

**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:
22-042

STATISTICAL REVIEW(S)



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

STATISTICAL REVIEW AND EVALUATION
SECONDARY REVIEW - CLINICAL STUDIES

NDA /Serial Number: 22-042/N_000

Drug Name: Evista ® (raloxifene hydrochloride HCl)

Applicant: Eli Lilly and Company

Indication(s): Reduction in Risk of Invasive Breast Cancer in Postmenopausal Women with Osteoporosis and Reduction in Risk of Invasive Breast Cancer in Postmenopausal Women at High Risk of Breast Cancer

Date(s): Submission Date: November 13, 2006

Review Priority: Standard Review

Biometrics Division: Division of Biometrics V (HFD-711)

Statistical Reviewer: Rajeshwari Sridhara, Ph.D.,
Deputy Division Director/Team Leader
Kun He, Ph.D., Primary Reviewer

Concurring Reviewer: Aloka Chakravarty, Ph.D., Division Director, DBV

Medical Division: Oncology Drug Products (HFD-150)

Clinical Team: Bhupinder Mann, M.B.B.S., Patricia Cortazar, M.D., and John Johnson, M.D.

Project Manager: Ms. Patricia Garvey

Keywords: Non-inferiority, log-rank test, Invasive breast cancer

Introduction

This is a secondary review to the primary statistical review by Kun He, Ph.D. I concur with Dr. He's conclusions and recommendations. Additional comments are presented here in our approach to the problem of collective evidence in this application as a whole.

Overview

In this application, the Applicant is seeking two indications:

- (1) Reduction in the risk of invasive breast cancer in postmenopausal women at high risk for invasive breast cancer and
- (2) Reduction in the risk of invasive breast cancer in postmenopausal women with osteoporosis.

In support of these claims, the Applicant has submitted 4 studies: STAR, RUTH, MORE and CORE. Specifically, the STAR study is intended to support claim of non-inferior efficacy compared to tamoxifen in postmenopausal women at high risk for breast cancer and the RUTH study is intended to support claim of superior efficacy compared to placebo in postmenopausal women with osteoporosis. Because reduction in the incidence of invasive breast cancer was not a pre-specified efficacy endpoint, the results from the MORE and its extension study, CORE conducted in a select sub-group of patients, are considered as hypothesis generating, supportive studies. The RUTH, MORE and CORE were placebo controlled studies and the STAR study compared raloxifene to tamoxifen. Please refer to the review by the primary reviewer for a detailed description and results of these studies.

Non-inferiority Considerations

In a study with non-inferiority hypothesis, three criteria have to be met: (1) that the active control has efficacy, (2) under the assumption that the efficacy has remained constant over time, the active control effect size can be estimated for patients with the indication under consideration, and (3) the percent of active control effect size to be retained can be pre-specified. The treatment effect size of the active control needs to be established based on meta-analysis of historical randomized studies with consideration of the between trial variability among historical studies. Because of the uncertainties in the analysis and interpretation of non-inferiority studies, in general for a new molecular entity to be considered for granting a non-inferiority efficacy claim, two studies are necessary to establish efficacy.

The STAR study was primarily designed to demonstrate superiority of raloxifen over tamoxifen in post-menopausal women with high risk for invasive breast cancer incidence. The non-inferiority analysis using percent retention approach was added to the statistical analysis plan at the end of the study. The study did not demonstrate superiority with respect to reduction in invasive breast cancer when treated with raloxifene compared to tamoxifen.

The historical NSABP P-1 study (N = 13,388) compared tamoxifen with placebo in both pre- and post-menopausal women with high risk for invasive breast cancer incidence. The results of the P-1 trial in all subjects studied are widely published in literature and were the basis for approval of tamoxifen in the treatment of reduction of breast cancer incidence. Tamoxifen is also approved for the treatment of breast cancer. In this application, for the non-inferiority analysis consideration, the tamoxifen effect size was estimated from one single study, based on un-published data on a sub-group of women who were older than 50 years from the P-1 study. Furthermore, the point estimate (RR = 0.47) from this subgroup was considered as the tamoxifen effect size and did not account for the variability in this estimate. The percentage of tamoxifen effect to be retained was also not pre-specified. Thus in the STAR study for non-inferiority consideration only the first of the three criteria mentioned above, namely, tamoxifen is efficacious, was met. It is to be noted that the STAR study was a large, multi-center, randomized study conducted in 19,747 postmenopausal women at high risk for invasive breast.

Collective Evidence

This application is a supplemental application for consideration of approval of raloxifen for the reduction of invasive breast cancer in postmenopausal women (1) with osteoporosis and (2) at high risk for invasive breast cancer. Raloxifen is an approved drug product and was first approved in 1997 for the prevention of osteoporosis and then approved in 1999 for the treatment of osteoporosis. Since then an estimated over 20 million women have been treated with raloxifen. Thus raloxifen has demonstrated benefit in the treatment and prevention of osteoporosis, albeit serious adverse reactions such as, deep vein thrombosis, pulmonary embolism, and possibly stroke death that have been observed and reported.

The data from the RUTH trial along with the supportive data from the MORE and CORE studies support the claim of efficacy with respect to reduction in invasive breast cancer in postmenopausal women with osteoporosis. It is to be noted that raloxifene is already an approved product for this population.

Although there appears to be raloxifene treatment effect with respect to reducing invasive breast cancer incidence in post-menopausal women at high risk for invasive breast cancer based on the results of the STAR study and supported by the three relatively large placebo controlled studies, the precise percentage of retention of tamoxifen effect is debatable as outlined above. Given that these women are healthy subjects, the benefit and risk have to be carefully considered. Whether the benefits outweigh the risks of the use of raloxifene in this population is deferred to clinical judgment.

Rajeshwari Sridhara, Ph.D.
Deputy Division Director/ Team Leader
Date:

Concur: Dr. Chakravarty, Division Director, DB5

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

/s/

Rajeshwari Sridhara
9/10/2007 01:57:35 PM
BIOMETRICS

Aloka Chakravarty
9/10/2007 02:13:10 PM
BIOMETRICS

Appears This Way
On Original



U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH
OFFICE OF TRANSLATIONAL SCIENCES
OFFICE OF BIostatISTICS

STATISTICAL REVIEW AND EVALUATION

Clinical Studies

NDA/Serial Number: 22-042 / N_000
Drug Name: Evista® (Raloxifene Hydrochloride HCl)
Indication: Reduction in Risk of Invasive Breast Cancer in Postmenopausal Women with Osteoporosis and Reduction in Risk of Invasive Breast Cancer in Postmenopausal Women at Risk of Breast Cancer
Applicant: Eli Lilly and Company
Date: November 13, 2006
Review Priority: Standard

Biometrics Division: V (HFD 711)
Statistical Reviewer: Kun He
Concurring Reviewers: Rajeshwari Sridhara, Ph.D., Team Leader
Aloka Chakravarty, Ph.D., Director

Medical Division: Oncology Drug Products (HFD 150)
Clinical Team: Bhupinder Mann, M.B.B.S., Clinical Reviewer
Patricia Cortazar, M.D., Clinical Reviewer
John Johnson, M.D., Team Leader

Project Manager: Patricia Garvey, R.Ph.

Keywords: Invasive breast cancer, log-rank test, non-inferiority analysis

TABLE OF CONTENTS

1. Executive Summary4

1.1 Conclusions and Recommendations..... 4

1.2 Brief Overview of Clinical Studies..... 4

1.3 Statistical Issues and Findings 6

2. Introduction14

2.1 Overview 14

2.2 Data Sources 16

3. Statistical Evaluation.....16

3.1 Evaluation of Efficacy16

 3.1.1 Study H3S-MC-GGIY (STAR).....16

 3.1.1.1 Objective 16

 3.1.1.2 Study Design 17

 3.1.1.3 Efficacy Measures 18

 3.1.1.4 Sample Size Considerations 19

 3.1.1.5 Interim Analysis 21

 3.1.1.6 Statistical Analysis Methods 22

 3.1.1.7 Applicant’s Results and Statistical Reviewer’s Findings/ Comments..... 22

 3.1.2 Study H3S-MC-GGIO (RUTH).....34

 3.1.2.1 Objective 34

 3.1.2.2 Study Design 34

 3.1.2.3 Efficacy Measures 35

 3.1.2.4 Sample Size Considerations 39

 3.1.2.5 Interim Analysis 39

 3.1.2.6 Statistical Analysis Methods 40

 3.1.2.7 Applicant’s Results and Statistical Reviewer’s Findings/ Comments..... 41

 3.1.3 Study H3S-MC-GGG (MORE).....62

 3.1.3.1 Objective 62

 3.1.3.2 Study Design 62

 3.1.3.3 Efficacy Measures 64

 3.1.3.4 Sample Size Considerations 64

 3.1.3.5 Interim Analysis 64

 3.1.3.6 Statistical Analysis Methods 65

 3.1.3.7 Applicant’s Results and Statistical Reviewer’s Findings/ Comments..... 66

 3.1.4 Study H3S-MC-GGJY (CORE) 80

 3.1.4.1 Objective 80

 3.1.4.2 Study Design 80

 3.1.4.3 Efficacy Measures 83

3.1.4.4 Sample Size Considerations83
3.1.4.5 Interim Analysis83
3.1.4.6 Statistical Analysis Methods84
3.1.4.7 Applicant’s Results and Statistical Reviewer’s Findings/ Comments.....84
3.2 Evaluation of Safety.....93
4. Findings in Special/Subgroup Populations.....93
4.1 Gender, Race, and Age93
4.2 Other Special/Subgroup Populations94
5. Summary and Conclusions.....94
5.1 Statistical Issues and Collective Evidence.....94
5.2 Conclusions and Recommendations.....101

Appears This Way
On Original

Statistical Review and Evaluation

1. Executive Summary

1.1 Conclusions and Recommendations

The applicant submitted the analyses and results of four trials, STAR, RUTH, MORE and CORE, to seek registration of raloxifene for two indications: “reduction in the risk of invasive breast cancer in postmenopausal women at high risk for breast cancer”, and “reduction in the risk of invasive breast cancer in postmenopausal women with osteoporosis”. Raloxifene is currently approved for the treatment of osteoporosis and prevention of osteoporosis.

The data and analyses from STAR trial, which supports the indication “reduction in the risk of invasive breast cancer in postmenopausal women at high risk for breast cancer”, failed to demonstrate the superiority of raloxifene over tamoxifen. The applicant performed a non-inferiority analysis, which was not planned in the original design, to compare raloxifene to tamoxifen. The result of the non-inferiority analysis indicated that raloxifene may lose up to 35% tamoxifen effect, but there were many problems involved in this non-inferiority analysis. Raloxifene had more events in several safety categories while had fewer events in other safety categories compared to tamoxifen.

The data and analyses from RUTH, MORE and CORE trials, which support the indication “reduction in the risk of invasive breast cancer in postmenopausal women with osteoporosis”, showed that there were fewer invasive breast cancer events in raloxifene-treated subjects than that in placebo-treated subjects. However, one should note that in RUTH trial the incidence of the invasive breast cancer was amended as a co-primary endpoint, in MORE trial the incidence of the invasive breast cancer was a secondary safety endpoint, and in CORE trial subjects were not randomized between two treatment arms. In addition, raloxifene-treated subjects had more exposure to thromboembolic adverse events than those placebo-treated subjects numerically.

This supplemental application was discussed at the Oncology Drugs Advisory Committee (ODAC) on July 24, 2007. The committee recommends approval for both indications.

The final regulatory action should be based on clinical judgment and acceptability of risk-benefit profile.

1.2 Brief Overview of Clinical Studies

The analyses and results of four trials, STAR, RUTH, MORE and CORE, were submitted in this supplemental application to support two indications: “reduction in the risk of invasive breast cancer in postmenopausal women at high risk for breast cancer”, and “reduction in the risk of invasive breast cancer in postmenopausal women with osteoporosis”.

