_i$$

Then the estimated fixed effects parameters can be obtained by

$$\hat{\beta} = \left(\sum_{i=1}^{K_j} X_i^T \hat{V}_i^{-1} X_i \right)^{-1} \left(\sum_{i=1}^{K_j} X_i^T \hat{V}_i^{-1} y_i \right) \text{ and } \operatorname{cov}(\hat{\beta}) = \left(\sum_{i=1}^{K_j} X_i^T \hat{V}_i^{-1} X_i \right)^{-1}$$

Note that (i) the random effects only contribute to the marginal covariance matrix, and not to the marginal means, I.e., V_i is the only function of random efects, and that the covariance structure will depend on a choice of random effects, Z_i, and that (ii) the misspecification of the marginal covariance matrix due to a incorrect choice of the random effects, Z_i, may lead to an underestimate of the variance of the estimated parameters.

The second approach is called a "random coefficient models". This approach is similar to a linear mixed effects model. The model can be defined as

$$y_i = X_i \beta_i + \varepsilon_i$$

where β_i and ε_i are $N(\beta, \Sigma_{\beta\beta})$ and $N(0, \sigma^2 I_i)$ respectively, and we assume that β_i and ε_i are independent each other.

Then a simple unweighted estimator can be defined as

$$b_u = \frac{1}{K_j} \left(\sum_{i=1}^{K_j} \hat{\beta}_i \right), \quad \text{where } \hat{\beta}_i = (X_i^T X_i)^{-1} (X_i^T y_i) \text{ and } \operatorname{cov}(\hat{\beta}_i) = \Sigma_{\beta\beta} + \sigma^2 (X_i^T X_i)^{-1} = W_i$$

And a weighted estimator can be defined as

$$b_w = \frac{1}{\sum_{i=1}^{K_j} W_i^{-1}} \left(\sum_{i=1}^{K_j} W_i^{-1} \hat{\beta}_i \right)$$

Note that for a balanced and a complete design we have $b_u = b_w$.

The main difference between the two approaches is that (I) a weighted least squares (a generalized least squares) is applied to each subject in a linear mixed effects model, and a simple least squares is applied to each subject in a random coefficient model, and (ii) the weighting scheme is different.

2. Population-Averaged Linear Models

In the PA approach to linear models we are interested in a model which is only a function of covariates without introducing subject to subject heterogeneity in the marginal covariance matrix. Therefore the model can be simply defined as

$$y_i = X_i \beta + \varepsilon_i$$

In the SS model, random effects variables are employed to describe the covariance structure.

This unknown correlation structure depends on the selection of Z_i matrix. Thus the selected covariance structure can be viewed as one of a number of possible alternatives. In applying the PA approach, Jennrich and Schluchter (1986) investigated a number of covariance structure (independent observations, compound symmetry, random-effects, first-order autoregressive structure, and so on), in a variety of situations (unbalanced and incomplete designs). They used a likelihood-based approach to the linear model. Therefore the only restriction required for the covariance matrix is a positive definite matrix. Note that the misspecification of of the covariance matrix may lead to an underestimate of the variance of the estimated parameters.

Another approach to the linear model, not requiring distributional assumptions on the error term, is the application of an estimating equation. Invoking M-estimation theory (Huber 1967, White 1982, Liang and Zeger, 1986), the estimating equation can be defined as

$$U(\beta) = \sum_{i=1}^{K} \mathbf{X}_i^T \mathbf{V}_i^{-1} (\mathbf{y}_i - \mathbf{X}_i\beta) = 0$$

