

CENTER FOR DRUG EVALUATION AND RESEARCH

APPROVAL PACKAGE FOR:

**APPLICATION NUMBER
NDA 50-778**

Statistical Review(s)

Statistical Review and Evaluation - NDA Review

NDA#: 21-010

DATE:

SPONSOR: Pharmacia & Upjohn Company

DRUG: Epirubicin Hydrochloride

INDICATION: Treatment of locally advanced or metastatic breast cancer

DATE: Date received by Medical Division (Stamp Date): 12/16/98
Date received by Division of Biometrics II: 1/28/99

DOCUMENTS REVIEWED:

Volumes 2.1-2.3 and 2.10-2.69 of NDA submission dated 12/15/98
Correspondence from Pharmacia & Upjohn dated 2/22/99, 3/9/99, 3/15/99, and 3/23/99
Electronic Data submitted 12/15/98 and 4/30/99

STATISTICAL REVIEWER: Ruthanna Davi, M.S.

MEDICAL REVIEWER: Susan Honig, MD

PROJECT MANAGER/CSO: Patrick Guinn

RELEVANT STATISTICAL ISSUES:

For study MA5, the protocol stated that Cox proportional hazard modeling would be used as a secondary analysis. However, it did not specify what covariates would be considered for the model. In addition, methods for development of the model (e.g., forward or backward stepwise regression, convergence criterion, maximum number of iterations, etc.) were not pre-specified. For study HEPI013, the protocol did not state that Cox proportional hazard modeling would be used. The results of both studies were presented using Cox proportional hazard modeling. In addition, the validity of the proportional hazards assumption was not addressed by the sponsor and was particularly questionable in study MA5.

An interim analysis was not planned for study MA5. However, an interim analysis was conducted when recruitment was complete and 263 relapses (86% of the total number of relapses that would occur by the end of the trial) had occurred. Using Lan and Demets alpha spending function procedures, an adjustment in the overall significance level (as determined by this reviewer) was to set $\alpha = 0.049$.

The quality of life measure used in study MA5 was an unweighted average of 30 questions. Justification for assuming equal weight for each question was not addressed by the sponsor. The sponsor did not assess whether or not the missingness mechanism for the quality of life data was informative, leaving the quality of life analyses open to question.

For study HEPI013, the protocol stated that the primary analysis of the time to progression (the primary efficacy endpoint) would be conducted using the intent-to-treat group. However, the study report showed only results for the "fully eligible, evaluable, treated as randomized" group which was smaller than the intent to treat group by about 15% in each treatment arm.

Introduction

The sponsor has submitted the results of six phase III studies in support of the efficacy of Epirubicin for the following proposed indication.

“TM Injection* is indicated as a component of adjuvant therapy in patients with evidence of axillary-node-tumor involvement following resection of primary breast cancer (Stage II & III). TM* is indicated for the therapy of patients with locally advanced or metastatic breast cancer.”

*The sponsor had not yet determined the trade name for Epirubicin at the time the NDA was submitted.

Three of the six studies submitted were conducted in previously untreated women with potentially curable adenocarcinoma of the breast. Of these three studies, one study, “MA-5”, is considered by the sponsor and the agency to be pivotal for substantiating the efficacy of Epirubicin in this patient group. Similarly, the results of three studies evaluating Epirubicin as a first line therapy were submitted and one trial, namely “HEPI013”, is considered pivotal for purposes of efficacy. Therefore, this document contains a **detailed statistical review of trials MA-5 and HEPI013**. For clarity, throughout this document, this reviewer’s comments will be italicized.

II. Study MA-5

II.A. Study Design / Analysis Plan – Study MA-5

Study MA-5 was a multicenter (39 centers), open label, phase III trial in premenopausal women with operable axillary node-positive breast cancer who had undergone complete resection of all known disease by means of total or partial mastectomy. Subjects were stratified by nodal status (1-3, 4-10, and >10 axillary nodes positive for tumor), type of initial surgery (total versus partial mastectomy), and estrogen/progesterone receptor status (either ≥ 10 , both < 10 , or unknown) and were randomized to either CEF or CMF therapy¹. Chemotherapy was repeated every four weeks for a total of six cycles. Follow-up of each subject continued until the end of the study or until death which ever occurred first.

The **primary efficacy endpoint** was designated in the protocol to be **relapse free survival** defined as the time from randomization until recurrence of disease. **Secondary efficacy endpoints** include **overall survival** (defined as the time from randomization until death from any cause) and a **quality of life indicator**.

Comparisons of relapse free survival and overall survival for the CEF group versus the CMF group were conducted using the protocol specified **stratified log rank test and plots of the Kaplan-Meier survival curves**. KM survival curves will be presented for each of the following subgroups.

1. Nodal status (1-3, 4-10, and > 10 axillary nodes positive for tumor),
2. Type of initial surgery (total versus partial mastectomy),
3. Estrogen/progesterone receptor status (either ≥ 10 , both < 10 , or unknown),
4. Menopausal status (pre- and perimenopausal), and
5. Tumor size (T0-T2, T3-T4, Tx or missing).

In addition, to evaluate the effect of covariates, **Cox proportional hazard modeling** was used. The use of Cox proportional hazard modeling was provisioned for in the protocol; however, the covariates for the model were not specified. Development of the model (i.e., determination of which covariates to include in the model) was based on forward stepwise procedures.

-
- 1 CEF Therapy: Cyclophosphamide 75 mg/m² orally, days 1 through 14
Epirubicin 60 mg/m² IV, days 1 and 8
5-FU 500 mg/m² IV, days 1 and 8
- CMF Therapy: Cyclophosphamide 100 mg/m² orally, days 1 through 14
Methotrexate 40 mg/m² IV, days 1 and 8
5-FU 600 mg/m², days 1 and 8

Quality of life was measured using the “**Breast Cancer Chemotherapy Questionnaire**”. This questionnaire was previously developed with the objective of measuring quality of life in women with stage II breast cancer.² The Breast Cancer Chemotherapy Questionnaire contains 30 questions addressing the impact of treatment for breast cancer on the physical, emotional, and social function of the patient. The areas of concern for breast cancer patients that are addressed on the questionnaire include (1.) consequences of hair loss, (2.) emotional dysfunction, (3.) physical symptoms, (4.) trouble and inconvenience associated with treatment, (5.) fatigue, (6.) nausea, and (7.) positive well-being. These items were chosen as a result of a review of the literature, the opinions of patients and clinicians, and formal interviews of women receiving treatment for breast cancer. There are at least four questions addressing each of the seven topics listed above on the Breast Cancer Chemotherapy Questionnaire. Response on an ordered seven-point scale is required for each question. The lowest score represents the worst possible outcome and the highest score represents the best possible outcome. The final score for the Breast Cancer Chemotherapy Questionnaire is the mean of a patient’s responses for all 30 questions. *Note that using the mean response as the final outcome results in each question being treated with equal weight or importance.*

The validity, reproducibility, and responsiveness of the questionnaire were previously evaluated in a group of 418 patients with stage II breast cancer. Two hundred sixteen of those patients were randomly assigned to receive chemotherapy for 12 weeks while 202 patients were receiving chemotherapy for 36 weeks. Since there is no gold standard for quality of life, an assessment of the validity of this questionnaire was made by comparing the results of the new questionnaire to other commonly used instruments for evaluating quality of life (i.e., Karnofsky, Rand physical Rand emotional, Spitzer QL). The correlation between the new questionnaire and the other instruments ranged from 0.41 to 0.62. It was also found that the new questionnaire correlated more closely with patients’ ratings as well as the physicians’ ratings of physical and emotional function than did any of the other instruments that were evaluated. Reproducibility of the results of the questionnaire was assessed using the change in the final score across a two-week period during which the patient reported no change in her global assessment of both physical and emotional function. This resulted in a mean value that was not statistically significantly different from zero indicating that the new questionnaire accurately reflected the absence of a change in the patient’s condition. (i.e., A patient in the same condition at two time points two weeks apart scored very similarly on the new questionnaire at each time point). Finally, the responsiveness of the questionnaire was evaluated by taking advantage of the fact that one group of subjects was receiving treatment for 12 weeks while the other group was receiving treatment for 36 weeks. One would expect similar quality of life in the two patient groups when receiving similar treatment and differing quality of life for the weeks when treatment differed. Statistically significant differences between patient groups for the new questionnaire final score were found for weeks ten through 24 while no significant differences between treatment groups were found for weeks eight to ten. The researchers concluded that, “strong evidence for the usefulness of this new questionnaire [the Breast Cancer Chemotherapy Questionnaire] in the setting of a clinical trial”, had been provided.²

The sponsor reports that initially, the Breast Cancer Chemotherapy Questionnaire (with 30 questions) was used for evaluating quality of life but that at some point during the trial, two additional questions were added to the questionnaire. *No further explanation for why or when the questions were added or if this change in protocol would be accounted for in the analysis were provided.* Each patient was expected to complete the questionnaire at baseline, monthly while on treatment, and at 9, 12, 15, 18, 21, and 24 months. According to the protocol, **repeated measures analysis of variance** will be used to summarize the quality of life data.

The **planned sample size** was based on (1.) a literature reference indicating that five-year relapse free survival with CMF was approximately 55%, (2.) the investigators expectation for five-year relapse free survival with CEF was 65%, and (3.) $\alpha = 0.05$, $\beta = 0.20$, and the use of one-sided tests. The resulting

² Levine MN, Guyatt GH, Gent M, et al. Quality of life in Stage II breast cancer: an instrument for clinical trials. J Clin Oncol 1988; 6: 1798-1810.

necessary sample size was **296 patients per treatment group (592 total)**. However, according to the sponsor since the subject accrual rate was much faster than originally anticipated, the targeted sample size of 592 patients was reached much more quickly than was expected. At that time a decision was made by the Study Steering Committee to continue recruitment for several more months until a new trial of adjuvant therapy in node-positive breast cancer which was under development was open for recruitment. The result was that for this trial, the **actual sample size was 716 patients (356 and 360 randomized to CEF and CMF, respectively)**.

The **protocol did not plan for an interim analysis**. However, after three years of follow-up when recruitment was already complete and 263 events had occurred, an **interim analysis was conducted**. The results of the interim analysis showed a statistically significant improvement in relapse free survival after a median follow up of four years in women who received CEF compared with women who received CMF (66% versus 56% respectively, $p=0.01$). Although a trend for improvement in overall survival with CEF versus CMF was observed, the difference was not statistically significant. *Substantiation of the result for overall survival would require further follow-up and more occurrences of events. Since this interim analysis was conducted after recruitment was complete, presumably the change in sample size discussed above was not impacted by the results of this analysis. Although not explicitly stated by the sponsor, it appears that the conduct of the study was not altered in any way as a result of this interim analysis.*

The only difference in the databases for the interim analysis and the final analysis was that follow up of all patients was complete at the time of the final analysis. At the time of the interim analysis 263 relapses had occurred. At the time of the final analysis 305 relapses had occurred. Since the amount of information known at the time of the interim analysis was nearly complete (i.e., 86% of relapses had already occurred) and using the Lan and DeMets alpha spending function procedure, only a very small correction in the overall significance level is warranted. In the opinion of this reviewer, an adjustment of the final α level to 0.049 (rather than 0.05) as a result of the interim analysis is appropriate.