STAR was a randomized, double-blind, active-controlled study to evaluate the effect of raloxifene versus tamoxifen in reducing the incidence of invasive breast cancer in postmenopausal women at increased risk of invasive breast cancer, and was conducted in North America. To be eligible for participation in STAR, a woman had to be at least 35 years of age, postmenopausal, and have either a Gail model-based 5-year predicted risk of invasive breast cancer of at least 1.66% or a history of LCIS treated by local excision alone. The study excluded patients with a history of invasive breast cancer, ductal carcinoma in situ (DCIS), deep vein thrombosis (DVT), pulmonary embolus (PE), stroke, transient ischemic attack (TIA), current use of coumadin, uncontrolled diabetes or hypertension, or atrial fibrillation. STAR randomized 19,747 postmenopausal women to receive either tamoxifen 20 mg/day (N=9872) or raloxifene HCl 60 mg/day (N=9875) for a maximum of 5 years of treatment. Randomization was stratified by age, 5-year predicted invasive breast cancer risk based on the Gail model, race, history of LCIS, and hysterectomy status. Analyses were performed using all randomized patients who had at least one post-baseline visit (primary analysis dataset [N=19,487]). Per protocol, the final intention-to-treat analysis was performed once a pre-specified number of breast cancers (at least 327) were observed. As of December 31, 2005, approximately 25% of the study population had completed 5 years of treatment. For the remainder of the patients, the planned 5-year treatment period is ongoing. The trial was designed as a superiority trial, and the sample size computations were based on demonstrating superiority. The primary objective of STAR was to determine if:

- 1) compared to tamoxifen, raloxifene significantly reduces the incidence rate of invasive breast cancer;
- 2) compared to raloxifene, tamoxifen significantly reduces the incidence rate of invasive breast cancer; or
- 3) the statistical superiority of one of the treatments cannot be demonstrated and the choice of therapy should be based on benefit/risk considerations.

STAR analyses presented in this document are based on follow-up data reported as occurring on or before December 31, 2005.

RUTH was a double-blind, randomized, placebo-controlled, multinational study examining the long-term effect of raloxifene HCl 60 mg/day versus placebo on the incidences of two primary endpoints: (1) a combined coronary primary endpoint (defined as coronary death, nonfatal [including silent] myocardial infarction [MI], or hospitalized acute coronary syndrome [ACS] other than MI) and (2) invasive breast cancer in postmenopausal women at risk for major coronary events. To adjust for multiplicity of these two co-primary endpoints, the coronary primary endpoint was tested at significance level of 0.0423 and the breast cancer endpoint at significance level of 0.0080. Incidence of all breast cancer was and continued to be a secondary endpoint in RUTH. To be eligible for participation in RUTH, women had to have been age 55 years or older, at least 1 year postmenopausal, and have established CHD or multiple CHD risk factors. Participants had to have a CV risk score ≥ 4 according to a point system that took into account established CHD (4 points), lower extremity arterial disease (4 points), age 70 years or older (2 points), cigarette smoking (1 point), hypertension (1 point), or hyperlipidemia (1 point). Women with a suspected breast

carcinoma or with a known history of breast carcinoma were not eligible to enroll. The 10,101 women enrolled in the study were randomized to treatment with placebo (N=5057) or raloxifene HCl 60 mg/day (N=5044). The active treatment phase ended after the last randomized patient had been followed for at least 5 years.

MORE was a double-blind, randomized, placebo-controlled multinational study that examined the use of raloxifene in postmenopausal women with osteoporosis. The study consisted of a 3-year treatment phase and a 1-year extension phase. The MORE data presented in this document were for the entire 4-year study period. The 7705 women enrolled in the study were randomized to treatment with placebo (N=2576), raloxifene HCl 60 mg/day (N=2557), or raloxifene HCl 120 mg/day (N=2572). To be eligible for participation in MORE, a woman had to have been age 80 years or younger, at least 2 years postmenopausal, and diagnosed with osteoporosis, defined as lumbar spine or femoral BMD more than 2.5 standard deviations (SD) below the mean for normal premenopausal women or at least one moderate or two mild vertebral fractures. Women with a history of breast cancer were not eligible to enroll. The primary objectives of MORE were to assess the effect of raloxifene treatment, compared with placebo, on the incidence of new vertebral fractures, lumbar spine and femoral neck BMD, and safety. A secondary safety objective was to assess the effect of raloxifene on the incidence of breast cancer, regardless of invasiveness status.

CORE was a double-blind, placebo-controlled, multinational study designed to collect long-term breast cancer and nonvertebral fracture data from the MORE cohorts. CORE enrolled women who had been randomized to treatment in MORE and who chose to enroll in CORE, i.e., subjects were not re-randomized. As per the CORE protocol, CORE enrollees received the same therapy they had received in MORE. Specifically, patients randomized to raloxifene HCl 60 mg/day (N=1355) or 120 mg/day (N=1370) in MORE were assigned to receive raloxifene HCl 60 mg/day in CORE (N=2725); those randomized to placebo in MORE were assigned to receive placebo in CORE (N=1286). Consequently, approximately twice as many women in CORE were assigned to receive raloxifene HCl 60 mg/day compared with placebo. For all CORE patients, a treatment gap occurred between the end of their participation in MORE and the start of their participation in CORE (the median time off therapy was approximately 10.6 months). During this gap, patients did not receive study drug but could have taken marketed raloxifene, tamoxifen, other SERMS, or a hormone. CORE was a follow-up study of MORE participants but with different primary and secondary objectives. The primary objective of CORE was to compare the long-term effect of raloxifene HCl 60 mg/day versus placebo on the reduction in incidence of invasive breast cancer in postmenopausal women with osteoporosis. A secondary objective was to assess the long-term effect of raloxifene on the incidence of invasive, ER-positive breast cancer.

1.3 Statistical Issues and Findings

The analyses and results of four trials, STAR, RUTH, MORE and CORE, were submitted in this supplemental application to support two indications.

Table 1.3.1 presents the analyses for the incidence of invasive breast cancer for four trials. The statistical test was a stratified log-rank test for STAR, and a log-rank test for RUTH, MORE and CORE.

Table 1.3.1 Analyses for the Incidence of Invasive Breast Cancer

Trial	Treatment	Number Of Subjects	Invasive Breast Cancer Event	Incidence Rate per 1000 patient-years	Relative Risk (95% CI)	P-value
STAR	Raloxifene	9,751	173	4.40	1.02 (0.82, 1.27)	0.9868
	Tamoxifen	9,736	168	4.30		
RUTH	Raloxifene	5,044	40	1.50	0.56 (0.37, 0.84)	0.0032
	Placebo	5,057	70	2.66		
MORE	Raloxifene	2,557	11	1.26	0.29 (0.13, 0.58)	<0.0001
	Placebo	2,576	38	4.36		
CORE	Raloxifene	2,716	19	2.43	0.45 (0.23, 0.89)	0.0092
	Placebo	1,274	20	5.41		

For STAR trial, the stratified log-rank test for the primary analysis was not statistically significant. A post-hoc non-inferiority analysis, which used the sub-population from the NSABP-P1 trial, was performed to estimate tamoxifen effect size. The NSABP-P1 trial was conducted to compare the incidence of invasive breast cancer between tamoxifen and placebo. The data of women 50 years of age or older from the NSABP-P1 trial showed that tamoxifen decreased the incidence of invasive breast cancer by 53%. If one assumes that tamoxifen would have the same effect in STAR trial if a placebo arm would be included in STAR trial, the non-inferiority analysis indicated that raloxifene may lose up to 35% of tamoxifen effect.

Since STAR trial failed to demonstrate superiority of raloxifene over tamoxifen which was the primary goal of the trial, any additional analyses used to support any claims violate the statistical principle. One may argue that the STAR trial was a large trial which should not be ignored and it will be almost impossible to repeat such a trial. However, a large trial, in many aspects, is not necessarily a better trial. The reason for a large trial in this case is that the effect is too small and events are rare. When the effect size is small and sample size is large, it is often difficult to control many confounding factors.

In a non-inferiority analysis setting, a percent retention needs to be pre-specified, the control effect used needs to be well estimated which is often derived from several similar historical trials, variability among trials and within trials needs to be considered, and the constancy assumption which assumes that tamoxifen would have the same effect over placebo between NSABP-P1 trial and STAR trial if a placebo arm would be included in STAR trial needs to be satisfied. In the current non-inferiority analysis, the percent retention was not pre-specified, the control effect was derived from a subpopulation of one trial, variability within the trial was not discussed, and the validity of the constancy assumption is very difficult or impossible to verify.

For RUTH trial, the co-primary endpoint of the incidence of invasive breast cancer was amended after initiation of the trial. The log-rank test for the primary analysis was statistically significant.

For MORE trial, the incidence of invasive breast cancer was a secondary safety endpoint. The trial was not designed to demonstrate the raloxifene effect on the incidence of invasive breast cancer. Although p-value of the log-rank test on the invasive breast cancer endpoint was less than 0.05, this statistical test was not planned, especially, not adjusted for multiple secondary and safety endpoints. One should realize that for a trial having many secondary and safety endpoints, one can always find some "significant" endpoints. Such results may really provide true discoveries, but one should not rely on this type of results because the significance level of the tests associated with this type of results is completely uncontrolled.

For CORE trial, the results are difficult to interpret since subjects were from a subgroup of patients from the MORE trial and were not re-randomized in the CORE trial. Although baseline characteristics between two treatment arms appeared to be balanced, however, the trial results are more appropriate to be used as exploratory due to lack of re-randomization. The key of the randomization is to control some "known" factors, and more importantly, to control many "unknown" factors. Although the CORE trial showed that there were less incidence of the invasive breast cancer in raloxifene arm than that in the placebo arm, one should be cautious in interpreting the trial results because it was not a randomized trial.

Table 1.3.2 presents the efficacy and important safety outcomes for STAR trial.

Appears This Way
On Original

Table 1.3.2 STAR: Efficacy and Important Safety Outcomes

Events	# Events (%)		IR ^a		RR (95% CI) ^b
	Tamoxifen N=9736	Raloxifene N=9751	Tamoxifen	Raloxifene	
All breast cancers	228 (2.3)	256 (2.6)	5.85	6.54	1.12(0.93,1.34)
Invasive	168 (1.7)	173 (1.8)	4.30	4.40	1.02(0.82,1.27)
Non-invasive	60 (0.6)	83 (0.9)	1.54	2.12	1.38(0.98,1.95)
Clinical vertebral fracture	58	58	1.47	1.46	0.99(0.68,1.46)
Death	109	104	2.76	2.62	0.95(0.72,1.25)
Death due to stroke	7	5	0.18	0.13	0.71(0.18,2.60)
Stroke	56	54	1.42	1.36	0.96(0.65,1.42)
Deep Vein Thrombosis	92	67	2.35	1.69	0.72(0.52,1.00)
Pulmonary Embolism	58	38	1.47	0.96	0.65(0.42,1.00)
Endometrial Cancer ^c	37/4739	23/4715	1.99	1.21	0.61(0.34,1.05)
Ovarian Cancer	14	18	0.52	0.66	1.27(0.60,2.76)
Cataracts	435	343	13.19	10.34	0.78(0.68,0.91)
Hysterectomy	246/4739	92/4715	13.25	4.84	0.37(0.28,0.47)
Hot Flashes	7170	6748	181.71	169.91	0.94(0.90,0.97)
Leg Cramps	5999	5373	152.03	135.29	0.89(0.86,0.92)
Edema ^d	664	741	16.83	18.66	1.11(1.00,1.23)
Cholelithiasis ^e	NA	NA	NA	NA	NA

^aIR=incidence rate per 1000 patient-years

^bRelative risk for raloxifene compared to tamoxifen.

Relative Risk >1 indicates higher incidence for raloxifene compared to tamoxifen

Relative Risk < 1 indicates lower incidence for raloxifene compared to tamoxifen

^c Only patients with a uterus at baseline (tamoxifen n = 4739; raloxifene n = 4715)

^c Hysterectomy was calculated as a risk ratio.

^d Peripheral edema is not a coding term in CTC v2.0.

^e Cholelithiasis is not a coding term in CTC v2.0.