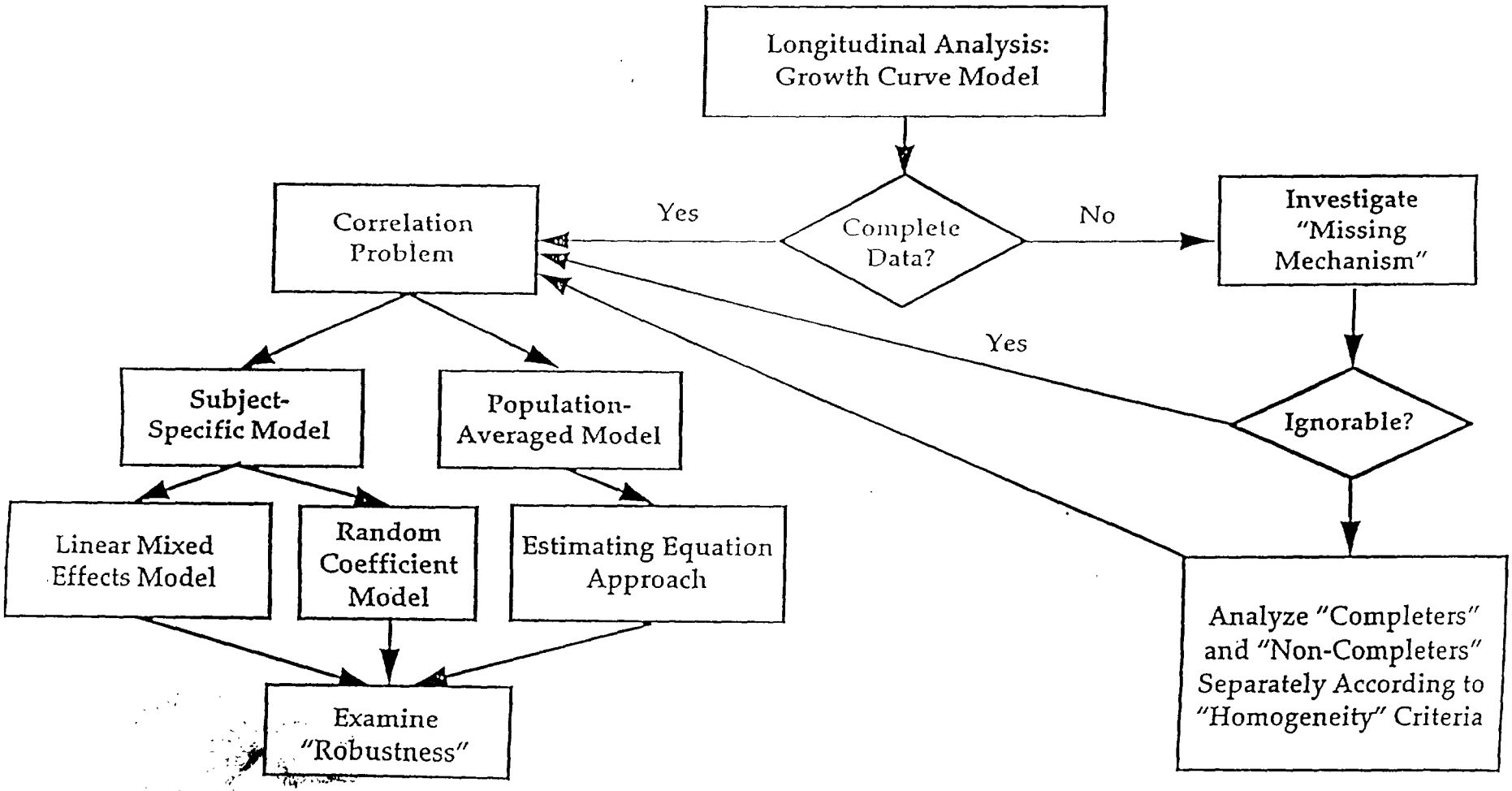
where $\mathbf{V}_i$ is known as a "working" covariance matrix. Note that the solution of the equation is consistent even if $\mathbf{V}_i$ is misspecified as long as the expected value of the estimating equation equal to 0. Liang and Zeger (1986) introduced the notion of a "working" correlation in the estimating equation --a parsimonious covariance structure. In addition, we can protect the underestimation of the variance of the estimators of the population parameters by introducing "sandwich" estimators of the variance, derived from M-estimation theory (Serfling, 1980). This is an important fact in a regulatory context in a sense that the variance estimator will be robust. The sandwich variance estimate of the parameters of interest can be given as

$$\hat{V}_\beta = (\sum_{i=1}^{K} \mathbf{X}_i^T \hat{\mathbf{V}}_i^{-1} \mathbf{X}_i)^{-1} (\sum_{i=1}^{K} \mathbf{X}_i^T \hat{\mathbf{V}}_i^{-1} (\mathbf{y}_i - \mathbf{X}_i\hat{\beta})(\mathbf{y}_i - \mathbf{X}_i\hat{\beta})^T \hat{\mathbf{V}}_i^{-1} \mathbf{X}_i)(\sum_{i=1}^{K} \mathbf{X}_i^T \hat{\mathbf{V}}_i^{-1} \mathbf{X}_i)^{-1}$$

Note that the asymptotic results will depend on having a large number of subjects, not on having a large number of data points per subject.

Reviewer's Figure 4.3.1

Overview: A Longitudinal Approach



Reviewer's APPENDIX 4.3.2:

SUMMARY RESULTS FROM A LONGITUDINAL ANALYSIS OF PAIN SCORE IN THE QLQ-C30 IN AR/B2 STUDY

0.5 mg Letrozole					
Completers: Maxi = 6 or 5					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.685)	inter	78.574	2.147	2.116	37.13
	slope*	-0.00537	0.007	0.008	-0.70
Patients with CR or PR within Completers					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.601)	inter	82.363	3.658	3.645	22.60
	slope*	-0.0080	0.012	0.013	-0.59
Patients with SD, PD, or Unknown within Completers					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.701)	inter	77.333	2.568	2.524	30.64
	slope*	-0.0044	0.008	0.009	-0.47
Dropouts: 2 ≤ Maxi ≤ 4					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.714)	inter	69.919	2.952	2.804	24.93
	slope	-0.105	0.031	0.036	-2.88
Comments: (i) a similar time trend was observed between patients with CR or PR and with SD, PD, or Unknown among completers. (ii) the estimated slope was statistically different between completers and dropouts, indicating that we have a nonignorable missing mechanism. (iii) pain in dropouts had tendency to become better than in completers. (iv) no statistical significance was found in slope*.					
Possible Missing Mechanism					
Nonignorable					

Note that wcor stands for a "working" correlation.
Note that N stands for a naive standard error calculated by the inverse of Fisher's information and R for a robust standard error calculated by a "sandwich" estimator.
Note that Z was calculated by the robust standard error.

Reviewer's APPENDIX 4.3.3:

SUMMARY RESULTS FROM A LONGITUDINAL ANALYSIS OF PAIN SCORE IN THE QLQ-C30 IN AR/B2 STUDY

2.5 mg Letrozole					
Completers: Maxi = 6 or 5					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.724)	inter	77.704	2.092	2.206	35.23
	slope*	0.000869	0.006	0.007	0.12
Patients with CR or PR within Completers					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.772)	inter	78.871	3.363	3.609	21.85
	slope*	0.00122	0.008	0.010	0.12
Patients with SD, PD, or Unknown within Completers					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.675)	inter	76.793	2.626	2.706	28.38
	slope*	0.000221	0.009	0.011	0.02
Dropouts: 2 ≤ Maxi ≤ 4					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.760)	inter	63.718	3.508	3.257	19.56
	slope	-0.176	0.031	0.040	-4.44
Comments: (i) a similar time trend was observed between patients with CR or PR and with SD, PD, or Unknown among completers. (ii) the estimated slope was statistically different between completers and dropouts, indicating that we have a nonignorable missing mechanism. (iii) pain in dropouts had tendency to become better than in completers. (iv) no statistical significance was found in slope*.					
Possible Missing Mechanism					
Nonignorable					

Note that wcor stands for a "working" correlation.
Note that N stands for a naive standard error calculated by the inverse of Fisher's information and R for robust standard error calculated by a "sandwich" estimator.
Note that Z was calculated by the robust standard error.

Reviewer's APPENDIX 4.3.4:
SUMMARY RESULTS FROM A LONGITUDINAL ANALYSIS OF PAIN SCORE IN THE QLQ-C30 IN AR/B2 STUDY

Megestrol Acetate					
Completers: Maxi = 6 or 5					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.695)	inter	73.521	2.399	2.401	30.62
	slope*	-0.0142	0.008	0.009	-1.50
Patients with CR or PR within Completers					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.726)	inter	81.854	4.907	4.579	17.88
	slope	-0.138	0.046	0.064	-2.14
	quad	0.000482	0.000	0.000	2.14
Patients with SD, PD, or Unknown within Completers					
WCOR*	ß	EST	N	R	Z*
AR-1	inter	72.394	2.754	2.834	25.55
	slope	-0.0319	0.015	0.014	-2.29
Dropouts: 2 ≤ Maxi ≤ 4					
WCOR*	ß	EST	N	R	Z*
AR-1 (1,0.71,0.50,0.35)	inter	67.838	3.322	2.963	22.90
	slope	-0.160	0.048	0.043	-3.71
Comments: (i) the estimated slope was statistically significant in maxi=5, which indicates, strictly speaking, that we have a nonignorable missing mechanism. For consistency the two categories were combined. (ii) a different time trend was observed between patients with CR, or PR and with SD, PD, or Unknown among completers. (iii) the estimated slope was statistically different between completers and dropouts, indicating that we have a nonignorable missing mechanism. (iv) pain in dropouts had tendency to become better than in completers. (v) no statistical significance was found in slope*.					
Possible Missing Mechanism					
Nonignorable					

Note that wcoɾ stands for a "working" correlation.
Note that N stands for a naive standard error calculated by the inverse of Fisher's information and R for robust standard error calculated by a "sandwich" estimator.
Note that Z was calculated by the robust standard error.