II.B. Sponsor's Results and Reviewer's Comments for Study MA-5

The **intention to treat group (ITT) includes 716 patients** that were accrued and randomized to receive treatment with **CEF (356 patients) or CMF (360 patients)**. Five of the 716 subjects randomized did not appropriately receive their assigned treatment. Two patients were enrolled (one in CEF, one in CMF) and subsequently determined to be ineligible since more than 10 weeks had passed since their surgery. One patient received CMF but had been randomized to CEF. Finally, two patients who were randomized to CEF were treated with CMF during certain cycles of therapy. The percentage of patients who completed the planned six cycles of treatment was comparable between treatment groups with the majority of patients (96% in the CEF group, 97% in the CMF group) completing all six treatment cycles. Since the ITT group differs from the per protocol group (PP) by only five subjects, efficacy results will be presented for the ITT group only. The results for the PP group would be nearly identical.

The **median follow up time was 54 months for both CEF and CMF**. The log rank test indicated that across time, the probability of remaining in the study was similar between treatment groups ($p=0.81$). No significant differences in demographic and baseline characteristics were found between treatment groups. The variables examined included: age, performance status (ECOG grade), menopausal status (pre vs. peri), number of positive nodes, receptor status, clinical stage, and type of surgery. *The lack of an association between treatment and each of these variables indicates that confounding of the treatment effect due to these variables is not likely.*

The results of the stratified log rank test (stratified by nodal status, total versus partial mastectomy, estrogen/progesterone receptor status) and KM curve depicted in Figure 1 show **statistically significant prolongation of relapse free survival in the CEF group compared to the CMF group** ($p=0.013$). The KM estimates of relapse free survival at five years were 62% and 54% in the CEF and CMF groups, respectively. Similar analysis and the graphical display in Figure 2 show a **statistically significant prolongation of overall survival in the CEF group compared to the CMF group** ($p=0.043$). The KM estimates of overall survival at five years were 77% and 70% for the CEF and CMF groups, respectively.

The **within strata relationships between CEF and CMF are very similar to the overall results** for both relapse free survival and overall survival. Table 1 shows numeric estimates of 5-year relapse free survival and overall survival within each stratum for each treatment. However, the relationship between treatments within strata is perhaps more clearly visualized by the KM survival curves in Figures 1A through 1J for relapse free survival and Figures 2A through 2J for overall survival. These relationships are not statistically significant findings (adequate power within each stratum was not expected and was not achieved) but are graphical displays indicating consistent trends.

Of the ten subgroups created for relapse free survival, there was only **one instance where the relapse free survival for the CEF group did not appear to differ from that of the CMF group** (illustrated by overlapping survival curves in Figure 1F, peri-menopausal patients). **In all other subgroups, relapse free survival appears to be prolonged in the CEF group versus the CMF group** (Figures 1A-E, G-J).

Since the number of deaths during the study was smaller than the number of relapses, the within strata relationships for overall survival graphically were not as clear cut as were the results for relapse free survival. Of the ten subgroups created for overall survival, there was **one instance where CMF appeared to provide slightly better survival than did CEF** (Figure 2F, overall survival for peri-menopausal patients). There is **one case where there is overlapping of the survival curves for CEF and CMF** (Figure 2C, negative receptor status). **In all other subgroups, overall survival appeared to be prolonged in the CEF group versus the CMF group** (Figures 2A-B, 2D-E, 2G-J).

Table 1: Relapse Free Survival and Overall Survival Estimates at Five Years Stratified by Covariates

	CEF			CMF		
	N (%)*	5-year relapse free survival	5-year overall survival	N (%)*	5-year relapse free survival	5-year overall survival
Total Patients	356 (100%)	62%	71%	360 (100%)	53%	70%
Positive Nodes						
1-3	218 (61%)	68%	82%	218 (61%)	62%	78%
4 or more	138 (52%)	52%	69%	142 (39%)	39%	58%
Surgery						
Partial mastectomy	175 (49%)	66%	82%	176 (49%)	56%	76%
Total mastectomy	181 (51%)	58%	73%	184 (51%)	50%	64%
Receptor Status						
Negative	75 (21%)	58%	61%	70 (19%)	55%	59%
Positive	241 (68%)	62%	63%	245 (68%)	51%	70%
Menopausal Status						
Peri-menopausal	80 (22%)	61%	74%	75 (21%)	60%	74%
Pre-menopausal	276 (78%)	62%	76%	285 (79%)	51%	69%
Tumor Size						
T0-T2	311 (87%)	63%	76%	315 (88%)	53%	71%
T3-T4	14 (4%)	17%**	41%	22 (6%)	18%**	46%

*N represents the number of patients randomized who fall in the indicated subcategories. % represents the percentage of patients in each category created by the indicated covariate.

**In the T3-T4 category, all the relapses observed occurred before completion of the fifth year. These estimates are at 4.3 years for the CEF group and at 4.9 years for the CMF group.

***In the T3-T4 category for CMF treatment, all the deaths observed occurred before completion of the fifth year. This estimate is at 3.0 years.

Although not the primary focus in evaluating the efficacy of CEF, using summaries in Table 1 and Figures 1A-1J and 2A-2J, the effect of the covariates within each treatment can be assessed. It is reassuring to note that most of these relationships trend in the scientifically anticipated direction. The following relationships are not statistically significant findings (adequate power within each stratum was not expected and generally was not achieved).

- (1.) Within both the CEF and CMF groups, 5-year relapse free survival and overall survival are better for the 1-3 positive nodes group than for the 4 or more positive nodes group.
68% vs. 52% and 62% vs. 39% for RFS for CEF and CMF, respectively and
82% vs. 69% and 78% vs. 58% for overall survival for CEF and CMF, respectively

- This result can be seen graphically by noting that the slopes of the survival curves in Figures 1B, 2B (4 or more nodes) are steeper than those in Figure 1A, 2A (1 – 3 nodes).
- (2.) Within both the CEF and CMF groups, 5-year relapse free survival and overall survival are slightly better for the partial mastectomy group than for the total mastectomy group.
 66% vs. 58% and 56% vs. 50% for RFS for CEF and CMF, respectively and
 82% vs. 73% and 76% vs. 64% for overall survival for CEF and CMF, respectively
 This result can be seen graphically by noting that the slopes of the survival curves in Figures 1H, 2H (total mastectomy patients) are slightly steeper than are those in Figure 1G, 2G (partial mastectomy patients).
- (3.) There was no clear indication that receptor status was indicative of relapse in either treatment group. However, overall survival seemed to be better for women with positive receptor status.
 61% vs. 83% and 59% vs. 70% for overall survival for CEF and CMF, respectively
 This result is illustrated in Figures 2C and 2D where the slope of the survival curves for women with positive receptor status is much more gradual than that for women with negative receptor status.
- (4.) Within the CEF group, there is no indication that menopausal status is indicative of relapse or death. Within the CMF group there is a suggestion that peri-menopausal patients do better in terms of relapse free survival and overall survival than do pre-menopausal patients.
 60% vs. 51% for relapse free survival in the CMF group
 74% vs. 69% for overall survival in the CMF group
 This result is illustrated in Figures 1E-F and 2E-F where the survival curve for CMF is slightly above CEF for peri-menopausal patients.
- (5.) Within both the CEF and CMF groups, 5-year relapse free survival and overall survival are better for the T0-T2 than for the T3-T4 group.
 63% vs. 17% and 53% vs. 18% for RFS for CEF and CMF, respectively and
 78% vs. 41% and 71% vs. 46% for overall survival for CEF and CMF, respectively
 This result can be seen graphically by noting that the slopes of the survival curves in Figures 1J, 2J (T3-T4 group) are steeper than are those in Figures 1I, 2I (T3-T4 group).

Note that the findings discussed in #1 to #5 above implies that number of positive nodes and tumor size will be predictive of relapse free survival. And that number of positive nodes, tumor size, and receptor status will be predictive of overall survival. For partial versus total mastectomy, findings point to a possible interaction between treatment and type of surgery, specifically that CMF effects relapse free survival and overall survival differently in peri-menopausal patients than in pre-menopausal patients. Except for the interaction between treatment and menopausal status (for which the power would be very low), these results are confirmed in the sponsor's Cox proportional hazards modeling.

**APPEARS THIS WAY
ON ORIGINAL**

Figure 1: Relapse Free Survival
(Overall Group)

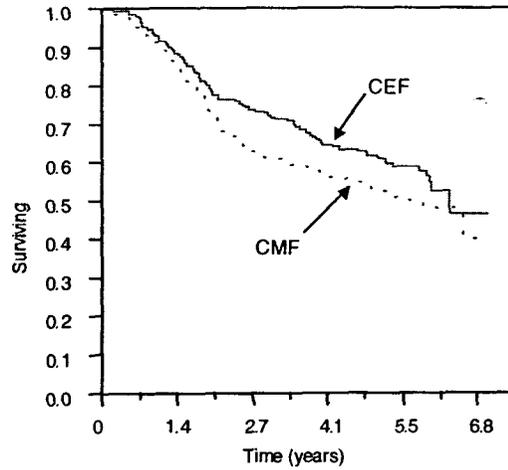


Figure 1A: Relapse Free Survival
(1 - 3 Positive Nodes)

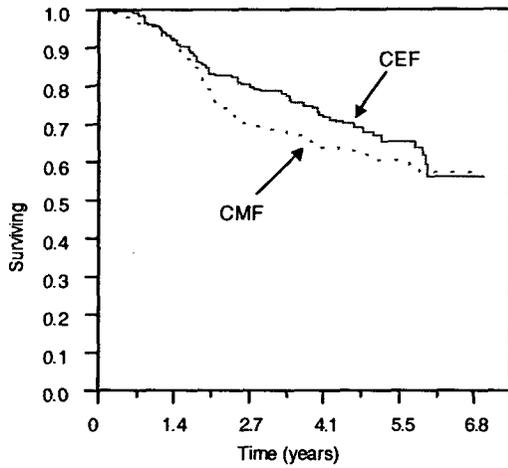


Figure 1B: Relapse Free Survival
(4 or More Positive Nodes)

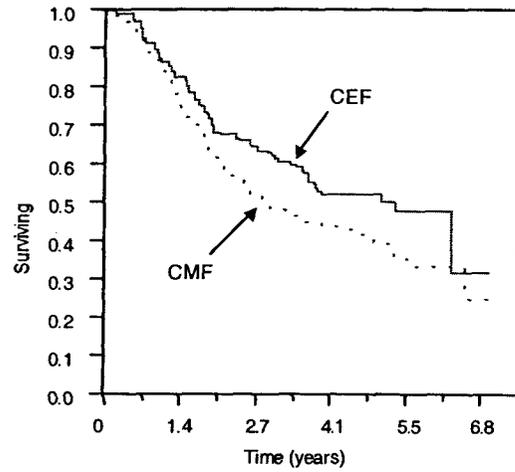


Figure 1C: Relapse Free Survival
(Negative ER/PR Receptor Status)