Table 1.3.3 presents the efficacy and important safety outcomes for RUTH trial.

Appears This Way
On Original

Table 1.3.3 RUTH: Efficacy and Important Safety Outcomes

Events	Raloxifene 5,044	Placebo 5,057	Raloxifene IR	Placebo IR	Absolute Risk Difference	Relative Risk (95% CI)
Invasive breast cancer	40	70	1.50	2.66	-1.16	0.56 (0.37, 0.84)
Noninvasive breast cancer	11	5	0.41	0.19	+0.22	2.18 (0.70, 7.99)
Invasiveness unknown	1	1	0.04	0.04	+0.00	NA
All breast cancers	52	76	1.95	2.89	-1.04	0.67 (0.46, 0.97)
Clinical vertebral fracture	64	97	2.40	3.70	-1.30	0.65 (0.47, 0.90)
Death	554	595	20.68	22.45	-1.77	0.92 (0.82, 1.04)
Death due to Stroke	59	39	2.20	1.47	+0.73	1.50 (0.98, 2.30)
Stroke	249	224	9.46	8.60	+0.86	1.10 (0.91, 1.32)
Deep vein thrombosis	65	47	2.44	1.78	+0.66	1.37 (0.94, 1.99)
Pulmonary embolism	36	24	1.35	0.91	+0.44	1.49 (0.89, 2.49)
Endometrial cancer ^a	21/3900	17/3882	1.01	0.83	+0.18	1.22 (0.61, 2.46)
Ovarian Cancer ^b	17/4559	10/4606	0.70	0.41	+0.29	1.71 (0.74, 4.17)
Hysterectomy ^a	58/3900	53/3882	2.79	2.60	+0.19	1.07 (0.73, 1.59)
Hot Flashes	397	241	14.82	9.09	+5.73	1.63 (1.39, 1.92)
Leg Cramps	483	334	18.03	12.60	+5.43	1.43 (1.24, 1.65)
Peripheral edema	706	583	26.36	22.00	+4.36	1.20 (1.07, 1.34)
Cholelithiasis ^c	168/4144	131/4111	7.83	6.20	+1.63	1.26 (1.00, 1.60)

Abbreviations: IR = Incidence Rate per 1000 Patient-years.

^a Only patients with an intact uterus were considered for the denominator (raloxifene denominator = 3900, placebo denominator = 3882).

^b Only patients with at least one ovary were considered for the denominator (raloxifene denominator = 4559, placebo denominator = 4606).

^c Only patients with an intact gallbladder at baseline (raloxifene n=4144, total person-years of follow-up=21467; placebo n=4111, total person-years of follow-up=21136).

Table 1.3.4 presents the efficacy and important safety outcomes for MORE trial.

Table 1.3.4 MORE: Efficacy and Important Safety Outcomes

Events ^a	Raloxifene 2,557	Placebo 2,576	Raloxifene IR	Placebo IR	Absolute Risk Difference	RR (95% CI)
Invasive breast cancer	11	38	1.26	4.36	-3.10	0.29 (0.13, 0.58)
Noninvasive breast cancer	3	5	0.34	0.57	-0.23	0.60 (0.09, 3.07)
Invasiveness unknown	3	1	0.34	0.11	+0.23	2.99 (0.24, 156)
All breast cancers	17	44	1.94	5.05	-3.11	0.38 (0.21, 0.69)
Clinical vertebral fracture	62	107	7.08	12.27	-5.19	0.58 (0.42, 0.80)
Death	64/5129	36	3.63	4.13	-0.50	0.88 (0.58, 1.36)
Death due to Stroke	9/5129	6	0.51	0.69	-0.18	0.74 (0.23, 2.52)
Stroke	91/5129	56	5.16	6.42	-1.26	0.80 (0.57, 1.14)
Deep vein thrombosis	44/5129	8	2.50	0.92	+1.58	2.72 (1.27, 6.68)
Pulmonary embolism	22/5129	4	1.25	0.46	+0.79	2.72 (0.92, 10.85)
Endometrial and uterine cancer ^b	8/3960	5/1999	0.59	0.74	+0.15	0.80 (0.23, 3.10)
Ovarian Cancer	6/5129	6/1999	0.34	0.69	-0.35	0.49 (0.13, 1.84)
Hysterectomy ^b	40/3960	22/1999	2.93	3.24	-0.31	0.90 (0.52, 1.60)
Hot Flashes	512/5129	151	29.04	17.31	+11.73	1.68 (1.40, 2.03)
Leg Cramps	443/5129	150	25.13	17.20	+7.93	1.46 (1.21, 1.77)
Peripheral edema	340/5129	134	19.29	15.36	+3.93	1.26 (1.03, 1.55)
Cholelithiasis ^c	93/5129	45	5.28	5.16	+0.12	1.02 (0.71, 1.50)

Abbreviations: IR = Incidence Rate per 1000 Patient-years; RR=Relative risk.

^a Breast cancer and clinical vertebral fracture events are for the raloxifene 60 mg/day arm only; denominator = 2557. For the safety events of death, death due to stroke, stroke, deep vein thrombosis, pulmonary embolism, and ovarian cancer, the raloxifene 60 and 120 mg/day arms were pooled to have the greatest opportunity to detect safety signals; thus, the denominator for these events is 5129.

^bOnly patients with a uterus at baseline (pooled raloxifene n=3960, total person-years of follow-up=13659.16; placebo n=1999, total person-years of follow-up=6791.41). "Hysterectomy" included MedDRA Preferred Terms of "Hysterectomy," "Hysterosalpingo-oophorectomy," and "radical hysterectomy."

^cGallbladder status at baseline was not ascertained in the MORE trial.

Table 1.3.5 presents the efficacy and important safety outcomes for CORE trial.

Table 1.3.5 CORE: Efficacy and Important Safety Outcomes

Events ^a	RLX 2,716	PLB 1,274	RLX IR	PLB IR	Absolute Risk Difference	Relative Risk (95% CI)
Invasive breast cancer	19	20	2.43	5.41	-2.98	0.45 (0.23, 0.89)
Noninvasive breast cancer	5	2	0.64	0.54	+0.10	1.18 (0.19, 12.44)
Invasiveness unknown	0	0	0.00	0.00	0.00	NA
All breast cancers	24	22	3.07	5.95	-2.88	0.52 (0.28, 0.96)
Clinical vertebral fracture ^b	65/2725	32/1286	8.28	8.56	-0.28	0.97 (0.62, 1.53)
Death	47/2725	29/1286	5.99	7.76	-1.77	0.77 (0.48, 1.27)
Death due to Stroke	6/2725	1/1286	0.76	0.27	+0.49	2.81 (0.34, 129)
Stroke	49/2725	14/1286	6.24	3.75	+2.49	1.65 (0.92, 2.98)
Deep vein thrombosis	17/2725	4/1286	2.17	1.07	+1.10	2.01 (0.68, 5.95)
Pulmonary embolism	9/2725	0/1286	1.15	0.00	+1.15	NA
Endometrial and uterine cancer ^c	4/2138	3/1008	0.65	1.02	-0.37	0.64 (0.11, 4.35)
Ovarian Cancer	2/2725	2/1286	0.25	0.54	-0.29	0.46 (0.03, 6.39)
Hysterectomy ^c	13/2138	10/1008	2.11	3.40	-1.29	0.62 (0.25, 1.58)
Hot Flashes	26/2725	11/1286	3.31	2.94	+0.37	1.13 (0.54, 2.52)
Leg Cramps	90/2725	36/1286	11.46	9.63	+1.83	1.19 (0.80, 1.80)
Peripheral edema	61/2725	30/1286	7.77	8.03	-0.26	0.97 (0.62, 1.55)
Cholelithiasis ^d	35/2725	12/1286	4.46	3.21	+1.25	1.39 (0.70, 2.94)

Abbreviations: IR = Incidence Rate per 1000 Patient-years; PLB = Placebo; RLX = Raloxifene.

^a Breast cancer events were for the patients who enrolled in CORE and had not been diagnosed with breast cancer prior to Visit 1.

^b Vertebral fractures were collected as adverse events.

^c Only patients with an intact uterus were considered for denominator (raloxifene denominator = 2138, placebo denominator = 1008).

^d Gallbladder status at baseline was not ascertained in the CORE trial.

Summary

The STAR trial failed to demonstrate superiority of raloxifene over tamoxifen, and had many problems in the post-hoc non-inferiority analysis. The RUTH, MORE and CORE trials showed that

there were fewer invasive breast cancer events in raloxifene-treated subjects than that placebo-treated subjects. However, one should be cautious in interpreting of the results of MORE and CORE trials. In addition, raloxifene-treated subjects had more exposure to thromboembolic adverse events than those placebo-treated subjects numerically.

This supplemental application was discussed at the Oncology Drugs Advisory Committee (ODAC) on July 24, 2007. The committee recommends approval for both indications.

The final regulatory action should be based on clinical judgment and acceptability of risk-benefit profile.

Appears This Way
On Original

2. Introduction

2.1 Overview

Evista® (raloxifene hydrochloride [HCl], hereafter referred to as raloxifene) was approved by Food and Drug Administration (FDA) on 09 December 1997 and 30 September 1999 for the prevention and treatment of osteoporosis in postmenopausal women, respectively. Since first approval through 30 November 2006, an estimated 22 million patients in 88 countries worldwide have received raloxifene, representing approximately 12 million patient-years of treatment.

The focus of this supplemental New Drug Application was on the use of raloxifene for the reduction in risk of invasive breast cancer in postmenopausal women. The applicant is seeking registration for the following indication statements for raloxifene:

The reduction in risk of invasive breast cancer in postmenopausal women at high risk for breast cancer.

The reduction in risk of invasive breast cancer in postmenopausal women with osteoporosis.

The submission contained the analyses and results of four trials, STAR, RUTH, MORE and CORE.

STAR was a randomized, double-blind, active-controlled study to evaluate the effect of raloxifene versus tamoxifen in reducing the incidence of invasive breast cancer in postmenopausal women at increased risk of invasive breast cancer, and was conducted in North America. To be eligible for participation in STAR, a woman had to be at least 35 years of age, postmenopausal, and have either a Gail model-based 5-year predicted risk of invasive breast cancer of at least 1.66% or a history of LCIS treated by local excision alone. The study excluded patients with a history of invasive breast cancer, ductal carcinoma in situ (DCIS), deep vein thrombosis (DVT), pulmonary embolus (PE), stroke, transient ischemic attack (TIA), current use of coumadin, uncontrolled diabetes or hypertension, or atrial fibrillation. STAR randomized 19,747 postmenopausal women to receive either tamoxifen 20 mg/day (N=9872) or raloxifene HCl 60 mg/day (N=9875) for a maximum of 5 years of treatment. Randomization was stratified by age, 5-year predicted invasive breast cancer risk based on the Gail model, race, history of LCIS, and hysterectomy status. Analyses were performed using all randomized patients who had at least one post-baseline visit (primary analysis dataset [N=19,487]). Per protocol, the final intention-to-treat analysis was performed once a pre-specified number of breast cancers (at least 327) were observed. As of December 31, 2005, approximately 25% of the study population had completed 5 years of treatment. For the remainder of the patients, the planned 5-year treatment period is ongoing. The trial was designed as a superiority trial, and the sample size computations were based on demonstrating superiority. The primary objective of STAR was to determine if:

- 1) compared to tamoxifen, raloxifene significantly reduces the incidence rate of invasive breast cancer;
- 2) compared to raloxifene, tamoxifen significantly reduces the incidence rate of invasive breast cancer; or
- 3) the statistical superiority of one of the treatments cannot be demonstrated and the choice of therapy should be based on benefit/risk considerations.

STAR analyses presented in this document are based on follow-up data reported as occurring on or before December 31, 2005.