Reviewer's APPENDIX 4.3.5:

SUMMARY RESULTS FROM A LONGITUDINAL ANALYSIS OF GLOBAL QUALITY OF LIFE IN QLQ-C30 IN AR/B2 STUDY

0.5 mg Letrozole					
Completers: Maxi = 6 or 5					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.534)	inter	71.262	1.862	1.986	35.88
	slope*	-0.00972	0.008	0.009	-1.06
Patients with CR or PR within Completers					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.626)	inter	77.222	3.677	3.745	20.62
	slope*	-0.0101	0.012	0.013	-0.79
Patients with SD, PD, or Unknown within Completers					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.491)	inter	69.358	2.106	2.299	30.17
	slope*	-0.0101	0.009	0.012	-0.85
Dropouts: 2 ≤ Maxi ≤ 4					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.613)	inter	63.120	2.422	2.249	28.07
	slope	-0.117	0.030	0.029	-4.03
Comments: (i) a similar time trend was observed between patients with CR or PR and with SD, PD, or Unknown among completers. (ii) the estimated slope was statistically different between completers and dropouts, indicating that we have a nonignorable missing mechanism. (iii) quality of life in dropouts had a tendency to become worse than in completers. (iv) no statistical significance was found in slope*.					
Possible Missing Mechanism					
Nonignorable					

Note that wcor stands for a "working" correlation.
Note that N stands for a naive standard error calculated by the inverse of Fisher's information and R for robust standard error calculated by a "sandwich" estimator.
Note that Z was calculated by the robust standard error.

Reviewer's APPENDIX 4.3.6:
SUMMARY RESULTS FROM A LONGITUDINAL ANALYSIS OF
GLOBAL QUALITY OF LIFE IN THE QLQ-C30 IN AR/B2 STUDY

2.5 mg Letrozole					
Completers: Maxi = 6 or 5					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.584)	inter	70.590	1.685	1.678	42.07
	slope*	-0.00250	0.006	0.007	-0.34
Patients with CR or PR within Completers					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.679)	inter	70.822	2.488	2.405	29.45
	slope	0.0641	0.025	0.022	2.87
	quad	-0.000212	0.000	0.000	-2.47
Patients with SD, PD, or Unknown within Completers					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.521)	inter	69.071	2.323	2.371	29.14
	slope*	-0.0129	0.010	0.012	-1.05
Dropouts: $2 \leq$ Maxi ≤ 4					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.669)	inter	59.577	2.677	2.562	23.25
	slope	-0.106	0.029	0.037	-2.91
Comments: (i) for maxi=6 linear and quadratic terms were statistically significant, which indicates, strictly speaking, that we have a nonignorable missing mechanism between maxi=6 and maxi=5. For consistency the two categories were combined. (ii) a different time trend was observed between patients with CR or PR and with SD, PD, or Unknown among completers. Linear and quadratic terms were statistically significant in patients with CR or PR. (iii) the estimated slope was statistically different between completers and dropouts, indicating that we have a nonignorable missing mechanism. (iv) quality of life in dropouts had a tendency to become worse than in completers. (v) no statistical significance was found in slope*.					
Possible Missing Mechanism					
Nonignorable					

Note that wcor stands for a "working" correlation.
Note that N stands for a naive standard error calculated by the inverse of Fisher's information and R for robust standard error calculated by a "sandwich" estimator.
Note that Z was calculated by the robust standard error.

Reviewer's APPENDIX 4.3.7:

SUMMARY RESULTS FROM A LONGITUDINAL ANALYSIS OF GLOBAL QUALITY OF LIFE IN THE QLQ-C30 IN AR/B2 STUDY

Megestrol Acetate					
Completers: Maxi = 6 or 5					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.538)	inter	68.221	1.723	1.682	40.56
	slope*	-0.00636	0.007	0.008	-0.77
Patients with CR or PR within Completers					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.678)	inter	70.000	3.442	3.228	21.68
	slope*	0.00241	0.010	0.011	0.22
Patients with SD, PD, or Unknown within Completers					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.463)	inter	67.528	1.948	1.973	34.22
	slope*	-0.0117	0.009	0.011	-1.03
Dropouts: 2 ≤ Maxi ≤ 4					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.586)	inter	60.657	2.582	2.443	24.83
	slope	-0.123	0.032	0.037	-3.30
Comments: (i) the estimated time trend was different between maxi=6 and maxi=5, which indicates, strictly speaking, that we have a nonignorable missing mechanism. For consistency the two categories were combined. (ii) a similar time trend was observed between patients with CR or PR and with SD, PD, or Unknown among completers. (iii) the estimated slope was statistically different between completers and dropouts, indicating that we have a nonignorable missing mechanism. (iv) quality of life in dropouts had tendency to become worse than in completers.					
Possible Missing Mechanism					
Nonignorable					

Note that wcor stands for a "working" correlation.
Note that N stands for a naive standard error calculated by the inverse of Fisher's information and R for robust standard error calculated by a "sandwich" estimator.
Note that Z was calculated by the robust standard error.