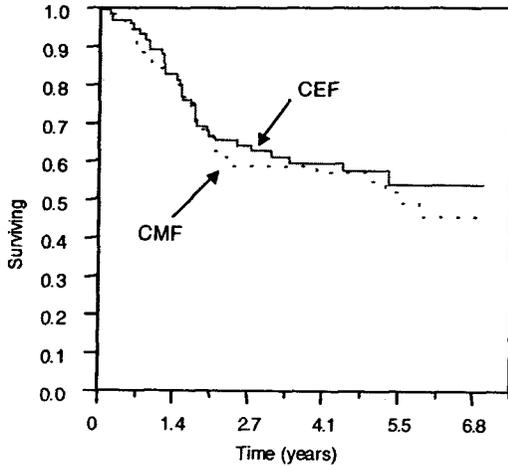


Figure 1D: Relapse Free Survival
(Positive ER/PR Receptor Status)

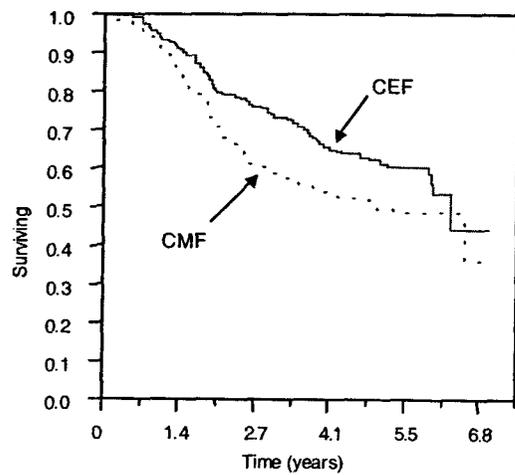


Figure 1E: Relapse Free Survival
(Pre-Menopausal Patients)

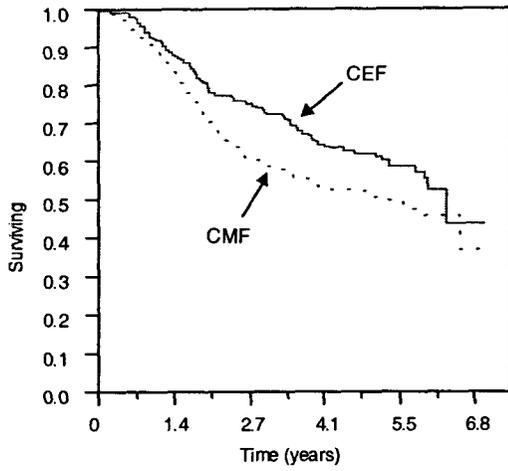


Figure 1F: Relapse Free Survival
(Peri-Menopausal Patients)

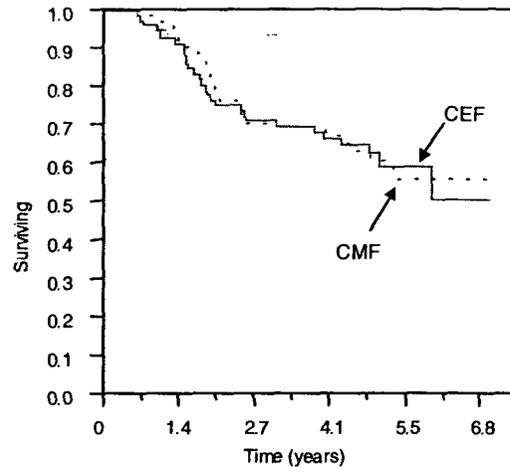


Figure 1G: Relapse Free Survival
(Partial Mastectomy Patients)

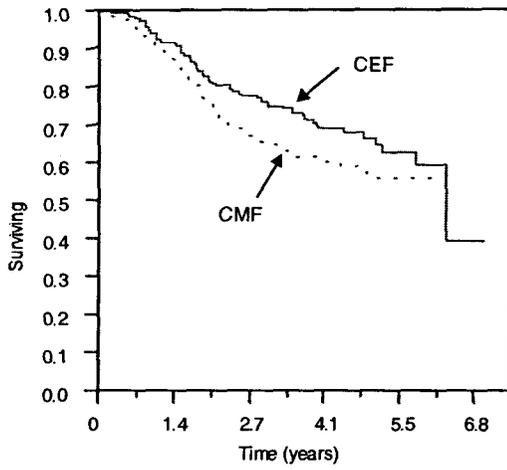


Figure 1H: Relapse Free Survival
(Total Mastectomy Patients)

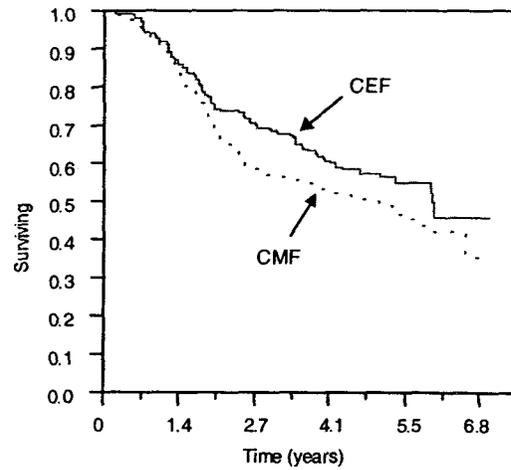


Figure 1I: Relapse Free Survival
(Clinical T Stage T0 - T2)

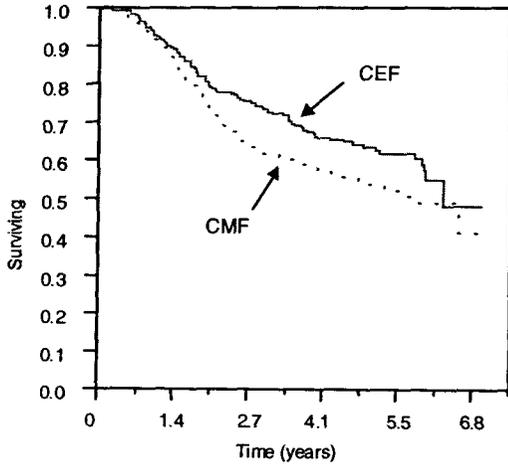


Figure 1J: Relapse Free Survival
(Clinical T Stage T3 - T4)

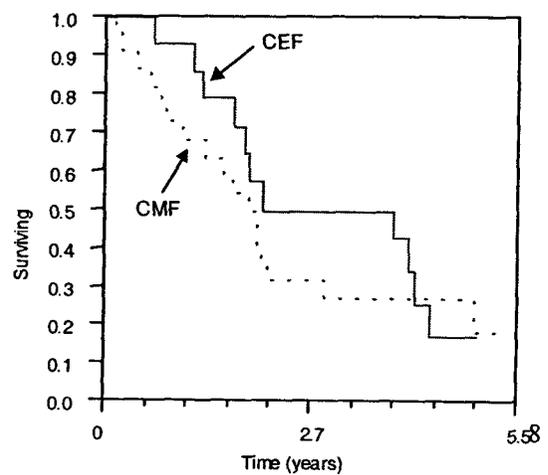


Figure 2: Overall Survival
(Overall Group)

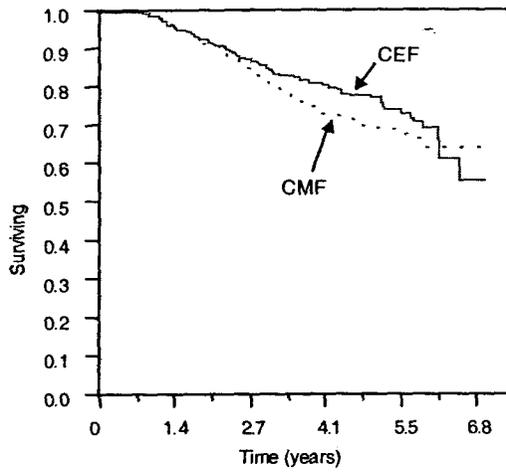


Figure 2A: Overall Survival
(1 - 3 Positive Nodes)

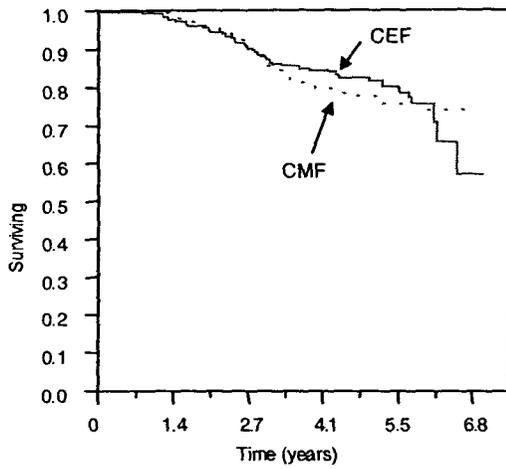


Figure 2B: Overall Survival
(4 or More Positive Nodes)

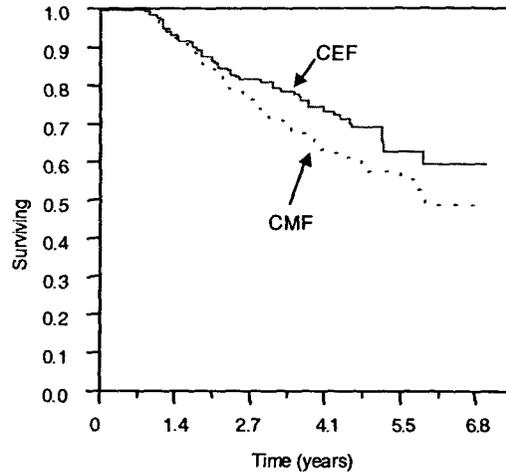


Figure 2C: Overall Survival
(Negative ER/PR Receptor Status)

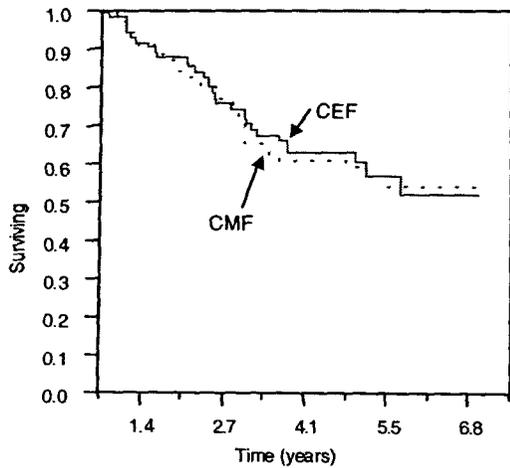


Figure 2D: Overall Survival
(Positive ER/PR Receptor Status)