RUTH was a double-blind, randomized, placebo-controlled, multinational study examining the long-term effect of raloxifene HCl 60 mg/day versus placebo on the incidences of two primary endpoints: (1) a combined coronary primary endpoint (defined as coronary death, nonfatal [including silent] myocardial infarction [MI], or hospitalized acute coronary syndrome [ACS] other than MI) and (2) invasive breast cancer in postmenopausal women at risk for major coronary events. To adjust for multiplicity of these two co-primary endpoints, the coronary primary endpoint was tested at significance level of 0.0423 and the breast cancer endpoint at significance level of 0.0080. Incidence of all breast cancer was and continued to be a secondary endpoint in RUTH. To be eligible for participation in RUTH, women had to have been age 55 years or older, at least 1 year postmenopausal, and have established CHD or multiple CHD risk factors. Participants had to have a CV risk score ≥ 4 according to a point system that took into account established CHD (4 points), lower extremity arterial disease (4 points), age 70 years or older (2 points), cigarette smoking (1 point), hypertension (1 point), or hyperlipidemia (1 point). Women with a suspected breast carcinoma or with a known history of breast carcinoma were not eligible to enroll. The 10,101 women enrolled in the study were randomized to treatment with placebo (N=5057) or raloxifene HCl 60 mg/day (N=5044). The active treatment phase ended after the last randomized patient had been followed for at least 5 years.

MORE was a double-blind, randomized, placebo-controlled multinational study that examined the use of raloxifene in postmenopausal women with osteoporosis. The study consisted of a 3-year treatment phase and a 1-year extension phase. The MORE data presented in this document were for the entire 4-year study period. The 7705 women enrolled in the study were randomized to treatment with placebo (N=2576), raloxifene HCl 60 mg/day (N=2557), or raloxifene HCl 120 mg/day (N=2572). To be eligible for participation in MORE, a woman had to have been age 80 years or younger, at least 2 years postmenopausal, and diagnosed with osteoporosis, defined as lumbar spine or femoral BMD more than 2.5 standard deviations (SD) below the mean for normal premenopausal women or at least one moderate or two mild vertebral fractures. Women with a history of breast cancer were not eligible to enroll. The primary objectives of MORE were to assess the effect of raloxifene treatment, compared with placebo, on the incidence of new vertebral fractures, lumbar spine and femoral neck BMD, and safety. A secondary safety objective was to assess the effect of raloxifene on the incidence of breast cancer, regardless of invasiveness status.

CORE was a double-blind, placebo-controlled, multinational study designed to collect long-term breast cancer and nonvertebral fracture data from the MORE cohorts. CORE enrolled women who had been randomized to treatment in MORE and who chose to enroll in CORE, i.e., subjects were not re-randomized. As per the CORE protocol, CORE enrollees received the same therapy they had received in MORE. Specifically, patients randomized to raloxifene HCl 60 mg/day (N=1355) or 120 mg/day (N=1370) in MORE were assigned to receive raloxifene HCl 60 mg/day in CORE (N=2725); those randomized to placebo in MORE were assigned to receive placebo in CORE (N=1286). Consequently, approximately twice as many women in CORE were assigned to receive raloxifene HCl 60 mg/day compared with placebo. For all CORE patients, a treatment gap occurred between the end of their participation in MORE and the start of their participation in CORE (the median time off therapy was approximately 10.6 months). During this gap, patients did not receive study drug but could have taken marketed raloxifene, tamoxifen, other SERMS, or a hormone. CORE was a follow-up study of MORE participants but with different primary and secondary objectives. The primary objective of CORE was to compare the long-term effect of raloxifene HCl 60 mg/day versus placebo on the reduction in incidence of invasive breast cancer in postmenopausal women with osteoporosis. A secondary objective was to assess the long-term effect of raloxifene on the incidence of invasive, ER-positive breast cancer.

The application was discussed at the Oncology Drugs Advisory Committee Meeting on July 24, 2007.

2.2 Data Sources

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

\\Cdsesub1\n22042\N_000\2006-11-13

3. Statistical Evaluation

3.1 Evaluation of Efficacy

Part of the text, tables and figures presented in this section were adapted from the applicant's Study Report.

3.1.1 Study H3S-MC-GGIY (STAR)

3.1.1.1 Objective

The primary objective of this study was to determine which of the following three statements was true:

- compared to tamoxifen, raloxifene significantly reduces the incidence rate of invasive breast cancer;

- compared to raloxifene, tamoxifen significantly reduces the incidence rate of invasive breast cancer; or
- the statistical superiority of one of the treatments cannot be demonstrated and the choice of therapy should be based on benefit/risk considerations.

3.1.1.2 Study Design

STAR was a multi-center, double-blind, double-dummy, active-controlled, randomized study enrolling 19,747 patients to one of two therapy groups: tamoxifen citrate 20 mg/day or raloxifene HCL 60 mg/day. Randomization was stratified by age (35-49, 50-59, >59), invasive breast cancer risk within 5 years from the Gail et al model (<2.0, 2.0-2.9, 3.0-4.9, ≥ 5.0), race (black, white, other), history of LCIS (yes, no), and prior hysterectomy status (yes, no). Figure 3.1.1.2.1 presents the study design.

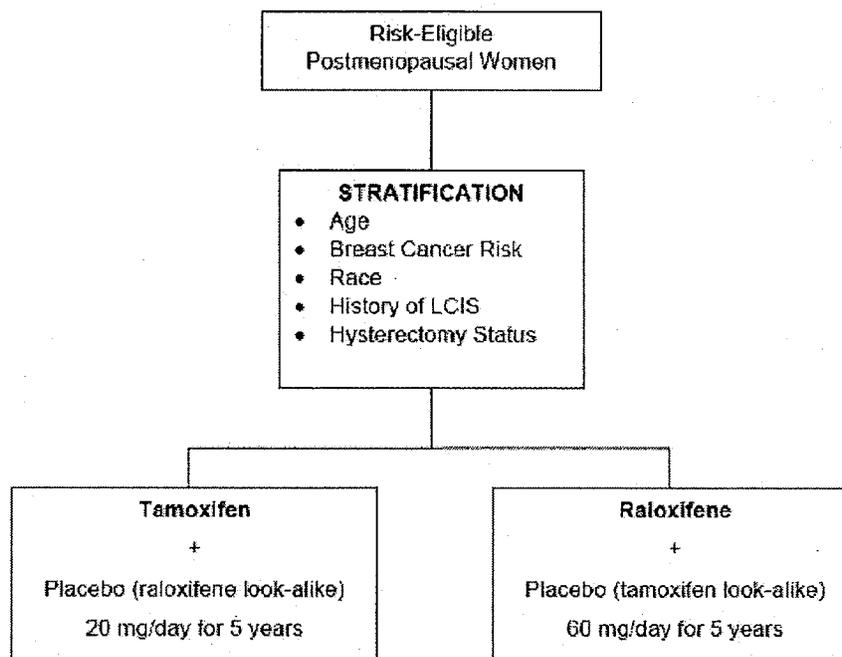
During the recruitment phase, potential patients consented to be evaluated for breast cancer risk and the breast cancer risk assessment profile was completed. At a subsequent visit, the risk assessment profile results were discussed with the potential patient, the protocol and protocol requirements were presented, and informed consent was obtained.

Prior to randomization, a medical history was obtained, including a specific assessment of breast cancer, cardiovascular (CV), and osteoporosis risk factors; detailed family history of breast and CV disease; demographic information; and existing symptoms. Within 180 days prior to randomization, a general physical examination, including a clinical breast examination, and analyses for complete blood count (CBC), differential, platelet count, alkaline phosphatase, alanine transaminase (ALT [SGPT]), aspartate transaminase (AST [SGOT]), total bilirubin, and serum creatinine were to have been performed. A gynecologic examination, including a bimanual pelvic examination and, if indicated, a pap smear, was to have been performed prior to entry (i.e., prior to randomization). If the patient had been asymptomatic and had a normal gynecologic examination within the past 365 days, a repeat examination prior to entry was not required. Patients with a prior hysterectomy and bilateral salpingoophorectomy were exempt from this requirement. Women with a prior hysterectomy who still had their ovaries were required to have the gynecologic examination. A preentry (i.e., prior to randomization) bilateral mammogram was performed within 365 days prior to randomization. Prior to entry, completion of a symptom checklist was required for all patients and completion of the QOL questionnaires was required for all patients at centers selected to participate in the QOL substudy. Prior to starting therapy, a blood sample for serum banking was collected from those patients who consented to have the sample collected and stored for future research.

After randomization, follow-up visits were scheduled every 6 months for 5 years and then annually thereafter. Patients were to have been contacted, either in person or by phone, at 3 months after randomization to monitor and promote compliance. A clinical breast examination was performed at each 6-month follow-up visit for 5 years and annually thereafter. A follow-up bilateral mammogram

was performed annually based on the pre-entry mammogram date. A follow-up gynecologic examination was performed annually for those patients who had not had a prior hysterectomy and bilateral salpingo-oophorectomy. Hematology and blood chemistry analyses, with the exception of the differential, were performed annually for 5 years.

Figure 3.1.1.2.1 Study Design



Abbreviation: LCIS = lobular carcinoma in situ.

3.1.1.3 Efficacy Measures

The primary endpoint was the time to the occurrence of invasive breast cancer. Secondary endpoints included the occurrence of non-invasive breast cancer and safety.

To define the incidence of invasive breast cancer, a clinical breast examination was performed at each 6-month follow up visit. A bilateral mammogram was required annually based on the pre-entry bilateral mammogram date. The results of all breast biopsies and cytologies including those diagnosed as benign and atypical hyperplasia were to have been reported. When the report was either positive or suspicious, all mammogram reports, operative reports, and pathology reports/materials must have been submitted to NSABP Biostatistical Center for medical review. A pathologic

diagnosis of invasive breast cancer, as indicated by the pathology report from the clinical center pathologist, was required. The medical reviewer confirmed the diagnosis of invasive breast cancer on medical review Form P2ER. Blocks of tumor tissue were to have been submitted to the NSABP Biostatistical Center for future analysis but not central pathology review.

To define in-situ (non-invasive) breast cancer, a pathologic diagnosis of non-invasive breast cancer (LCIS or DCIS) and/or atypical hyperplasia, as indicated by the pathology report from the clinical center pathologist, was required. Blocks of tissue were to have been submitted to the NSABP Biostatistical Center for future analysis but not central pathology review.

3.1.1.4 Sample Size Considerations

An original sample size requirement of approximately 22,000 women, with 327 incidence of invasive breast cancers at the final analysis, was calculated based upon several factors. These included: 1) the expected hazard rate for invasive breast cancer; 2) the dropout rate for the study; 3) the rate of patient non-compliance with protocol therapy; 4) the duration of treatment effect; and 5) the anticipated pattern of patient accrual (the number of years of accrual and the number of patients accrued by year).

Table 3.1.1.4.1 provides the probabilities, for selected scenarios of raloxifene breast cancer effect, of concluding: 1) that the superiority of one treatment for its effectiveness in reducing breast cancer incidence was sufficient to make it the preferred treatment for women eligible for this trial, or 2) that neither treatment had met this criteria and that other factors may result in each treatment being recommended for certain subsets of patients.

Table 3.1.1.4.1 Probability of Concluding Superiority for Reducing the Incidence of Breast Cancer for Several Possible Scenarios of Raloxifene Breast Cancer Effect

True (not observed) breast cancer Incidence rate per 1000 person-years				Probability of concluding superiority for tamoxifen, raloxifene, or neither treatment for reducing the incidence of breast cancer		
Raloxifene	Tamoxifen	Difference in rate	Ratio of rates	Tamoxifen	Neither	Raloxifene
5.50	3.21	2.29	1.71	.99	.01	.00
5.00	3.21	1.79	1.56	.95	.05	.00
3.21	3.21	0.00	1.00	.025	.95	.025
2.14	3.21	-1.07	0.67	.00	.15	.85
1.61	3.21	-1.61	0.50	.00	.003	.997

In computing the sample size for the trial, it was particularly important to assure that one would not conclude that the two treatments were equivalent if the overall increase in annual incidence rate

associated with raloxifene (versus tamoxifen) would negate half of the gain obtained from tamoxifen versus placebo. This would have occurred if the incidence rate of invasive breast cancer in those receiving raloxifene increased (relative to the rate in those receiving tamoxifen) by 56%. (The overall incident rate for placebo was 112% greater than that for tamoxifen for postmenopausal P-1 patients). With the proposed sample size of 22,000, there was 95% power to detect this alternative (see the second row of Table 3.1.1.4.1).