Reviewer's APPENDIX 4.3.8:

SUMMARY RESULTS FROM A LONGITUDINAL ANALYSIS OF PERFORMANCE STATUS IN AR/B2 STUDY

0.5 mg Letrozole					
Completers: Maxi = 6 or 5					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.586)	inter	0.491	0.061	0.063	7.83
	slope*	0.000575	0.000	0.000	1.88
Patients with CR or PR within Completers					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.583)	inter	0.355	0.104	0.118	3.01
	slope*	-0.000107	0.000	0.000	-0.28
Patients with SD, PD, or Unknown within Completers					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.579)	inter	0.524	0.071	0.072	7.24
	slope	0.000831	0.0000	0.000	2.14
Dropouts: 2 ≤ Maxi ≤ 4					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.517)	inter	0.803	0.097	0.082	9.75
	slope	0.0729	0.001	0.001	5.33
Comments: (i) for maxi=5 the estimated slope was statistically significant, which indicates strictly speaking that we have a nonignorable missing mechanism. For consistency two categories (maxi=6 and maxi=5) were combined. (ii) a similar time trend was observed between patients with CR or PR and with SD, PD, or Unknown among completers. (iii) the estimated slope was statistically different between completers and dropouts, indicating that we have a nonignorable missing mechanism. (iv) performance status in dropouts had tendency to become worse than in completers.					
Possible Missing Mechanism					
Nonignorable					

Note that wcor stands for a "working" correlation.
Note that N stands for a naive standard error calculated by the inverse of Fisher's information and R for robust standard error calculated by a "sandwich" estimator.
Note that Z was calculated by the robust standard error.

Reviewer's APPENDIX 4.3.9:

SUMMARY RESULTS FROM A LONGITUDINAL ANALYSIS OF PERFORMANCE STATUS IN AR/B2 STUDY

2.5 mg Letrozole					
Completers: Maxi = 6 or 5					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.662)	inter	0.517	0.058	0.059	8.75
	slope*	0.0000413	0.000	0.000	0.17
Patients with CR or PR within Completers					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.760)	inter	0.458	0.090	0.094	4.85
	slope*	-0.000205	0.000	0.000	-0.69
Patients with SD, PD, or Unknown within Completers					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.594)	inter	0.554	0.074	0.076	7.33
	slope*	0.000290	0.000	0.000	0.73
Dropouts: 2 ≤ Maxi ≤ 4					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.623)	inter	0.861	0.108	0.095	9.05
	slope	0.00581	0.001	0.001	4.93
Comments: (i) a similar time trend was observed between patients with CR or PR and with SD, PD, or Unknown among completers. (ii) the estimated slope was statistically different between completers and dropouts, indicating that we have a nonignorable missing mechanism. (iii) performance status in dropouts had tendency to become worse than in completers. (v) no statistical significance was found in slope*.					
Possible Missing Mechanism					
Nonignorable					

Note that wcor stands for a "working" correlation.
Note that N stands for a naive standard error calculated by the inverse of Fisher's information and R for robust standard error calculated by a "sandwich" estimator.
Note that Z was calculated by the robust standard error.

Reviewer's APPENDIX 4.3.10:
SUMMARY RESULTS FROM A LONGITUDINAL ANALYSIS OF PERFORMANCE STATUS IN AR/B2 STUDY

Megestrol Acetate					
Completers: Maxi = 6 or 5					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.561)	inter	0.488	0.062	0.054	9.09
	slope	0.00148	0.000	0.000	4.39
Patients with CR or PR within Completers					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.577)	inter	0.281	0.089	0.071	3.95
	slope	0.000941	0.000	0.000	2.56
Patients with SD, PD, or Unknown within Completers					
WCOR*	ß	EST	N	R	Z*
Exchangeable (0.534)	inter	0.565	0.076	0.068	8.33
	slope	0.00180	0.000	0.000	3.74
Dropouts: 2 ≤ Maxi ≤ 4					
WCOR*	ß	EST	N	R	Z*
Excahngeable (0.576)	inter	0.782	0.101	0.085	9.23
	slope	0.00829	0.001	0.001	6.11
Comments: (i) for maxi=5, the estimated slope is larger than for maxi=6. Strictly speaking nonignoable missing mechanism was observed, indicating that we can not combine these two categories. But for consistency, the two categories were combined. (ii) a different time trend was observed between patients with CR or PR and with SD, PD, or Unknown among completers, indicating that performance status score tended to be worse in patients with SD, PD, or Unknown than for those with CR or PR. (iii) the estimated slope was statistically different between completers and dropouts, indicating that we have a nonignorable missing mechanism. (Iv) performance status in dropouts had tendency to become worse than in completers.					
Possible Missing Mechanism					
Nonignorable					

Note that wcor stands for a "working" correlation.
Note that N stands for a naive standard error calculated by the inverse of Fisher's information and R for robust standard error calculated by a "sandwich" estimator.
Note that Z was calculated by the robust standard error.