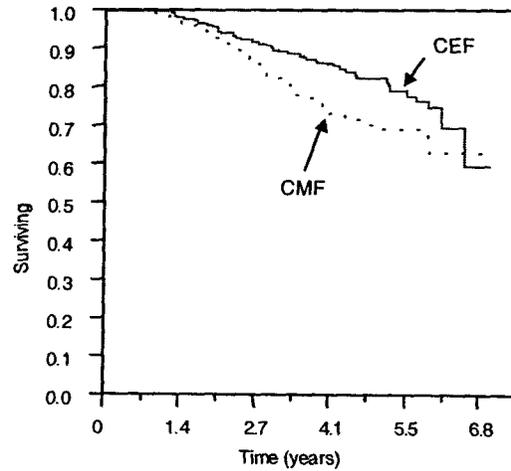


Figure 2E: Overall Survival
(Pre-Menopausal Patients)

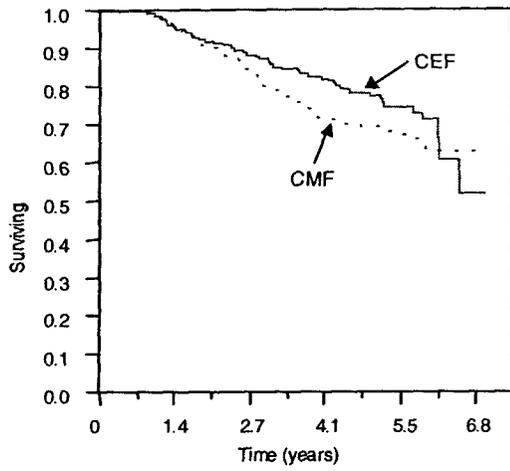


Figure 2F: Overall Survival
(Peri-Menopausal Patients)

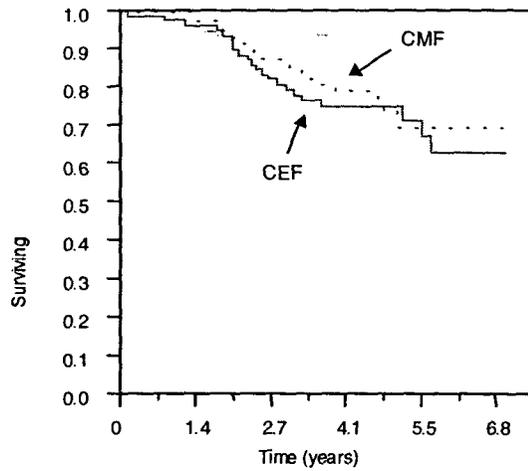


Figure 2G: Overall Survival
(Partial Mastectomy Patients)

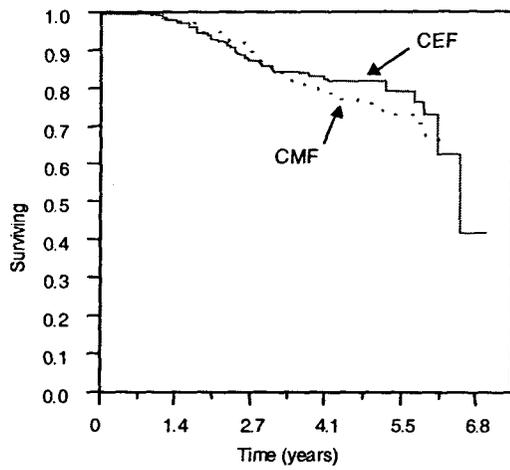


Figure 2H: Overall Survival
(Total Mastectomy Patients)

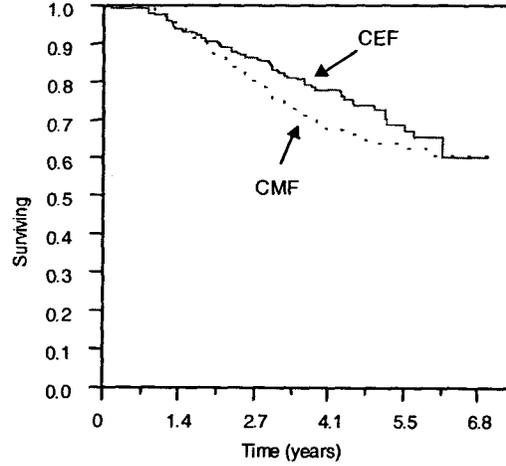


Figure 2I: Overall Survival
(Clinical T Stage T0 - T2)

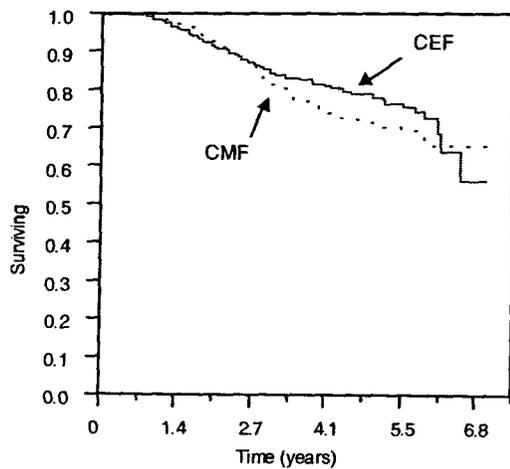
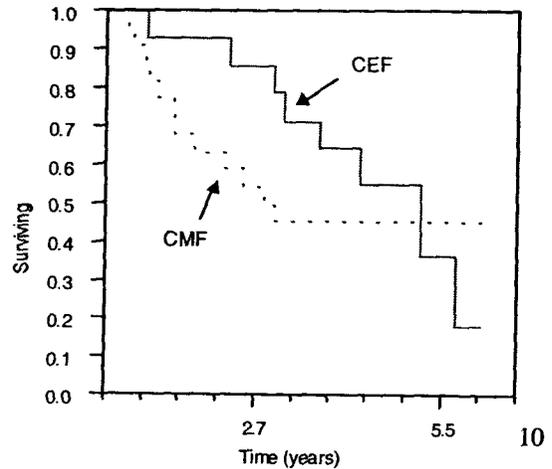


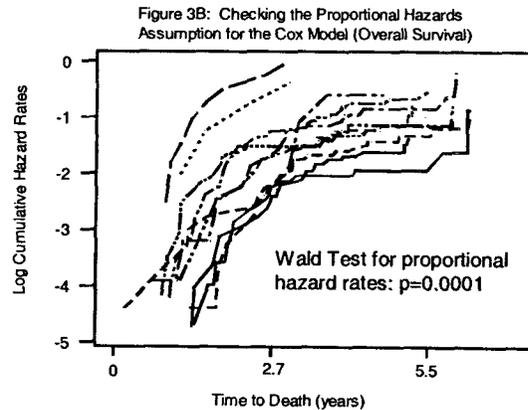
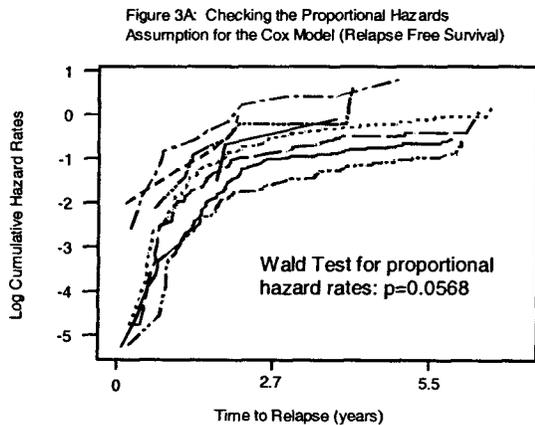
Figure 2J: Overall Survival
(Clinical T Stage T3 - T4)



Sponsor's Adjusted Analysis and Reviewer's Comments:

To evaluate the effect of covariates the sponsor used Cox proportional hazard modeling. The use of Cox proportional hazard modeling was provisioned for in the protocol; however, the covariates to be used in the model were not specified. Forward stepwise procedures were used to build the model at the time of data analysis. These procedures resulted in a **model for relapse free survival containing factors for treatment, number of positive nodes, and tumor size**. The sponsor's model for **overall survival contained factors for treatment, number of positive nodes, tumor size, and receptor status**.

When using Cox proportional hazards modeling, it is necessary to assume that the hazard functions for all strata are proportional to one another. This is a strong assumption and must be verified. Figures 3A and 3B provide a graphical assessment of the validity of this assumption. If the assumption is met, we would expect to see parallel lines in these plots. The Wald chi-squared test was used to formally test the proportional hazards assumption. The p-value for that test is displayed in each graph. Based on Figures 2A and 2B and the Wald test, it is the opinion of this reviewer that the validity of the proportional hazards assumption is questionable (which is not too surprising given the large number of strata being examined). In addition, the terms to be used in the model were not specified in the protocol. Therefore, it is the opinion of this reviewer that the log rank test and results by strata previously presented in this review are preferable to the modeling results. For completeness, however, the results of the Cox modeling are being presented below. Note that a Cox model with treatment as the only covariate will give essentially equivalent results to the log rank test in the case where few ties in survival times are observed.



Since the covariates for the Cox models were not pre-specified, there are concerns regarding their possible data dependency. To address this issue, it is important to note first that since none of these factors were related to treatment at baseline, confounding of the treatment effect from these factors is not likely. The fact that these variables are important for predicting relapse free survival or overall survival appears to be independent of the treatment effect. To illustrate this, the model containing only treatment is also presented here showing results fairly consistent with the full model.

The results in Table 2 indicate that risk of relapse in the CEF group is a statistically significantly lower than the risk of relapse in the CMF group using either model. The full model shows a more favorable result for the treatment effect on overall survival than does the treatment only model; the treatment effect is statistically significant in the full model ($p=0.0204$) but not in the treatment only model ($p=0.1330$). *Since the covariates were not prespecified in the protocol, greater weight should be given to the unadjusted analysis (i.e., treatment only model) and the protocol specified stratified log rank procedure.*

Table 2: Cox Proportional Hazards Modeling – Relapse Free Survival, Overall Survival

Relapse Free Survival			
	Hazard Ratio*	95% Confidence Limits	Wald Chi-Square p-value
Full Model			
Treatment (CEF/CMF)	0.752	(0.595, 0.951)	0.0175
Number of Positive Nodes	1.699	(1.344, 2.146)	0.0001
Tumor Size	2.491	(1.679, 3.694)	0.0001
Treatment Only Model			
Treatment (CEF/CMF)	0.765	(0.601, 0.959)	0.0204
Overall Survival			
	Hazard Ratio*	95% Confidence Limits	Wald Chi-Square p-value
Full Model			
Treatment (CEF/CMF)	0.714	(0.523, 0.976)	0.0344
Number of Positive Nodes	1.706	(1.251, 2.327)	0.0007
Tumor Size	2.477	(1.542, 3.978)	0.0002
Receptor Status	2.000	(1.447, 2.764)	0.0001
Treatment Only Model			
Treatment (CEF/CMF)	0.805	(0.607, 1.068)	0.1330

*Definition of hazard ratios

Treatment: CEF/CMF

Number of Positive Nodes: 4 or more + nodes / 1-3 + nodes

Tumor Size: T3-T4 / T0-T2

Receptor Status: Negative / Positive

Sponsor's Quality of Life Analysis:

Regarding the analysis of the quality of life endpoint, the sponsor acknowledges that, “the implications of missing values in the analysis of longitudinal data and their impact on conclusions driven by ignoring cases with missing observations ... remain a major methodological challenge”. In that spirit, the sponsor conducted analyses on the overall data set as well as on two defined subsets of the data set. The conclusions from each of the three data sets were very similar.