It was also of considerable interest if the incidence rate of breast cancer in those receiving raloxifene decreased by 1/3 (relative to the rate in those receiving tamoxifen). This would be a substantial enough reduction that one would not want to fail to identify the benefit, as such a benefit would be important to women with or without hysterectomy. The power to detect this alternative was 85% (see the fourth row of Table 3.1.1.4.1).

The third row of Table 3.1.1.4.1 represented the case that raloxifene and tamoxifen were equivalent for reducing breast cancer incidence. The probability of concluding that the two treatments were equivalent was 0.95. Although the expected incident rate of 3.21 per 1000 person-years was used for this example, the probability of concluding that the two treatments were equivalent remained at 0.95, provided that the incident rates for raloxifene and tamoxifen were equal.

The two other alternatives presented in Table 3.1.1.4.1 referred to extreme cases. The first row in the table represented a case in which the benefit associated with tamoxifen was such that any reduction in endometrial cancers associated with raloxifene (assuming raloxifene caused no increase risk of endometrial cancer relative to placebo) would have been completely offset by a comparable increase in invasive breast cancers. The last row in the table represented a case in which the benefit for raloxifene versus tamoxifen was essentially equivalent to the benefit associated with tamoxifen versus placebo. The power for both alternatives exceeded 99%.

The sample size and number of events required to have adequate statistical power, were based on numerous assumptions that represented "best-guess" estimates of the relevant design factors. The experience of the P-1 trial was the source for most of these estimates. However, the population of women attracted to the STAR trial might have been different from the P-1 population in average risk of breast cancer, therapy compliance, and study retention. Therefore, the sample size requirement was reassessed at the DMC meeting subsequent to the time when there had been 2 years of follow-up on the first 5,000 women randomized or at the time of the DMC meeting closest to the 3-year anniversary of the start of randomization, whichever came first. At that time, if it appeared there were errors in the original assumptions, the sample size goal may have been modified to assure that there would be 327 incidence of invasive breast cancers at the time of final analysis. The reassessment was based on the invasive breast cancer rate for patients receiving tamoxifen but did not utilize the invasive breast cancer rate for patients receiving raloxifene. This analysis was performed by the Biostatistical Center staff and submitted for review by an external DMC. Per the specification defined a priori in the protocol, the protocol statistician performed a reassessment of the projected sample size requirements for the study. The findings from the reassessment were presented

to the DMC for their review and comment at the October 04, 2002 meeting of the Committee. The findings were summarized as "As defined in the protocol, 327 cases of invasive breast cancer were needed to achieve the statistical power for the test of the primary hypothesis. As described in the protocol, it was originally estimated that after accruing 22,000 patients in 5 years, 327 events would be observed at 6.6 years from the time randomization was initiated. When performed in accordance to the scenario described in the protocol, the recalculations of the sample size requirements incorporating the observed values of the parameters used for sample size determination indicated that the needed 327 events would be observed at about the same time planned earlier (6.6 years) if 19,000 patients were randomized within the planned 5-year accrual period." In light of this information, the DMC submitted to the NSABP leadership a written recommendation that the target accrual for the trial was reduced from 22,000 to 19,000 patients. This recommendation was based on four main points of consideration. Compared to the original sample size and statistical analysis methods, the new sample size provided a trial: 1) with the same number of observed breast cancer events for final analysis; 2) with the same statistical power for the test of the primary hypothesis; 3) of the same duration as originally planned; and 4) that avoided the situation where the study was still in the accruing phase at or near the time when the final analysis was conducted (327 breast cancer events observed). The recommendation of the DMC and the considerations upon which the recommendation was made were reviewed by STAR Steering Committee at their meeting on November 01, 2002. The Steering Committee unanimously accepted the DMC recommendation and the target accrual were modified to 19,000 patients.

Reviewer's Comments

Based on the statistical inference principle, one can not conclude that two treatments are equivalent if the null hypothesis is not rejected.

3.1.1.5 Interim Analysis

The DMC for the STAR trial consisted of 7 individuals who were external to the NSABP. The DMC met approximately every 6 months to: a) review available safety and outcome data and; b) ensure that the event rates, enrollment rates, and dropout rates met protocol projections.

Six interim efficacy analyses of the breast cancer primary endpoint were planned and conducted when 47, 93, 140, 187, 234, and 280 invasive cancers have been observed. The interim analyses were presented at the DMC meeting on 26 April 2002, 11 April 2003, 12 September 2003, 22 October 2004, 5 May 2005 and 28 October 2005. The final analysis was presented to the DMC at their meeting on 14 April 2006.

The primary endpoint of the trial, incidence of invasive breast cancer, was the basis for the formal interim analyses. The difference between treatment groups in the incidence rate of invasive cancer was analyzed to assess if there was a higher-than-anticipated potential benefit from either treatment. This was accomplished by using the general stopping rule proposed by Fleming et al. using a two-

tailed log-rank test. The boundaries for the interim and final analyses were 0.00161, 0.00197, 0.00221, 0.00289, 0.00288, 0.00366 and 0.04628, respectively.

3.1.1.6 Statistical Analysis Methods

The primary analysis for the incidence of invasive breast cancer was a stratified log-rank test, stratified by the stratification factors used in the randomization: age (35-49, 50-59, >59), invasive breast cancer risk within 5 years from the Gail et al model (<2.0, 2.0-2.9, 3.0-4.9, ≥ 5.0), race (black, white, other), history of LCIS (yes, no), and prior hysterectomy status (yes, no). Secondary and other endpoints were compared using an un-stratified log-rank test. Cumulative incidence curves through 72 months of study follow-up were generated for the primary endpoint and other secondary endpoints.

3.1.1.7 Applicant's Results and Statistical Reviewer's Findings/ Comments

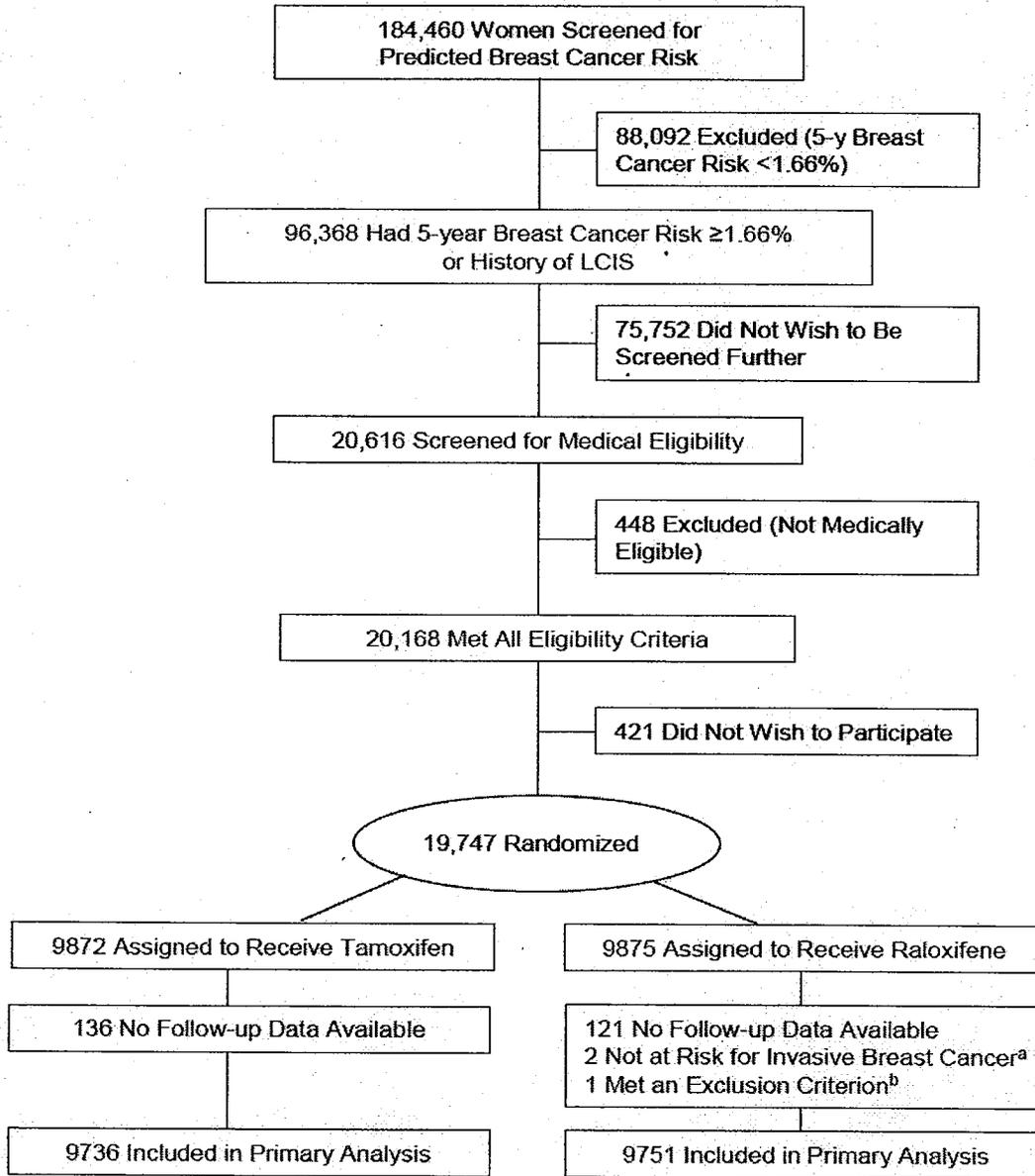
3.1.1.7.1 Study Population

A total of 184,460 women were screened. Of these, 96,368 had a predicted 5-year risk of at least 1.66% or a history of LCIS. From this group, 20,616 agreed to be screened to determine full eligibility for the trial based on the medical criteria defined below; 20,168 were found to meet all eligibility criteria of the study. Of this latter group, 19,747 women expressed a desire to go forward with participation in the trial, signed a consent form, and were randomized to receive either tamoxifen (N=9872) or raloxifene (N=9875). Of those randomized to tamoxifen, 136 patients had no follow-up data available after randomization. Of those randomized to raloxifene, 121 patients had no follow-up data available after randomization, 2 patients who had a history of bilateral mastectomy were not at risk for invasive breast cancer, and 1 patient who had a history of invasive breast cancer prior to randomization met an exclusion criterion. Thus, 9736 women randomized to tamoxifen and 9751 women randomized to raloxifene were included in the primary analysis dataset.

Table 3.1.7.1.1 presents the subject disposition.

Appears This Way
On Original

Table 3.1.7.1.1 Subject Disposition



^a History of bilateral mastectomy.

^b History of invasive breast cancer prior to randomization.

3.1.1.7.2 Demographic and Baseline Characteristics

Table 3.1.1.7.2.1 summarizes the treatment groups by the randomization strata and other patient characteristics at time of randomization.

Table 3.1.1.7.2.1 Patient Characteristics at Time of Randomization

Patient Characteristic	Tamoxifen (N=9736)		Raloxifene (N=9751)	
	n	%	n	%
Age (years)				
≤49	884	9.1	878	9.0
50-59	4856	49.9	4852	49.8
60-69	3136	32.2	3174	32.6
≥70	860	8.8	847	8.7
Race/ethnicity				
Caucasian	9105	93.5	9112	93.4
African American	233	2.4	243	2.5
Hispanic	192	2.0	193	2.0
Other	206	2.1	203	2.1
No. of first-degree relatives with breast cancer				
0	2838	29.1	2791	28.6
1	5046	51.8	5132	52.6
2	1532	15.7	1561	16.0
≥3	320	3.3	267	2.7
History of hysterectomy				
No	4739	48.7	4715	48.4
Yes	4997	51.3	5036	51.6
History of lobular carcinoma in situ				
No	8845	90.8	8859	90.9
Yes	891	9.2	892	9.1
History of breast atypical hyperplasia				
No	7546	77.5	7512	77.0
Yes	2190	22.5	2239	23.0
5-year predicted breast cancer risk (%)				
≤2.00	1055	10.8	1101	11.3
2.01-3.00	2993	30.7	2892	29.7
3.01-5.00	3042	31.2	3085	31.6
≥5.01	2646	27.2	2673	27.4
History of bilateral oophorectomy				
Yes	2923	30.0	2964	30.4
No	6813	70.0	6787	69.6
History of cataracts				
Yes	1394	14.3	1418	14.5
No	8342	85.7	8333	85.5

Abbreviation: N = patients comprising the primary analysis dataset; n = number of patients.