Reviewer's APPENDIX 6.1:

Estimated Odds Ratios, 95% CIs and Associated P-Values in Each Category Stratified by a Hormone Receptor Status in Each Treatment Comparison

0.5 mg letrozole vs. 2.5 mg letrozole			
	RS = 1	RS = 2	RS = 3
Est. OR	0.585	1.159	1.305
95% CI	(0.226 , 1.459)	(0.242 , 5.552)	(0.523 , 3.366)
P - Value	p = 0.292	p = 1.000	P = 0.682
Homogeneity: Breslow - Day: p = 0.380 Zelen: p = 0.383 Note: Hormone receptor status can be considered as an 'effect modification' factor. Note: Mantel-Haenszel OR = 0.914, 95% CI = (0.535, 1.562), p = 0.743			
0.5 mg letrozole vs. Aminoglutethimide			
	RS = 1	RS = 2	RS = 3
Est. OR	1.359	1.087	2.027
95% CI	(0.443 , 4.455)	(0.210 , 6.066)	(0.785 , 5.558)
P - Value	p = 0.738	p = 1.000	p = 0.164
Homogeneity: Breslow - Day: p = 0.715 Zelen: p = 0.772 Note: Mantel-Haenszel OR = 1.582, 95% CI = (0.869 , 2.882), p = 0.133			
2.5 mg letrozole vs Aminoglutethimide			
	RS = 1	RS = 2	RS = 3
Est. OR	2.324	0.938	1.553
95% CI	(0.832 , 7.198)	(0.182 , 5.200)	(0.546 , 4.532)
P - Value	p = 0.121	p = 1.000	p = 0.491
Homogeneity: Breslow - Day: p = 0.560 Zelen: p = 0.553 Note: Hormone receptor status can be considered as an 'effect modification' factor. Note: Mantel-Haenszel OR = 1.681, 95% CI = (0.921 , 3.065), p = 0.09			

Note: RS =1 stands for ER/PR+, RS = 2 stands for ER or PR+, and RS = 3 stands for Unknown hormone receptor status.

Reviewer's APPENDIX 6.2:

Estimated Risk Ratios and P-Values by Logrank Tests in Each Category Stratified by Hormone Receptor Status and Associated P-Value by a Stratified Logrank Test in Each Treatment Comparison

0.5 mg letrozole vs 2.5 mg letrozole			
	RS =1	**RS = 2**	**RS = 3**
RR	1.360	1.333	0.856
P-Value (logrank)	P = 0.122	P = 0.296	P = 0.423
Stratified Logrank Test: P = 0.354			
0.5 mg letrozole vs Aminoglutethimide			
	RS =1	RS = 2	RS =3
RR	1.135	0.989	0.659
P-Value (logrank)	P = 0.532	P = 0.969	P = 0.021
Stratified Logrank Test: P = 0.629			
2.5 mg letrozole vs Aminoglutethimide			
	RS = 1	RS = 2	RS = 3
RR	0.795	0.714	0.777
P-Value (logrank)	P = 0.280	P = 0.234	P = 0.175
Stratified Logrank Test: P = 0.039			

Note: the estimated RRs are derived from Cox regression with only treatment effect in a model.

Note: RS = 1 stands for ER/PR+, RS = 2 stands for ER or PR+, and RS =3 stands for Unknown hormone receptor status.

Reviewer's APPENDIX 6.1:

Estimated Odds Ratios, 95% CIs and Associated P-Values in Each Category Stratified by Hormone Receptor Status for Each Treatment Comparison

0.5 mg letrozole vs. 2.5 mg letrozole			
	RS = 1	RS = 2	RS = 3
Est. OR	0.585	1.159	1.305
95% CI	(0.226 , 1.459)	(0.242 , 5.552)	(0.523 , 3.366)
P - Value	p = 0.292	p = 1.000	P = 0.682
Homogeneity: Breslow - Day: p = 0.380 Zelen: p = 0.383 Note: Hormone receptor status can be considered as an 'effect modification' factor. Note: Mantel-Haenszel OR = 0.914, 95% CI = (0.535, 1.562), p = 0.743			
0.5 mg letrozole vs. Aminoglutethimide			
	RS = 1	RS = 2	RS = 3
Est. OR	1.359	1.087	2.027
95% CI	(0.443 , 4.455)	(0.210 , 6.066)	(0.785 , 5.558)
P - Value	p = 0.738	p = 1.000	p = 0.164
Homogeneity: Breslow - Day: p = 0.715 Zelen: p = 0.772 Note: Mantel-Haenszel OR = 1.582, 95% CI = (0.869 , 2.882), p = 0.133			
2.5 mg letrozole vs Aminoglutethimide			
	RS = 1	RS = 2	RS = 3
Est. OR	2.324	0.938	1.553
95% CI	(0.832 , 7.198)	(0.182 , 5.200)	(0.546 , 4.532)
P - Value	p = 0.121	p = 1.000	p = 0.491
Homogeneity: Breslow - Day: p = 0.560 Zelen: p = 0.553 Note: Hormone receptor status can be considered as an 'effect modification' factor. Note: Mantel-Haenszel OR = 1.681, 95% CI = (0.921 , 3.065), p = 0.09			

Note: RS =1 stands for ER/PR+, RS = 2 stands for ER or PR+, and RS = 3 stands for Unknown hormone receptor status.