The sponsor’s “first step analysis” was performed on 117 patients all of whom had complete quality of life data. Using a repeated measures analysis of variance model with terms for treatment, visit, and the interaction between treatment and visit, the difference between treatments was not statistically significant but the interaction between treatment and visit was marginally statistically significant (p=0.08). Note that in the sponsor’s FIG 1a and FIG 1b, there is an initial decrease in the quality of life score from baseline after starting chemotherapy that tended to be more pronounced in the CEF arm. During the six monthly cycles of treatment, the curves for each treatment arm are roughly parallel with the CEF results being lower. After therapy completion, a large increase in the score was observed for both treatment groups and the quality of life in the CEF and CMF arms appear to be very similar to each other.

The “second step analysis” involved 211 patients including the 117 patients in the first step analysis and 94 subjects without lost visits but with some incompleteness in the questionnaire responses. The general patterns observed in this analysis are similar to those observed in the first step analysis (see sponsor’s FIG 2a and FIG 2b). There were lower quality of life scores in the CEF group during chemotherapy administration and the treatment by visit interaction became statistically significant (p=0.0014).

The “third step analysis” included all available data (715 patients). The patterns seen in FIG 3a and FIG 3b are similar to those seen previously but the quality of life scores during chemotherapy treatment appear slightly lower than in the previous analyses, especially in the CEF group. Overall, the sponsor reports that the CEF group’s quality of life is statistically significantly lower than that of the CMF group (p=0.001). Overlapping of the quality of life measures for each treatment group is postponed until approximately 15 months after baseline (rather than at 9 months as was indicated in the two previous analyses). As expected, the interaction between treatment and visit was statistically significant

(p=0.0001).

Overall, the sponsor concludes that, "In all the three analyses the estimated BQC-score means of CEF were consistently lower than the ones of CMF at the treatment cycles and at the immediately subsequent follow up visit, approximately until the 15th month since randomization".

Reviewer's Comments on Sponsor's Quality of Life Analysis:

(1) *The QOL measure used was an unweighted average of 30 questions. This gives all of the component questions equal weight. This justification for assuming equal weights was not discussed by the sponsor.*

(2) *The primary analysis for quality of life was not specified in the protocol. They present three analyses: (i) complete cases only (ii) complete cases + intermittent missingness due to incomplete questionnaires and (iii) all available data.*

- (i.) *The analysis including complete cases only is very likely biased. Only 117/715 (16%) of patients had complete data.*
- (ii.) *Analysis of the complete cases and cases with intermittent missingnesses will also be biased depending on the amount of missing data and type of missing data mechanism.*
- (iii.) *The analysis involving all available data will also be problematic if the missing data mechanism is informative. The sponsor states that "if the missingness process were informative, one could expect, as a whole, lower estimates." In the opinion of this reviewer, this may or may not be true. Prediction of the size or directionality of the estimates in such a case; would depend on the specific missingness pattern by treatment arm.*

(3) *The sponsor cites the statistical literature on the pattern mixture model (Little, 1993 and 1995) for assessing whether or not the missingness mechanism is informative, but they did not employ it. Without such an assessment, all of the analyses are open to question. However, all three analyses do show a consistent pattern over time which indicates poorer performance for the CEF arm.*

APPEARS THIS WAY
ON ORIGINAL

FIG 1a. 1° Step Analysis (including only Complete cases)

Model-based response profiles (marginal means and std. err.)

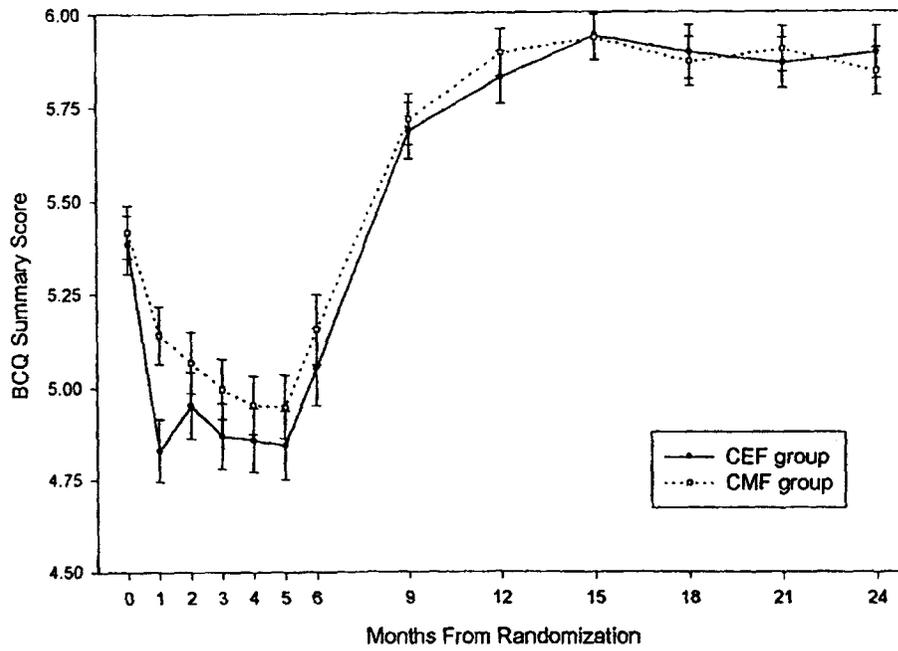


FIG 1b. 1° Step Analysis (Including only Complete cases)

Empirical response profiles (Raw means and std. err.)

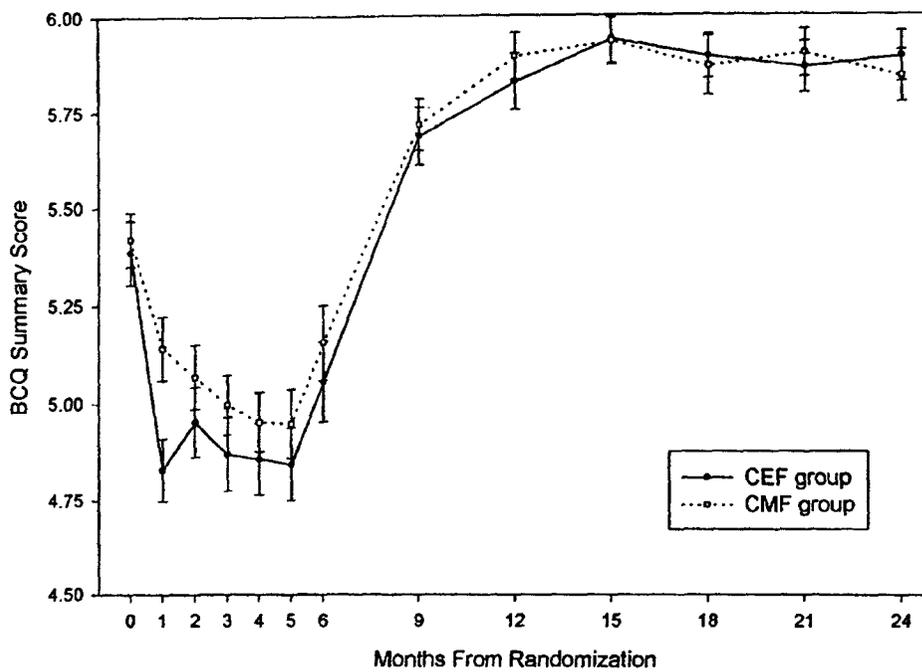


FIG 2a. 2° Step Analysis (Including intermittent missing due to incomplete questionnaire)

Model-based response profiles (marginal means and std. err.)

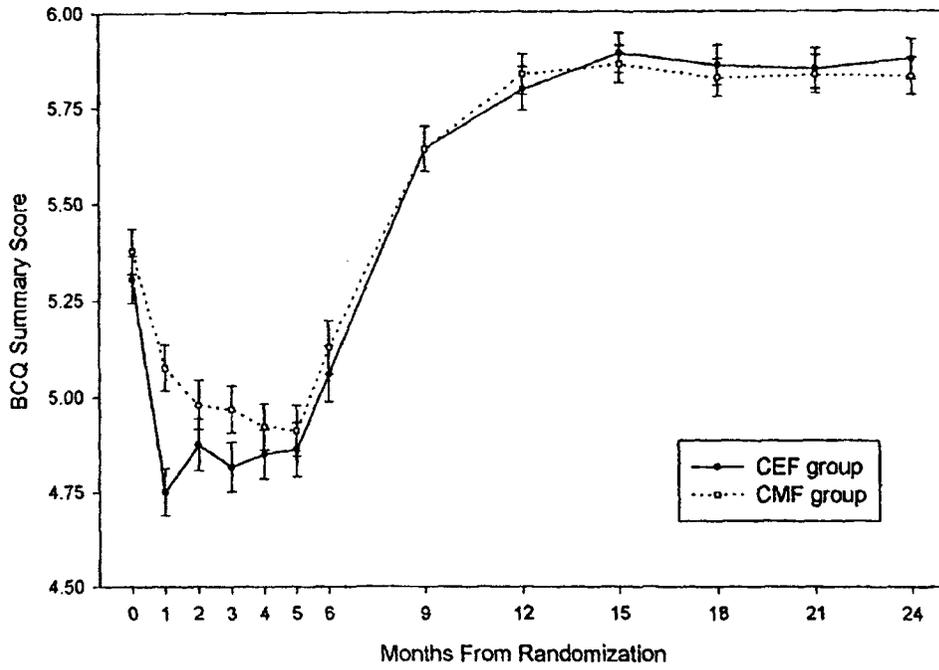


FIG 2b. 2° Step Analysis (Including intermittent missing due to incomplete questionnaire)

Empirical response profiles (raw means and std. err.)