Table 3.1.1.7.2.2 summarizes the total patient-years of study follow up.

Table 3.1.1.7.2.2 Total Patient-Years of Study Follow Up

Patient-Years	Tamoxifen (N=9736)	Raloxifene (N=9751)	Total (N=19,487)
Mean	4.05	4.07	4.06
Standard deviation	1.62	1.62	1.62
Median	4.29	4.34	4.32
Minimum	0.08	0.07	0.07
Maximum	6.50	6.50	6.50

3.1.1.7.3 Applicant's Efficacy Analyses

The efficacy analyses used the primary analysis dataset and compared the two treatment groups by a stratified log-rank test based on cumulative incidence. All investigator-reported breast cancers were reviewed and confirmed by a physician blinded to individual patient treatment assignment. Results presented were based on confirmed events.

There were 168 cases of invasive breast cancer reported in patients assigned to tamoxifen and 173 cases reported in those assigned to raloxifene (Table 3.1.1.7.3.1). The rate per 1000 was 4.30 in the tamoxifen group and 4.40 in the raloxifene group (RR 1.02, 95% CI 0.82-1.27). For the primary analysis, the effects of treatment with tamoxifen and raloxifene on the incidence of invasive breast cancer were compared by a stratified log-rank test based on the randomization stratification factors, which gave a p-value 0.99. There was no difference between treatment groups on the incidence of invasive breast cancer. Figure 3.1.1.7.3.1 presents the Kaplan-Meier curve for the incidence of invasive breast cancer.

*Appears This Way
On Original*

Table 3.1.1.7.3.2 presents results of the time to event analyses of invasive breast cancer by ER status.

Table 3.1.1.7.3.2 Invasive Breast Cancer by ER Status (Primary Analysis Dataset)

Breast Cancer Category	Tamoxifen 20 mg/day N=9726		Raloxifene HCl 60 mg/day N=9745		RR (95% CI)
	n	IR	n	IR	
Invasive	163	4.30	168	4.41	1.02 (0.82, 1.28)
ER-positive	115	3.04	109	2.86	0.93 (0.72, 1.24)
ER-negative	44	1.16	51	1.34	1.15 (0.75, 1.77)
ER unknown	4	0.11	8	0.21	1.99 (0.53, 9.02)

Table 3.1.1.7.3.3 presents results of the time to event analyses of in situ (non-invasive) breast cancer by treatment group. There were 60 incident cases of noninvasive breast cancer among patients assigned to tamoxifen and 83 incident cases among patients assigned to raloxifene. The rate per 1000 women was 1.54 for the tamoxifen group and 2.12 for the raloxifene group (RR 1.38, 95% CI 0.98-1.95). Figure 3.1.1.7.3.2 presents the Kaplan-Meier curve for the incidence of non-invasive breast cancer.

Table 3.1.1.7.3.3 In Situ [Non-invasive] Breast Cancer by Treatment Group (Primary Analysis Dataset)

Noninvasive Breast Cancer	Number of Events		Average Annual Rate per 1000			RR (95% CI) ^b
	Tamoxifen	Raloxifene	Tamoxifen	Raloxifene	Difference ^a	
Overall	60	83	1.54	2.12	-0.58	1.38 (0.98-1.95)
DCIS	32	47	0.82	1.20	-0.38	1.46 (0.91-2.37)
LCIS	23	29	0.59	0.74	-0.15	1.26 (0.70-2.27)
Mixed	5	7	0.13	0.18	-0.05	1.39 (0.38-5.57)

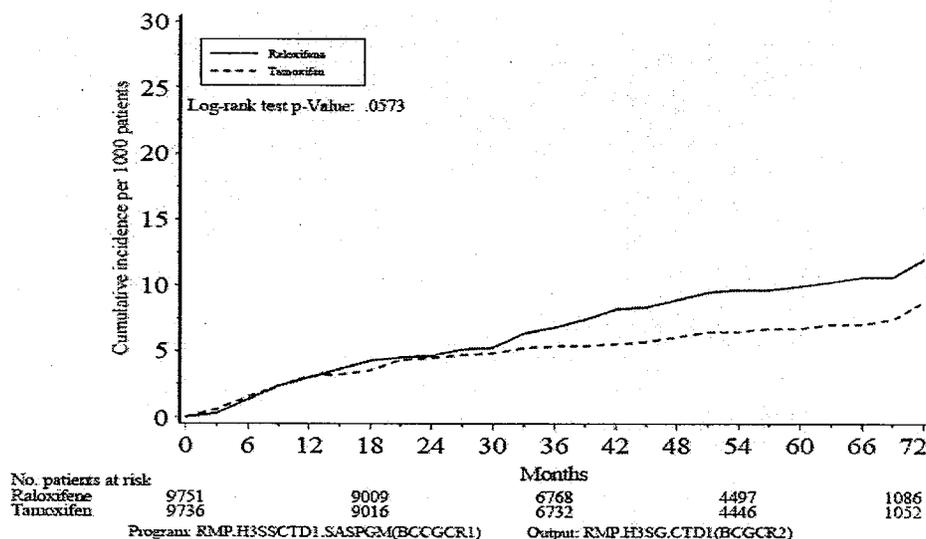
Abbreviations: CI = confidence interval; DCIS = intraductal carcinoma in situ; LCIS = lobular carcinoma in situ; RR = risk ratio.

^a Rate in the tamoxifen group minus rate in the raloxifene group.

^b Risk ratio for patients in the raloxifene group compared to patients in the tamoxifen group.

Appears This Way
On Original

Figure 3.1.1.7.3.2 Cumulative Incidence of In Situ (Non-invasive) Breast Cancer



For all breast cancer events, there were 228 incident cases among patients assigned to tamoxifen and 256 incident cases among patients assigned to raloxifene. The rate per 1000 women was 5.85 for the tamoxifen group and 6.54 for the raloxifene group (RR 1.12, 95% CI 0.93-1.34).

The applicant also performed a non-inferiority analysis, using Rothmann's method to evaluate the proportion of tamoxifen's effect maintained by raloxifene and to calculate a putative hazard ratio of raloxifene compared to placebo had a placebo been included in STAR. The following assumptions were made in applying Rothmann's method:

- 1) tamoxifen decreases the incidence of invasive breast cancer in women 50 years of age or older by 53% based on the P-1 trial
- 2) tamoxifen maintains the same treatment effect in women age of 50 years or older in P-1 trial and in STAR trial.

Under these assumptions, a calculation was performed quantifying the treatment effect of raloxifene on invasive breast cancer compared with tamoxifen by using the observed relative risk of raloxifene versus tamoxifen in STAR (RR 1.02, 95% CI 0.82-1.27). The analysis indicated that raloxifene maintained at least 65% of the effect of tamoxifen on invasive breast cancer (point estimate = 97%, 95% CI 65%-128%). Had a placebo group been included in STAR, the putative relative risk of raloxifene versus placebo would have been 0.48 (95% CI 0.32-0.72).

Reviewer's Comments

1. The STAR trial was conducted by NSABP. The trial protocol was not reviewed by the

Agency prior to start of the trial.

2. The data and analyses in the STAR demonstrated that raloxifene was not superior to tamoxifen. Therefore, the STAR trial was a failed superiority trial.
3. The applicant submitted the statistical framework to the Agency before submitting NDA package and documented the non-inferiority analysis plan. However, at the time, the trial was already completed, and part of the information might not be completely unknown. In addition, the Agency didn't have any agreement on the non-inferiority analysis plan, and only agreed to review the submission.
4. Non-inferiority analysis for invasive breast cancer: because the trial failed to demonstrate the superiority, the applicant performed a non-inferiority analysis.

Historically, there were four tamoxifen prevention trials which aimed to evaluate the effect of tamoxifen on the risk reduction of breast cancer for women: Royal Marsden (N=2471 European trial), Italian (N=5408, European trial), P-1 (N=13388, North American trial) and IBIS-1 (N=7139, European trial). The participant populations of P-1 and the other three studies were very different in terms of risk profiles and the use of estrogen replacement therapy during the study periods. The three European studies allowed women to use estrogen therapy while women in the P-1 did not. The STAR trial was similar to P-1 trial in the following aspects: both were conducted in North America; neither trial allowed the use of estrogen replacement therapy; and both trials were based on the 5-year predicted risk of invasive breast cancer as enrollment criteria. Therefore, data from the P-1 trial was considered to be the most relevant to the population studied in STAR. Since the P-1 trial included both premenopausal and postmenopausal women and the information regarding menopausal status was not available, the risk reduction of tamoxifen for women 50 years of age or older from the P-1 was used in the non-inferiority analysis. According to the applicant, this data is not published in literature. Based on the P-1 data, RR of tamoxifen versus placebo was 0.47 with 95% CI (0.33-0.66).

Using Rothmann et al. method, a point estimator for the proportion of tamoxifen effect retained by raloxifene was calculated as

$$\delta = [\ln(1/0.47) - \ln(1.02)] / \ln(1/0.47) = 0.97.$$

Limits of 95% CI were calculated as solutions of two equations

$$[\ln(1.02) - (1 - \delta) * \ln(1/0.47)] / \sqrt{(4/341) + (1 - \delta)^2 * (0.1807)^2} = \pm 1.96,$$

which were 0.65 and 1.28. Here 0.1807 was derived as the square root of

$$\sqrt{[\ln(0.67) - \ln(0.33)] / (2 * 1.96)} = 0.1807,$$

assuming a normal distribution for the calculation of CI.

It is not known as to what percent retention should be used in the current setting. In the adjuvant breast cancer setting, the Agency has required at least a 75% retention of an active control effect for an efficacy claim based on a non-inferiority trial. In a prevention trial, it is not clear what the

minimum percent retention of an active control effect should be for an efficacy claim. Since 75% was not in the 95% CI of δ , one can not conclude that raloxifene was non-inferior to tamoxifen if 75% was used as a percent of retention in this non-inferior trial.

The 95% CI (0.65-1.28) was interpreted as raloxifene retained at least 65% tamoxifen effect, or equivalently, may lose up to 35% tamoxifen effect. Using direction of whether raloxifene is superior to tamoxifen, the probability (one-sided) of retaining at least 65% is equal to 97.5%.

The following table presents various percent retention and its corresponding probabilities (one-sided) of retaining at least the percent retentions:

Percentage retention δ	Probability of retaining at least δ
0.65	0.975
0.70	0.9561
0.75	0.9251
0.80	0.8747
0.85	0.7987
0.90	0.694
0.95	0.5656

The above table may serve as guidance in considering different percent retention for decision making.

A second important aspect in a non-inferiority trial to be considered is the variability within the trial in calculating the control effect used in the above calculation. In a non-inferiority analysis, estimation of the control effect from historical trials is usually done by a meta-analysis to consider any variability among trials. Since only a subpopulation of P-1 was used in the current non-inferiority analysis, the variability within trial should be considered. The RR of tamoxifen vs. placebo from a subpopulation of the P-1 study was 0.47 with 95% CI (0.33-0.66). The following table presents the lower confidence limit of a 95% CI for the percent retention if other estimates within the 95% CI of (0.33-0.66) were used instead of the point estimate of 0.47.