Reviewer's APPENDIX 6.2:

Estimated Risk Ratios and P-Values by Logrank Tests in Each Category Stratified by Hormone Receptor Status and Associated P-Value by a Stratified Logrank Test for Each Treatment Comparison

0.5 mg letrozole vs 2.5 mg letrozole			
	RS =1	**RS = 2**	**RS = 3**
RR	1.360	1.333	0.856
P-Value (logrank)	P = 0.122	P = 0.296	P = 0.423
Stratified Logrank Test: P = 0.354			
0.5 mg letrozole vs Aminoglutethimide			
	RS =1	RS = 2	RS =3
RR	1.135	0.989	0.659
P-Value (logrank)	P = 0.532	P = 0.969	P = 0.021
Stratified Logrank Test: P = 0.629			
2.5 mg letrozole vs Aminoglutethimide			
	RS = 1	RS = 2	RS = 3
RR	0.795	0.714	0.777
P-Value (logrank)	P = 0.280	P = 0.234	P = 0.175
Stratified Logrank Test: P = 0.039			

Note: the estimated RRs are derived from Cox regression with only treatment effect in a model.
Note: RS = 1 stands for ER/PR+, RS = 2 stands for ER or PR+, and RS =3 stands for Unknown hormone receptor status.

SE- 30 ...

STATISTICAL SUMMARY
45-DAY REVIEW OF NDA 20-726

I. General Information:

NDA: 20-726
Sponsor: Ciba-Geigy Corporation
Drug Name: Femara™ (letrozole)
Proposed Indication: Advanced breast cancer in postmenopausal women
Dosage Form: Tablet
Review statisticians: Masahiro Takeuchi and Roswitha Kelly

II. Summary of the Controlled Clinical Trial

There is one controlled clinical trial, AR/BC2, in the submission to support effect of letrozole for "treatment of advanced breast cancer in postmenopausal women who have previously received antiestrogen therapy." The following table summarizes the trial.

Title of the Study	AR/BC2
Study Design	double-blind, randomized, multicenter, comparative between daily oral doses of 0.5 mg and 2.5 mg of letrozole versus megestrol acetate 160 mg once daily
Diagnosis	"postmenopausal patients with locally advanced or loco-regionally recurrent or metastatic breast cancer (measurable and/or evaluable) who previously progressed under adjuvant or therapeutic anti-estrogen"
Objectives	Primary efficacy variables: response rate (CR+PR), duration of response, time to progression (TTP), time to treatment failure (TTF), time to death (TTD) Secondary variables: performance status, severity of pain, quality of life, ethical code-breaks, endocrine response, and trough levels

Number of Subjects	Total enrolled/randomized: 552 0.5 mg letrozole: 188 2.5 mg letrozole: 174 megestrol acetate 160 mg: 190
Statistical Methods	-logistic regression analysis for response rate -Cox proportional regression analysis for TTP, TTF, and TTD Note: both adjusted and unadjusted analyses were performed for prognostic factors
Results	ITT 0.5mg 2.5 mg MA SS 188 174 189 CR+PR 11.7% (12.8%)* 23.6% 16.4% Med. dur. 474 (555)* + 561 (546)* Med. TTP 154 170 168 Med. TTF 98 168 (155)* 118 Med.TTD 604 (633)* 650 (731)* + (660)*
Conclusion (by sponsor)	"2.5 mg letrozole can be recommended as a legitimate alternative to megestrol acetate for the treatment of advanced breast cancer in postmenopausal women"

Note: the results by ()* were obtained from AR/BC2 Extension (six months additional follow-up of AR/BC2).

Two diskettes with SAS data sets and statistical efficacy analysis programs for AR/BC2 and AR/BC2 Extension and paper copy of file documentation were provided by the sponsor.

III. Statistical Issues:

a. Covariate Adjustment for a logistic regression analysis and a Cox regression analysis.
b. No longitudinal analyses (repeated measurements per subject and a dropout problem) were performed for secondary variables.

IV. Comments:

There are no major deficiencies for statistical review. The application is fileable from a statistical standpoint.

9/6/96

Masahiro Takeuchi Sc.D.
Mathematical Statistician

4/6/96

Roswitha Kelly M.S.
Mathematical Statistician

9/30/96

Clare Gnecco Ph.D.
Team Leader

Statistical Review and Evaluation

APR 18 '97

DATE:

NDA#: 20-726

APPLICANT: Novartis Pharmaceuticals Corp.