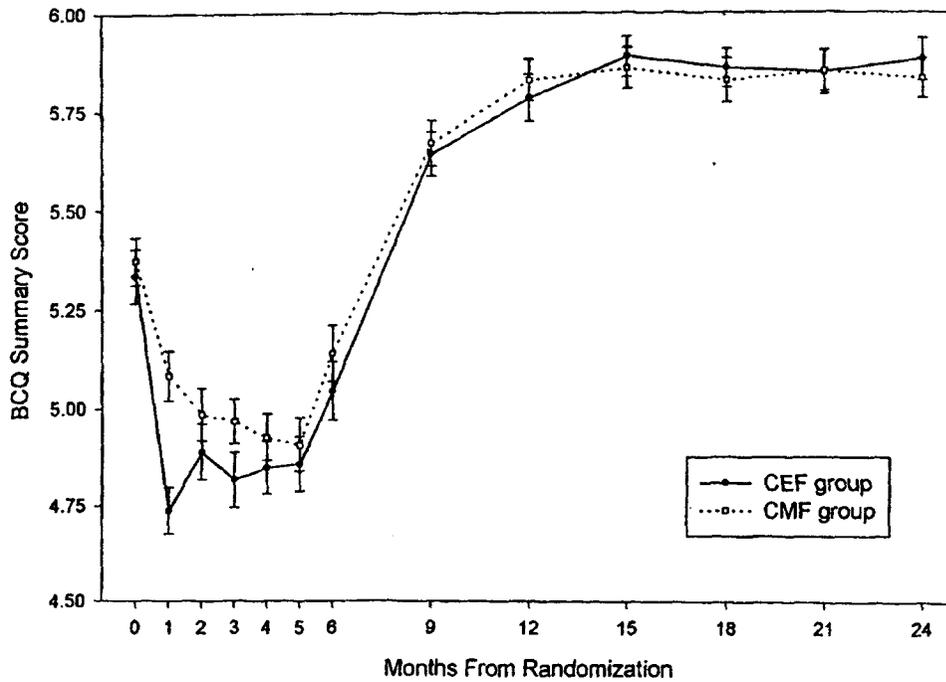


FIG 3a. 3° Step Analysis (Including all available data)

Model-based response profiles (marginal means and std. err.)

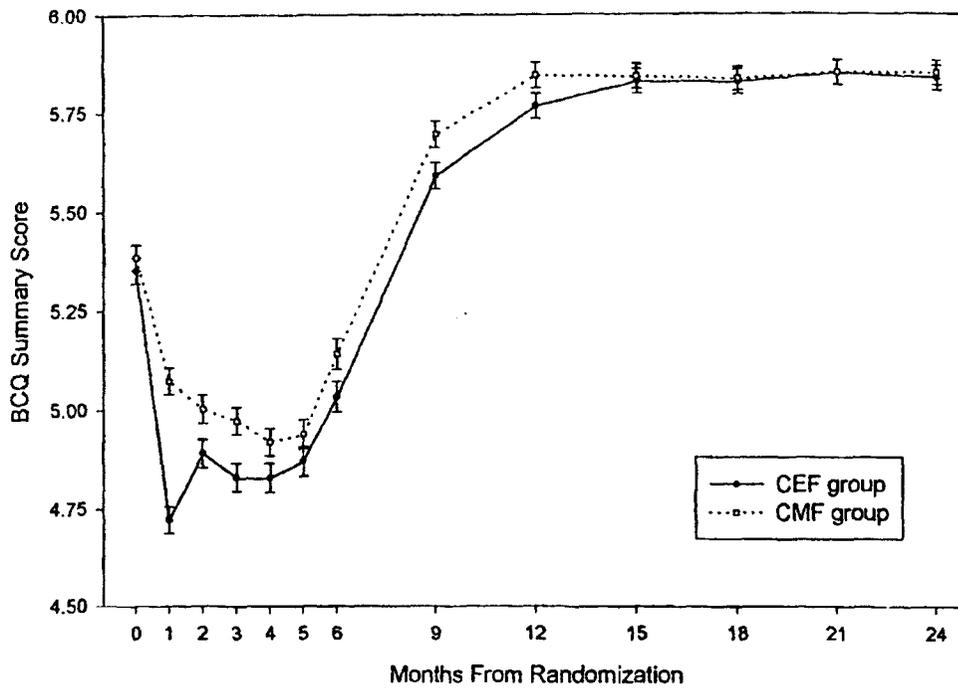
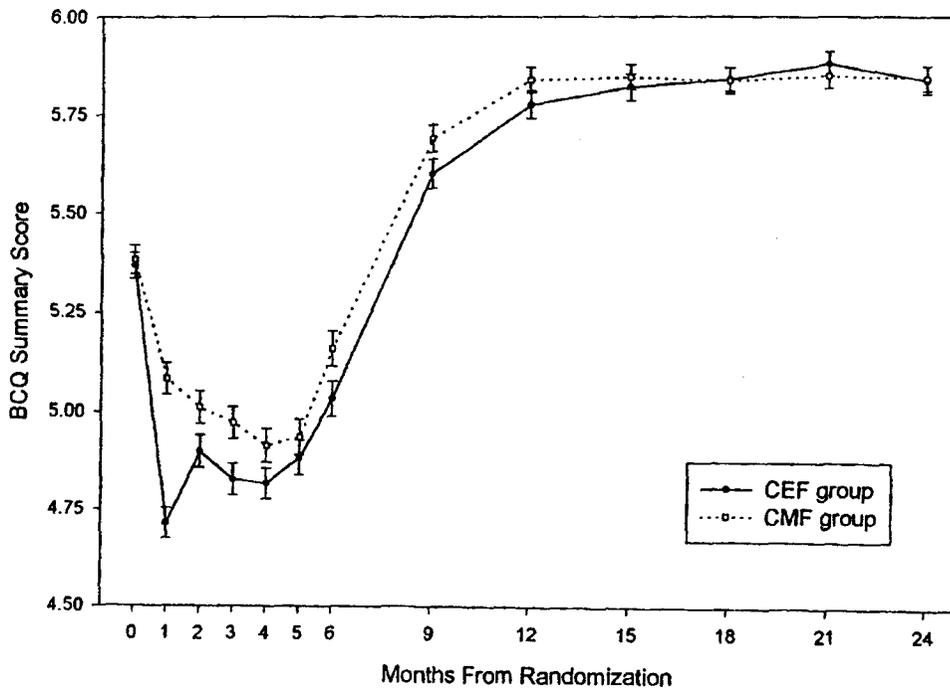


FIG 3b. 3° Step Analysis (Including all available data)

Empirical response profiles (Raw means and std. err.)



III. Study HEPI013

III.A. Study Design / Analysis Plan – Study HEPI013

Study HEPI013 was a multicenter (48 centers), open label, phase III trial in adult women. Subjects were required to have histologically proven breast cancer with measurable and/or evaluable metastatic disease (stage IV excluding inflammatory breast cancer) at diagnosis or recurrent disease following total mastectomy and axillary dissection located outside of previously irradiated fields. Subjects were stratified by center, prior adjuvant chemotherapy (yes, no), presence of visceral metastases (yes, no), and number of organs involved by distant metastases (1-2 vs. >2) and were randomized to either CEF or CMF therapy.³ All drugs were administered on day 1 and day 8 of a 3-weekly cycle. Patients were to receive six cycles of therapy. Follow-up of each subject continued until the end of the study or until death which ever occurred first.

The **primary efficacy endpoint** was designated in the protocol to be **time to progression** defined as the time from randomization until progression of disease or death due to any cause whichever occurred first.

Secondary efficacy endpoints include:

- (1.) **response rate** (the proportion of patients with complete or partial response out of the total number of patients considered in the analysis),
- (2.) **time to failure** (failure is defined as disease progression, death, treatment discontinuation due to patient refusal, toxicity or loss to follow-up), and
- (3.) **overall survival** (time from randomization to death).

Comparisons of **time to progression** for the CEF group versus the CMF group were conducted using the **protocol specified log rank test and plots of the Kaplan-Meier survival curves**. To evaluate the effect of covariates, Cox proportional hazard modeling was used. The use of Cox proportional hazard modeling was not provisioned for in the protocol. The **response rates** in the two treatment groups were compared using the **chi-square test**. **Odds ratios (and 95% confidence intervals)** were also calculated for the response rates (CEF/CMF). **Time to failure and overall survival** were summarized using **Kaplan Meier estimates** and the two treatment arms compared by the **log-rank test**.

All of the case report forms were reviewed by an **independent oncologist** who re-assessed patient eligibility and evaluability. From the case report forms, the oncologist reviewer assessed the patient's response⁴, date of response, date of progression, and date of death. The sponsor made this statement regarding the handling of discrepancies between the reviewer and the original investigators.

"In case of discrepancy between the reviewer's and the investigator's judgement, as reported in the case report form, the investigator was provided with the reviewer's comments and was given the opportunity to respond. If no reply was received or if it was not considered convincing by the reviewer, the **reviewer's assessment was the one taken into account in the final analysis.**"

The **protocol planned sample size** was based on the following assumptions.

- (1.) median time to progression with CMF was believed to be 8 months,
- (2.) median expected time to progression with CEF was 11 months,
- (3.) accrual time was expected to be 15 months,
- (4.) time to progression was expected to be analyzed 12 months after the last patient was randomized and

3 **CEF Therapy:** Cyclophosphamide 400mg/m²

Epirubicin 50 mg/m²

5-Fluorouracil 500 mg/m²

CMF Therapy: Cyclophosphamide 500 mg/m²

Methotrexate 40 mg/m²

5-Fluorouracil 500 mg/m²

4 Possible categories of response were defined a priori and included: complete response, partial response, no change, or progressive disease.

Response rate = Proportion of patients with complete or partial response

(5.) $\alpha = 0.05$, $\beta = 0.20$, and the use of two-sided tests.

The calculated sample size was 210 patients per treatment group (420 total).

III.B. Sponsor's Results and Reviewer's Comments for Study MA-5

Four hundred sixty women were enrolled in the trial and randomized to receive either CEF (223 patients) or CMF (237 patients). Six (5 in the CEF arm and 1 in the CMF arm) did not receive treatment. The percentage of patients who completed the per-protocol treatment was comparable between groups (33% in CEF group, 30% in CMF group). The reasons for treatment discontinuation are presented in Table 3. Except for progressive disease and cardiac toxicity, the reasons for treatment withdrawal were fairly balanced across treatment groups.