RR of Tamoxifen vs. placebo	Lower limit of 95% CI for the percent retention
0.50	0.60
0.53	0.55
0.56	0.48
0.59	0.39
0.62	0.24
0.65	-0.00

Note that the lower limit of 95% CI is very sensitive to the point estimate of control effect RR.

The third aspect in a non-inferiority trial is to consider whether the constancy assumption holds. Performing a non-inferiority analysis in the current setting, a key assumption is that tamoxifen versus placebo effect would be the same between the subpopulation of P-1 and STAR if a placebo arm was included in STAR. In the P-1 trial, RR was 0.57 with 95% CI (0.39-0.84) for whole population, while the RR was 0.47 with 95% CI (0.33-0.66) for the subpopulation of subjects whose ages were greater than or equal to 50 years old. In STAR trial, the subpopulation of subjects whose ages were greater than or equal to 50 years old had 160 and 165 incidences of invasive breast cancer in tamoxifen and raloxifene arms with IR 4.5 and 4.6 per 1000 patient-years, and RR 1.02 with 95% CI (0.82-1.28). Notice that for whole population in STAR, IR was 4.3 and 4.4, and RR was 1.02 with 95% CI (0.82-1.27). It appears that in STAR trial, the subpopulation and the whole population was more similar than that in the P-1 trial.

IR for the subpopulation from the P-1 were 3.21 for tamoxifen arm and 6.80 for placebo arm, and for whole population of P-1 were 9.6 for tamoxifen arm and 16.73 for placebo arm. Comparing IR between the subpopulation and STAR for tamoxifen arm, it is not clear whether the constancy assumption will hold. In addition, the IRs in the subpopulation of the P-1 and whole population for tamoxifen in the P-1 trial was different.

A putative placebo analysis was performed: if there was a placebo arm included in STAR trial, the point estimate of a putative relative risk of raloxifene versus placebo would have been

$$\exp(\ln 1.02 - \ln(1/0.47)) = 0.48$$

with a 95% CI calculated as

$$\exp[\ln 1.02 - \ln(1/0.47) \pm 1.96 * \sqrt{(4/341) + (0.1807)^2}]$$

which were 0.32 to 0.72.

However, since model assumptions are impossible to verify plus the variability among trials and within the trial are very difficult to evaluate, the result of the putative analysis is hard to interpret.

5. Analyses for the incidence of invasive breast cancer by invasiveness and estrogen receptor status are presented below.

Appears This Way
On Original

Breast Cancer Category	Number Events		IR ^a			RR (95% CI) ^b
	Tam	Evista	Tam	Evista	Difference ^c	
Invasive	168	173	4.30	4.40	-0.10	1.02 (0.82, 1.27)
ER Pos	120	115	3.07	2.93	0.14	0.95 (0.73, 1.24)
ER Neg	46	52	1.18	1.32	-0.14	1.12 (0.74, 1.71)
ER Unkn	2	6	0.05	0.15	-0.10	2.98 (0.53, 30.21)
Non-Invasive	60	83	1.54	2.12	-0.58	1.38 (0.98, 1.95)
DCIS	32	47	0.82	1.20	-0.38	1.46 (0.91, 2.37)
LCIS	23	29	0.59	0.74	-0.15	1.26 (0.70, 2.27)
Mixed	5	7	0.13	0.18	-0.05	1.39 (0.38, 5.57)

^aIR=rate per 1000 patients years

^bRisk ratio for Evista compared to tamoxifen

^cRate in tamoxifen group minus rate in Evista group

6. Analyses of important safety outcomes are presented below.

Events	# Events (%)		IR		RR (95% CI)
	Tamoxifen N=9736	Raloxifene N=9751	Tamoxifen	Raloxifene	
Clinical vertebral fracture	58	58	1.47	1.46	0.99(0.68,1.46)
Death	109	104	2.76	2.62	0.95(0.72,1.25)
Death due to stroke	7	5	0.18	0.13	0.71(0.18,2.60)
Stroke	56	54	1.42	1.36	0.96(0.65,1.42)
Deep Vein Thrombosis	92	67	2.35	1.69	0.72(0.52,1.00)
Pulmonary Embolism	58	38	1.47	0.96	0.65(0.42,1.00)
Endometrial Cancer ^a	37/4739	23/4715	1.99	1.21	0.61(0.34,1.05)
Ovarian Cancer	14	18	0.52	0.66	1.27(0.60,2.76)
Cataracts	435	343	13.19	10.34	0.78(0.68,0.91)
Hysterectomy	246/4739	92/4715	13.25	4.84	0.37(0.28,0.47)
Hot Flashes	7170	6748	181.71	169.91	0.94(0.90,0.97)
Leg Cramps	5999	5373	152.03	135.29	0.89(0.86,0.92)
Edema ^b	664	741	16.83	18.66	1.11(1.00,1.23)

^a Only patients with a uterus at baseline (tamoxifen n = 4739; raloxifene n = 4715)

^a Hysterectomy was calculated as a risk ratio.

^b Peripheral edema is not a coding term in CTC v2.0.

7. Analysis for incidence of all breast cancer

The following table presents the analysis for the incidence of all breast cancer. The incidence rate in raloxifene arm was greater than that in tamoxifen arm.

Number of Events (%)		IR		RR (95% CI)
Tamoxifen N=9736	Raloxifene N=9751	Tamoxifen	Raloxifene	
228 (2.3%)	256 (2.6%)	5.85	6.54	1.12 (0.93, 1.34)

Similar to the non-inferiority analysis of the incidence of invasive breast cancer, the point estimate of percent retention is 0.85 with 95% CI (0.53, 1.09).

Summary

For STAR trial, the stratified log-rank test for the primary analysis was not statistically significant. A post-hoc non-inferiority analysis, which used the sub-population from the NSABP-P1 trial, was performed to estimate tamoxifen effect size. The NSABP-P1 trial was conducted to compare the incidence of invasive breast cancer between tamoxifen and placebo. The data of women 50 years of age or older from the NSABP-P1 trial showed that tamoxifen decreased the incidence of invasive breast cancer by 53%. If one assumes that tamoxifen would have the same effect in STAR trial if a placebo arm would be included in STAR trial, the results of the non-inferiority analysis indicated that raloxifene may lose up to 35% of tamoxifen effect.

Since STAR trial failed to demonstrate superiority of raloxifene over tamoxifen which was the primary goal of the trial, any additional analyses used to support any claims violate the statistical principle. One may argue that the STAR trial was a large trial which should not be ignored and it will be almost impossible to repeat such a trial. However, a large trial, in many aspects, is not necessarily a better trial. The reason for a large trial in this case is that the effect is too small and events are rare. When the effect size is small and sample size is large, it is often difficult to control many confounding factors.

In a non-inferiority analysis setting, a percent retention needs to be pre-specified, the control effect used needs to be well estimated which is often derived from several similar historical trials, variability among trials and within trials needs to be considered, and the constancy assumption which assumes that tamoxifen would have the same effect over placebo between NSABP-P1 trial and STAR trial if a placebo arm would be included in STAR trial needs to be satisfied. In the current non-inferiority analysis, the percent retention was not pre-specified, the control effect was derived from a subpopulation of one trial, variability within the trial was not discussed, and the validity of the constancy assumption is very difficult or impossible to verify.

Appears This Way
On Original

3.1.2 Study H3S-MC-GGIO (RUTH)

3.1.2.1 Objective

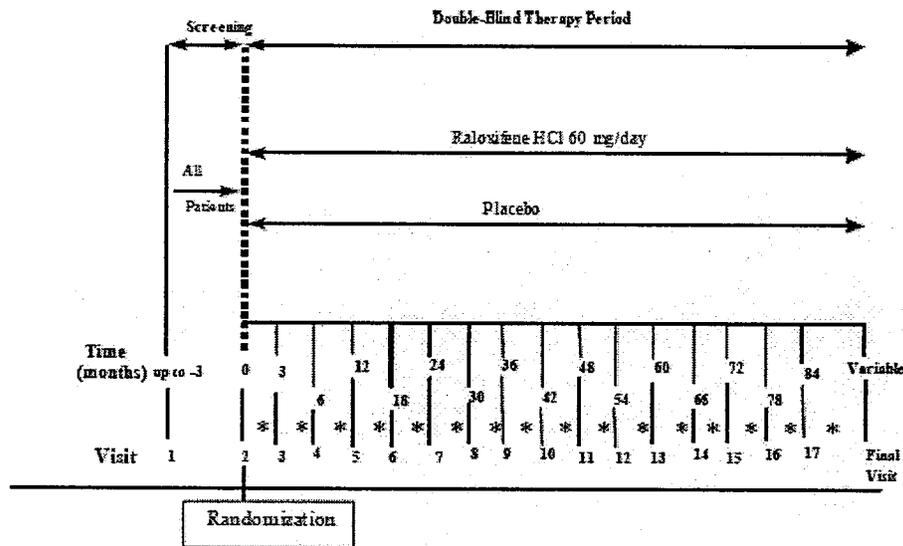
The two primary objectives of RUTH were to assess whether chronic oral treatment with raloxifene HCl 60 mg/day, compared with placebo, reduced the incidence of

1. combined coronary endpoint events of coronary death, nonfatal (including silent) myocardial infarction (MI), or hospitalized acute coronary syndrome (ACS) other than MI; or
 2. Invasive breast cancer
- in postmenopausal women at risk for major coronary events.

3.1.2.2 Study Design

RUTH was a randomized, double-blind, placebo-controlled, multinational study conducted in postmenopausal women at risk for major coronary events. A total of 10,101 postmenopausal women with established CHD or at increased risk for CHD were randomly assigned to either placebo (N=5057) or raloxifene HCl 60 mg/day (N=5044). The active treatment phase ended after the last randomized patient had been followed for at least 5 years.

Figure 3.1.2.2.1 Study design for RUTH



*Retention-related telephone contacts may occur between scheduled visits.

Women aged 55 years or older who were at least 1 year postmenopausal, and who had established CHD or multiple CHD risk factors, were eligible to enroll. A CV risk score of 4 or greater was required for enrollment, using the following point system: established CHD (4 points); lower extremity arterial disease (4 points); diabetes mellitus (3 points); age 70 years or greater (2 points); current smoker (1 point); hypertension (1 point); or hyperlipidemia (1 point). Exclusion criteria were an MI, a coronary artery bypass graft (CABG), or a percutaneous coronary intervention (PCI) within 6 months of randomization; being investigated for suspected breast carcinoma or with a known history of breast carcinoma; a history of cancer or VTE; unexplained uterine bleeding within 6 months of randomization; class III or IV heart failure; chronic liver or renal disease; use of oral or transdermal estrogens within 6 months of randomization; or concurrent use of other sex hormones or selective estrogen receptor modulators (SERMs).

Each patient's 5-year predicted risk of invasive breast cancer was calculated at baseline using the modified Gail model. The breast cancer risk factors in the model include: current age; age at menarche; nulliparity or age at first live birth; number of female first degree relatives diagnosed with breast cancer; number of breast biopsies; presence of atypical hyperplasia in a biopsy sample; and race.

Study drug was permanently discontinued when a participant was unblinded or diagnosed with breast cancer or venous thromboembolism. The use of CV medications to treat CHD or CHD risk factors was encouraged.

Bilateral mammograms were performed at baseline (within 12 months before randomization), every 2 years thereafter, and at the final visit; clinical breast examinations were performed at baseline (within 3 months before randomization) and every 2 years thereafter. All investigator-reported cases of breast cancer were reviewed and adjudicated by a board of physicians who were blinded to patient treatment assignment and who were not employed by Applicant. The adjudicators determined: 1) whether the patient had a primary breast cancer, and whether it was invasive or noninvasive; 2) what the ER status was; 3) whether the cancer may have been pre-existing (i.e., evident on the baseline mammogram) or was new (i.e., identified on a post baseline mammogram). The diagnosis of a breast cancer was based on the findings reported in the local pathology report (or equivalent document describing the pathology findings). The following items were used to determine whether or not a breast cancer may have been preexisting: mammogram films from baseline through diagnosis, related radiology reports, and any reports provided for additional studies performed, such as magnification views or an ultrasound. Estrogen receptor status of the tumor was ascertained from the pathology report (i.e., immunocytochemical assay). Non-cancer cases were not included as controls in the adjudication process; therefore, ascertainment bias was a possibility.