NAME OF DRUG: Femara (letrozole) 2.4 mg Tablets

DOCUMENTS REVIEWED: Copies of Original Amendments B2 and BZ Dated 12/30/96 and ~~03/03/97~~ 02/27/97, Respectively.

I. Background (Lispillman) 4/22/97

Dr. Paul Dietze (HFD-150) requested the Division of Biometrics to review the sponsor's stability submission in support of a 24 months expiration dating period. The 12/30/96 submission contained nine months of primary stability data, the ~~03/03/97~~ 02/27/97 submission contained 12 months. This review will address the updated information.

II. Sponsor's Results

There were eight primary stability `studies'. Each `study' consisted of one of the three batches being filled into either 30 count or 100 count bottles. In addition, one of the batches was also filled into 30 count and 100 count bottles which had closures from a different manufacturer.

The sponsor combined the data of all eight studies testing for significant differences of slopes and intercepts at p=.25. As none of the effects were significant, a single regression line was fitted. The one-sided lower confidence band did not intersect with the specification limit of 95 % label claim within the plotted 24 months. The sponsor therefore concluded that the requested 24 months expiration dating period is supported.

III. Reviewer's Results

The sponsor's statistical approach is generally valid. However, there are two points on which this reviewer disagrees with the sponsor: 1) that the potential effects of the different bottle sizes and closures were ignored in the model, and 2) that only one-sided confidence limits were used.

1) There are only three batches on stability. These three batches were each filled into 30 count and 100 count bottles of the same type. In addition, one batch was also filled into 30 count and 100 count bottles with a different closure. If one does not assume a priori that the degradation pattern of the different size bottles

or of the bottles with different closures is identical, then allowing all data into one model would result in narrower confidence bands than appropriate. In addition, the results of each batch are highly correlated across bottle size and closure manufacturers, so that the sponsor's results estimate more how a batch performs in different containers than how stable the manufacturering process is. This reviewer chose to analyze each container set separately with the following result:

Study Number	Bottle Size	Closure Type	Est. Expiration Period
28059 28060 28123	30	Kerr Kerr Kerr	32 months
28059 28060 28123 28110	30	Kerr Kerr Kerr Owens/Illinois	40 months
28054 28055 28122	100	Kerr Kerr Kerr	28 months
28054 28055 28122 28111	100	Kerr Kerr Kerr Owens/Illinois	33 months
28110 28111	30 100	Owens/Illinois Owens/Illinois	24 months

These data show that the extrapolated estimated expiration dating period is at least 24 months, but this extrapolation assumes that the degradation pattern observed for the first 12 months (actually only nine months for study numbers 28122 and 28123) will be maintained in the future. This concern supports the custom of granting an expiry period only six months beyond the actual data.

2) Potency usually has a lower and an upper specification limit and it is not proper to ignore this fact when establishing an expiration dating period. In addition, though the product theoretically degrades in the direction of losing potency, statistical variation can cause the upper confidence bound to intersect with the upper specification limit before the lower

confidence band intersects with the lower specification limit. The *Guideline for Submitting Documentation for the Stability of Human Drugs and Biologics* (FDA, 1987) states that the expiration dating period is set by the earliest intersection of a confidence band with the corresponding specification limit.

Whereas the sponsor used only the lower confidence band this reviewer used both. This procedure as well as the fact that this reviewer based her estimation of expiration dating periods on smaller groupings of the three batches than did the sponsor, resulted in the shorter estimated expiry periods.

IV Summary and Conclusion

Though the statistical approach used by the sponsor was valid, this reviewer took issue with the sponsor's not accounting for the different bottle sizes and closure manufacturers and with using only a one-sided confidence band around the regression line. This reviewer's approach resulted in shorter expiration dating periods for the various groupings of the three batches than did the sponsor's. Her findings do support an extrapolated expiration dating period of 24 months, but this extrapolation relies on the assumption that the product will continue to degrade linearly to the same degree as observed for the first nine to twelve months. Granting an expiration dating period of six months beyond the actual data rather than the requested 24 months seems more appropriate given the above mentioned factors bearing on the adequacy of the submitted data.

Roswitha E. Kelly

Concur:

Clare Gnecco 4/18/97

Clare Gnecco, Ph.D.
Team Leader

4/18/97

George Chi, Ph.D.
Director, DB I

cc: Archival NDA 20-726 Femara (letrozole) 2.5 mg Tablets, Novartis
STABILITY
HFD-150/Dr. Dietze, Dr. Zhou
HFD-150/Dr. Tolgyesi
HFD-150/D. Spillman, CSO
HFD-710/Dr. Chi
HFD-710/Dr. Gnecco
HFD-710/R. Kelly
HFD-710/Chron

This review consists of 4 pages of text.
RKELLY/04/18/97/wp-femastab.rev