Table 3: Reasons for and Frequency (and Percent) of Treatment Withdrawals by Treatment Group

Reason for Treatment Discontinuation	CEF N (%)	CMF N (%)
Progressive disease	73 (33.5%)	104 (44.1%)
Relapse	1 (0.5%)	0 (0.0%)
Toxicity	11 (5.0%)	9 (3.8%)
Adverse reaction	1 (0.5%)	0 (0.0%)
Cardiac toxicity	15 (6.9%)	1 (0.4%)
Patient refusal	26 (11.9%)	30 (12.7%)
Protocol violation	1 (0.5%)	0 (0.0%)
Death	7 (3.2%)	9 (3.8%)
Lost to follow-up	5 (2.3%)	5 (2.1%)
Other	6 (2.8%)	7 (3.0%)

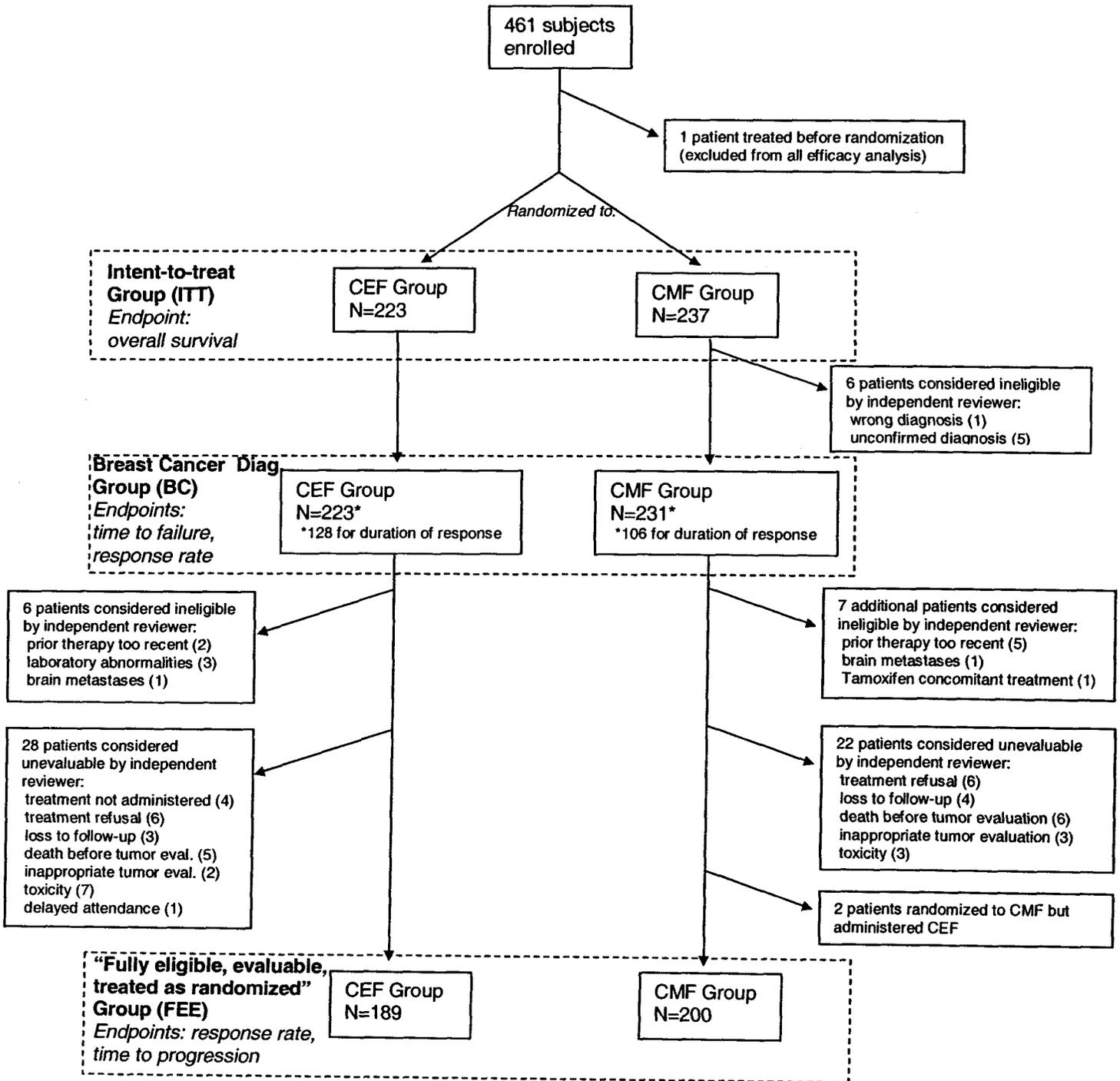
Patients' baseline characteristics were compared across treatment groups. The demographic factors examined included: age at study entry, age at first diagnosis, performance status, ethnic group, menopausal status, histological subtypes, receptor status, tumor staging, disease free interval (time from surgery to first relapse), prior treatment, and disease characteristics. No significant differences were found between the two treatment groups, except for age at the time of first diagnosis. The median ages at first diagnosis were 53 years (range 21-70) and 51 years (range 23-70) in the CEF and CMF groups, respectively. The number of patients whose age at first diagnosis fell into each of the following categories is indicated by treatment arm. A higher frequency of patients sixty years old or over were in the CEF group compared with the CMF group ($p=0.009$).

Age at first diagnosis < 50 years –	CEF: 39.0%, CMF: 46.4%
Age at first diagnosis 50-59 years –	CEF: 31.4%, CMF: 36.7%
Age at first diagnosis \geq 60 years –	CEF: 28.7%, CMF: 16.9%
Data not available –	CEF: 0.9%, CMF: 0.0%

The protocol stated that primary analyses of time to progression (i.e., the protocol specified primary endpoint), response rate, and overall survival would be conducted using the intent-to-treat group but that the same analyses would be repeated in an "evaluable group". The protocol indicated that the analysis of duration of response would include all patients who achieve complete response or partial response. However, instead of presenting the result of all the intent-to-treat analyses, the study report presents results for certain endpoints using various subsets of the original set of patients. Overall survival was analyzed using the intent-to-treat group. Response rate and time to failure results are based on the group of patients who have histologically proven breast cancer. Duration of response is summarized for the subset of the histologically proven breast cancer patients who achieved complete or partial response. **Time to progression (primary endpoint) and response rate were evaluated in the eligible and evaluable group.** Figure 1 illustrates how these data sets were established. The "fully eligible, evaluable, treated as randomized" group is smaller than the intent-to-treat group by about 15% in each treatment arm.

Since excluding patients from analyses (especially post-hoc) can introduce bias in the results, this reviewer conducted the previously described analyses in the intent-to-treat group. Although the following section will primarily present the sponsor's analyses (conducted in the indicated subgroups), differences between the subset analyses and the ITT analyses will be pointed out when they seem to be material.

Figure 1: Efficacy Analysis Groups



Efficacy Results and Reviewer's Comments for Time to Progression (primary endpoint)

The results of the log rank test and the KM curves depicted in Figure 4 show a **statistically significant prolongation of time to progression in the CEF group compared to the CMF group** (FEE group, $p=0.0064$; ITT group, $p=0.0204$). In the FEE analysis group, the median time to progression was 8.9 months in the CEF group and 6.3 months in the CMF group. The median time to progression was 8.6 months in the CEF group and 6.3 months in the CMF group for the ITT analysis.

The **within strata relationships between CEF and CMF are very similar to the overall results** for time to progression. Figures 4A through 4D show within-strata KM survival curves for the FEE analysis group by treatment.

Although not the primary focus in evaluating the efficacy of CEF, using Figures 4A through 4D the effect of covariates can be assessed. It is reassuring to note that these relationships trend in the scientifically anticipated direction. Comparison of Figures 4A and B show a steeper slope in the survival curves (in both the CEF or CMF group) for patients with visceral metastases compared to those without visceral metastases. In addition, a steeper slope for the survival curve is evident for the group with >2 sites than for the group with 1-2 sites (Figures 4C and D). The impact of the covariates discussed here is confirmed using Cox modeling.

APPEARS THIS WAY
ON ORIGINAL

APPEARS THIS WAY
ON ORIGINAL

Figure 4: Time to Progression
(Fully Eligible, Evaluable, Treated
as Randomized (FEE) Group)

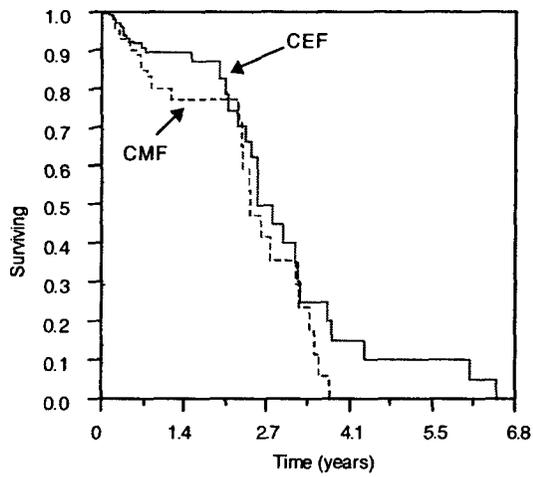


Figure 4A: Time to Progression
(FEE Group without Visceral Metastases)

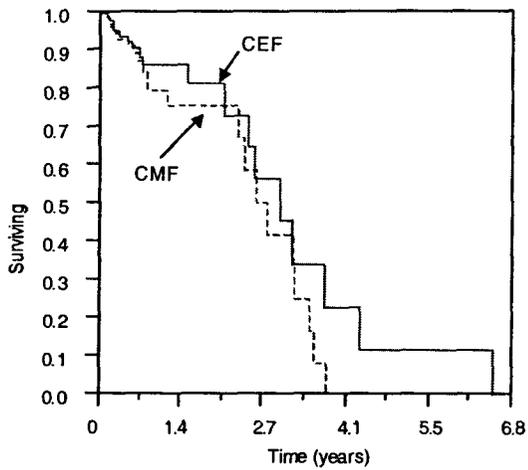


Figure 4B: Time to Progression
(FEE Group with Visceral Metastases)

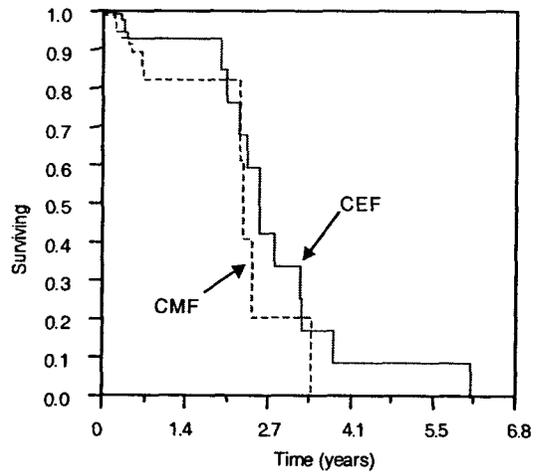


Figure 4C: Time to Progression
(FEE Group with 1-2 Sites)

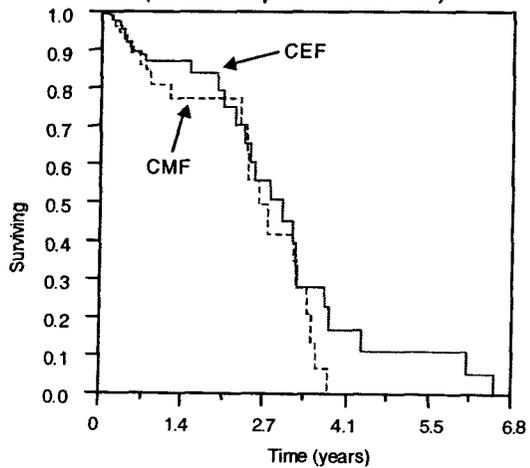
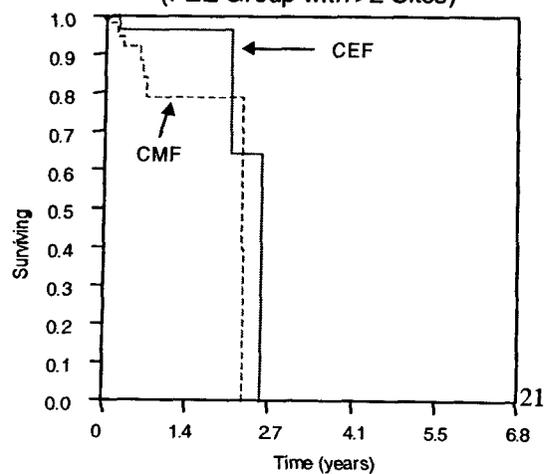


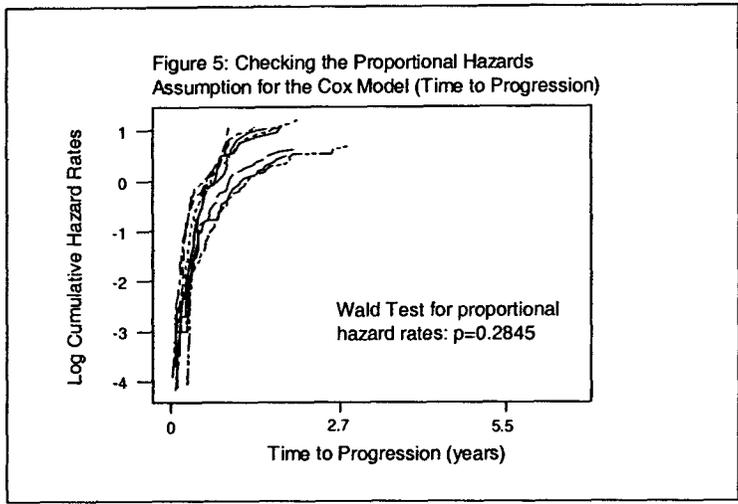
Figure 4D: Time to Progression
(FEE Group with >2 Sites)



Sponsor's Adjusted Analysis:

To evaluate the effect of covariates the sponsor used Cox Proportional Hazards modeling. The use of Cox proportional hazard modeling was not provisioned for in the protocol. Backwards stepwise procedures were used to build the model at the time of data analysis. These procedures resulted in a **model for time to progression containing factors for treatment, dominant metastases (visceral vs. non-visceral), and number of sites (1-2 sites vs. more than 2 sites).**

When using Cox proportional hazards modeling, it is necessary to assume that the hazard functions for all strata are proportional to one another. This is a strong assumption which needs to be verified. Figure 5 provides a graphical assessment of the validity of this assumption. When the assumption is met, one would expect to see parallel lines in this plot. The Wald chi-squared test was used to formally test the proportional hazards assumption. The p-value for that test is displayed on the graph in Figure 5. Based on Figure 5 and the Wald test, it is the opinion of this reviewer that the proportional hazards assumption may be valid in this case. However, since the use of Cox modeling (and the terms used in the model) was not specified in the protocol, it is the opinion of this reviewer that the log rank test and results by strata previously presented in this review are preferable to the modeling results. For completeness however, the results of the Cox modeling are being presented below.



Since the covariates for the Cox model were not pre-specified, there are concerns regarding their possible data dependency. To address this issue, it is important to note first that since neither visceral metastases or number of sites were related to treatment at baseline, confounding of the treatment effect from these factors is not likely. The fact that these variables are important for predicting time to progression appears to be independent of the treatment effect. To illustrate this, the model containing only treatment is also presented here showing results fairly consistent with the full model. The results for both the FEE group and the ITT group are included in Table 4 and are fairly consistent with one another. Table 4 indicates that the risk of progression in the CEF group is statistically significantly lower than the risk of progression in the CMF group.

Table 4: Cox Proportional Hazards Modeling – Time to Progression

	Hazard Ratio		95% Confidence Limits		Wald Chi-Square p-value	
	FEE group	ITT group	FEE group	ITT group	FEE group	ITT group
Full Model						
Treatment (CEF/CMF)	0.720	0.751	(0.578, 0.898)	(0.608, 0.927)	0.0036	0.0076
Visceral Metastases (with/without)	1.395	1.442	(1.106, 1.759)	(1.153, 1.802)	0.0050	0.0013
Number of sites (>2/1-2)	1.680	1.723	(1.326, 2.128)	(1.376, 2.157)	0.0001	0.0001
Treatment Only Model						
Treatment (CEF/CMF)	0.737	0.780	(0.591, 0.919)	(0.632, 0.963)	0.0067	0.0209

Efficacy Results for Response Rate, Overall Survival, and Time to Failure (secondary endpoints)

Using the BC analysis group, the treatment **response rate** (complete response + partial response) was **statistically significantly higher in the CEF group compared to the CMF group** ($p=0.01$). The odds ratio (CEF/CMF) was 1.6 with 95% confidence interval (1.1, 2.3).

	CEF	CMF	Row Totals
Responded*	128	106	234
Failed to respond**	95	125	220
Column Totals	223	231	454

*Patients who achieved complete or partial response to treatment were considered responders.

**Patients who had stable disease, progressive disease, or were not evaluable were considered treatment failures.

Using the ITT analysis group, **overall survival was not statistically significantly different for the CEF and CMF groups** ($p=0.2372$). Figure 6 displays the KM estimates by treatment. Using the BC analysis group, **time to treatment failure was statistically significant longer in the CEF group than in the CMF group** ($p=0.0008$). Figure 7 displays the KM estimates by treatment.

Figure 6: Overall Survival

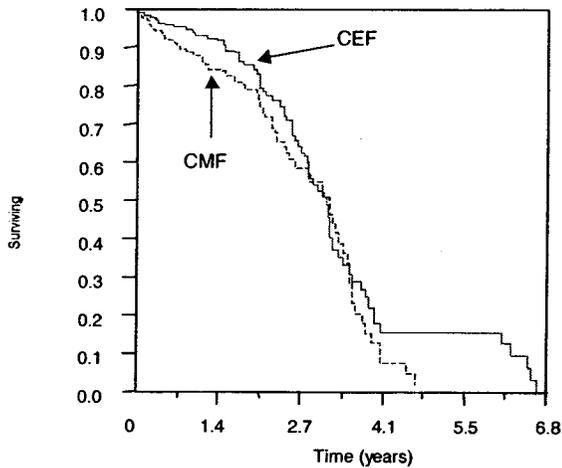
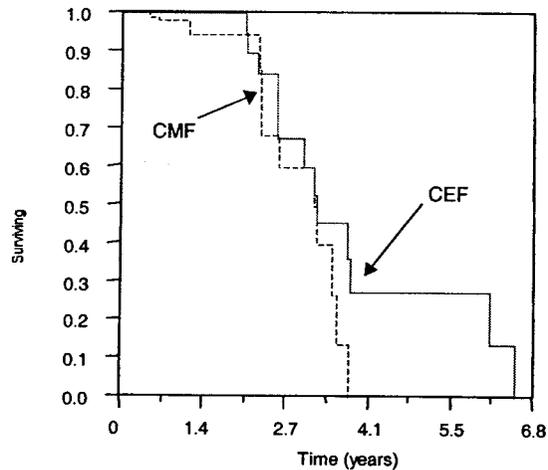


Figure 7: Time to Treatment Failure



Meta-Analysis Comparing the Efficacy of Epirubicin to Doxorubicin

As doxorubicin is believed to convey approximately six months in survival benefit in the first-line treatment of breast cancer, a new product should not lose that advantage. Therefore, FDA requested that the sponsor compare (using existing literature) the efficacy of Epirubicin to Doxorubicin. The sponsor identified over 2800 papers through a literature search utilizing MEDLINE, CANCELIT, BIOS, DERWENT, and EXCEPTA MEDICA databases. Six papers were ultimately selected for inclusion in the meta-analysis based on the following factors:

- (1.) Papers in English language
- (2.) Phase III randomized trials
- (3.) Epirubicin versus Doxorubicin as single agent and/or in combination
- (4.) Chemotherapy-naïve patients with advanced breast cancer
- (5.) All dose and schedules
- (6.) No abstract or review articles

A total of 1257 randomized patients were reported in these six studies. Six hundred thirty five patients had been randomized to receive Epirubicin and 622 to receive Doxorubicin. Of these, a total of 575 (90%) in the Epirubicin group and 566 (90%) in the Doxorubicin group were evaluable for efficacy. The sponsor reported an overall odds ratio for response rate of Doxorubicin versus Epirubicin which was 0.87 with 90% confidence interval (0.72, 1.06). In addition, an overall odds ratio for overall survival was reported. This odds ratio for Doxorubicin:Epirubicin was 0.98 with 90% confidence interval (0.80, 1.20). Both of these odds ratios were calculated at the median survival time (which was approximately 18 months).

It is the opinion of this reviewer as well as being noted by the sponsor that this analysis may be subject to criticism for several reasons.

(1.) The analysis uses only one point on the survival curve (i.e., 18 months).

(2.) This approach ignores the effect of censoring.

(3.) And finally, the usual problems associated with a meta-analysis such as publication bias, selection bias, etc. are of concern.

Therefore it is the opinion of this reviewer that this meta-analysis should be considered an exploratory analysis and the results should therefore be interpreted and utilized cautiously.

IV. Overall Conclusions

In the opinion of this reviewer, the results from study MA5 demonstrate that CEF provides statistically significantly longer relapse free survival and overall survival in premenopausal women with operable axillary node-positive breast cancer when compared to CMF based on the per protocol analysis. The sponsor's adjusted analyses were not prespecified and should be considered exploratory. The quality of life average scores for CEF patients are lower than those for the CMF patients for the duration of treatment based on the sponsor's analysis.

Study HEPI013 showed a statistically significantly prolonged time to progression of disease for the CEF group when compared to the CMF group in patients with stage IV disease (excluding inflammatory breast cancer) or recurrent disease following total mastectomy and axillary dissection located outside previously irradiated fields. Overall survival tended to favor the CEF group over the CMF group but the relationship was not statistically significant. The time to treatment failure was statistically significantly longer in the CEF group than in the CMF group.

Ruthanna C. Davi
Mathematical Statistician

Concur:

Gang Chen, Ph.D.

George Chi, Ph.D.

cc: Archival NDA#21-010
HFD-150/P. Guinn
HFD-150/S. Honig
HFD-150/G. Williams
HFD-344/B. Barton
HFD-710/G. Chen
HFD-710/G. Chi
HFD-715/R. Davi
HFD-715/M. Welch
HFD-715/File Copy
R. Davi/x73122/Word/5/5/99

This review contains 24 pages of text, tables, and figures.

/s/

Ruth Davi
8/3/1999 10:37:20 AM
BIOMETRICS

Gang Chen
3/3/00 10:59:35 AM
BIOMETRICS

George Chi
11/6/00 09:17:47 AM
BIOMETRICS