3.1.2.3 Efficacy Measures

Invasive Breast Cancer Primary Endpoint

Mammograms were obtained at randomization (Visit 2), or within 12 months prior to randomization, and every 2 years thereafter throughout the study; the maximum number of scheduled mammograms that could have been performed for each patient during the study was 4. Clinical breast exams were performed at baseline and every 2 years thereafter throughout the study except at final visit.

All cases of an investigator-reported breast cancer were adjudicated in a blinded fashion by the Breast Cancer Endpoint Committee (BCEC), composed of a medical oncologist (Chair), a surgical oncologist, and a radiologist, none of whom were employees of the applicant. The BCEC adjudicated: 1) whether the patient had a primary breast cancer, and whether it was invasive or non-invasive; 2) what the estrogen receptor (ER) status was; 3) whether the cancer was pre-existing (i.e., evident on the baseline mammogram) or was new (i.e., identified on a post baseline mammogram).

Mammogram films for only those patients diagnosed with a breast cancer, either originals or copies, were sent to the BCEC radiologist for central reading. The BCEC radiologist interpreted the films, and determined if a mammographic abnormality was preexisting based on comparison to baseline films. Mammogram films from patients not diagnosed with a breast cancer were not sent to the radiologist for review. Since no internal control was employed in the interpretation of the films, bias may have been introduced in the determination of whether a breast cancer was pre-existing.

An endpoint package, including copies of the mammograms and the BCEC radiologist's interpretation of the films, was sent to the remaining two committee members for review. If the two adjudication committee members agreed on an event, the case decision was documented in the clinical trial database. If the two adjudication committee members disagreed on an event, the Chair re-reviewed the endpoint package and reviewed both adjudication forms and made the final decision. In the event that a case was not adjudicated to be a primary breast cancer, the applicant notified the investigator of any information included in the endpoint package which could be pertinent to the care of the patient.

The diagnosis of a breast cancer was based on the findings reported in the local pathology report (or equivalent document describing the pathology findings). The following items were used to determine whether or not a breast cancer was preexisting: mammogram films from baseline through diagnosis, the related radiology reports, and any reports provided for additional studies performed, such as magnification views, or an ultrasound. Estrogen receptor status of the tumor was ascertained from the pathology report (i.e., immunocytochemical assay).

Coronary Primary Endpoint

The Coronary Primary Endpoint Committee (CPEC) was comprised of 10 cardiologists, including the Chair, none of whom were employees of Applicant. This committee adjudicated, in a blinded fashion, all investigator-reported events of coronary death, MI, and hospitalized ACS other than MI. Endpoint packages for investigator-reported MI, hospitalized ACS other than MI, and coronary death were sent to two CPEC members, chosen by a statistically generated random rotation schedule, for

review. If the two adjudication committee members were able to adjudicate and agree on an event, the case decision was recorded in the clinical trial database. If the two adjudication committee members were unable to adjudicate, or if they disagreed on an event, the case was forwarded to the Chair for either final adjudication or the determination that a full committee review was necessary. Any committee member could request review of an event by the entire adjudication committee, but the Chair ultimately determined if a full committee review was warranted.

The CPEC may have adjudicated an investigator-reported MI as a hospitalized ACS, or an investigator-reported hospitalized ACS other than MI as an MI, if appropriate clinical criteria were met. An investigator-reported coronary death may have been adjudicated to another cause of death if criteria for a coronary death were not met.

The CPEC reviewed investigator-reported CV non-coronary and non-CV deaths before they were sent to Secondary Endpoint Committee Chair for adjudication. The purpose of this review was to determine if the cause of death was due to a coronary etiology, and if it met the criteria for a coronary death, as defined for this study. An investigator-reported death may have been adjudicated as a coronary death if the criteria for a coronary death were met.

MI was diagnosed if a patient had at least one of the following criteria:

Criteria I: Ischemic symptoms in the presence of abnormal cardiac enzymes or markers. Abnormal serum enzymes or markers were defined as follows:

- Creatine kinase isoenzyme MB (CK-MB) greater than 2 times the upper limit of normal local laboratory value
- If CK-MB not measured, total creatine kinase (CK) greater than 2 times the upper limit of normal local laboratory value
- If CK-MB or CK not measured, lactate dehydrogenase (LDH) and/or aspartate transaminase/alanine transaminase (AST/ALT) greater than 2 times the upper limit of normal local laboratory value
- Elevation in other biochemical markers including troponin.

Criteria II: Ischemic symptoms with new equivocal ECG changes, indicating ischemia and abnormal cardiac enzymes or markers. Abnormal serum enzymes or markers were defined as for Criteria I above.

Criteria III: New pathological Q wave on ECG, with ischemic symptoms and/or abnormal cardiac enzymes or markers. Abnormal serum enzymes or markers were defined as follows:

- CK-MB greater than the upper limit of normal local laboratory value
- If CK-MB not measured, CK greater than the upper limit of normal local laboratory value
- If CK-MB or CK not measured, LDH and/or AST/ALT greater than the upper limit of normal local laboratory value
- Elevation in other biochemical markers including troponin.

Criteria IV: New pathological Q waves on ECG in the absence of cardiac enzyme or marker changes or ischemic symptoms (silent MI)

Criteria V: Markedly abnormal cardiac enzymes or markers or new pathological Q waves following invasive coronary procedures. Abnormal cardiac enzymes defined as follows:

- Percutaneous coronary intervention: (balloon angioplasty, stent, atherectomy, laser, etc): CK-MB or CK greater than 3 times the upper limit of normal local laboratory value
- Coronary artery bypass graft: CK-MB or CK greater than 5 times the upper limit of normal local laboratory value

Hospitalized ACS other than MI was diagnosed if the patient was admitted to the hospital for, or if she developed during a hospitalization, cardiac symptoms with new ST-T changes on ECG or abnormal cardiac-specific enzymes or troponin levels defined as follows:

- CK-MB greater than the upper limit of normal for the local laboratory but less than or equal to 2 times that limit
- If CK-MB not measured or available, total CK greater than the upper limit of normal for the local laboratory but less than or equal to 2 times that limit.
- Abnormal troponin was defined as troponin (either I or T) greater than the upper limit of normal for the local laboratory.

Deaths for this study were classified as either CV (coronary or non-coronary) or non-CV in etiology. The cause of death was attributed to a coronary etiology when evidence surrounding the death suggested one of the following:

- acute MI (definite or probable),
- sudden death within 24 hours of being seen by another person (i.e., family, friends, neighbors, physicians),
- an unwitnessed death in the absence of other likely non-coronary etiologies,
- death related to undergoing a coronary artery procedure,
- death due to heart failure in the presence of coronary artery disease.

The cause of death was attributed to a CV non-coronary etiology when evidence surrounding death was secondary to one of the following: cerebrovascular disease (stroke or other cause), peripheral vascular disease, a non-coronary arterial procedure, venous thromboembolic event, endocarditis/myocarditis, valvular disease, or a CV cause not otherwise specified on the CRF.

Secondary endpoints included all breast cancer (All breast cancer cases included invasive breast cancers, non-invasive breast cancers, and any other breast cancers for which invasiveness could not be determined with certainty as indicated by the BCEC on the adjudication form), death, stroke, VTE, biochemical markers of CV risk, and safety.

3.1.2.4 Sample Size Considerations

A sample size requirement of approximately 10,000 patients was calculated, which provided sample size estimates for tests based on a log-rank statistic, after adjusting for complex trial characteristics, such as staggered accrual, time-varying hazard ratios, treatment benefit lag, losses to follow-up, and noncompliance. The sample size and other trial characteristics, such as interim analysis power, were also validated through simulation. Trial assumptions were based on information from raloxifene cardiovascular advisors and a review of the relevant literature. The following assumptions were used: (1) final analysis significance level (two-sided type I error) of 0.0423 for the coronary primary endpoint and 0.008 for the breast cancer primary endpoint; (2) 80% power for the coronary primary endpoint and 80% power for the breast cancer primary endpoint; (3) uniform patient accrual over 2.25 years; (4) annual placebo-group event rate of 2.0% for the coronary primary endpoint and 0.3% for the breast cancer primary endpoint; (5) raloxifene treatment benefit lag of 9 months for the coronary primary endpoint and 3 months for the breast cancer primary endpoint; (6) after the lag period, 20% risk reduction for the coronary primary endpoint and 58.5% risk reduction for the breast cancer primary endpoint with raloxifene, before adjusting for losses and noncompliance; (7) annual loss to follow-up rate of 0.8% in each treatment group (incorporates loss rate due to documented noncardiovascular deaths); (8) drop-out (permanent discontinuation of raloxifene therapy) rate of 8% in the first year and 2% per year thereafter; (9) among those assigned to placebo, an annual drop-in (receiving a drug with efficacy assumed to be similar to that of raloxifene) rate of 1%.

3.1.2.5 Interim Analysis

The DSMB met approximately every 6 months to (1) review available safety and efficacy data and (2) ensure that the event rates, enrollment rates, and dropout rates met protocol projections. In addition, the DSMB chair or designee reviewed trial data between meetings of the full DSMB to ensure that data were reviewed at 3-month intervals.

One interim efficacy analysis of the breast cancer primary endpoint was planned. This interim analysis was performed on January 27, 2003. Only adjudicated invasive breast cancer endpoints were used in the analysis. The statistical guideline for a conclusion of significant evidence of efficacy was a two-sided test with significance level $\alpha=0.000001$. The final breast cancer analysis was conducted at $\alpha=0.008$ level.

Three interim efficacy analyses were planned for the coronary primary endpoint in this study. Two interim analyses actually occurred. Only adjudicated coronary endpoint events were used in the analyses. The first coronary interim analysis was conducted on January 28, 2002, at which time 328 patients (26% of 1268) had experienced a coronary primary endpoint event. The second interim analysis occurred on September 8, 2003, at which time 708 patients (56% of 1268) had experienced a coronary primary endpoint event. At each of the coronary interim analyses, the statistical guideline for a conclusion of significant evidence of efficacy was a two-sided test with significance level

$\alpha=0.0001$. The final analysis was conducted at $\alpha=0.0423$ level. This ensured an overall type I error rate of 0.04234 for the coronary primary endpoint.

In 2004, the applicant identified a discrepancy between the protocol and the informed consent document (ICD). The ICD stated that patients would participate in the study for between 5 and 7.5 years, while the protocol stated that patients would be followed until a minimum of 1268 patients experience a coronary primary endpoint event, centrally adjudicated as such. A blinded analysis of coronary primary endpoint event rate conducted by the applicant indicated that it would not be possible to attain 1268 endpoints within the maximum follow-up period of 7.5 years, thus requiring patients to re-consent to extend their duration of participation in the trial beyond 7.5 years. The applicant, Executive Committee (EC), and DSMB agreed that it was not in the best interest of patients to extend their participation beyond 7.5 years. Given that the 1268 coronary primary endpoint events specified in the protocol could not be achieved as described above, the applicant and EC requested that the DSMB assessed two other options: continuing the trial until each woman randomized and surviving had been followed for a minimum of 5 years or continuing the trial until each woman had been followed for, at most, 7.5 years, as specified in the ICD. On 06 December 2004, the DSMB met to review the accumulating data, and accepted the 5-year option (i.e., the trial would conclude after all surviving women in the trial have been followed for a minimum of 5 years). Based on this, patient follow-up ranged from 5 to 7 years (median of 5.6 years). Given this change, the planned third coronary primary endpoint interim analysis was not performed.

3.1.2.6 Statistical Analysis Methods

All analyses were performed according to the intent-to-treat (ITT) principle, which included all randomized patients.

The primary analysis for the breast cancer primary endpoint was a log-rank test, based on time-to-first invasive breast cancer, and planned to be performed to compare the survival functions between treatment groups. The log-rank test was conducted at the two-sided significance level of 0.008.

The primary analysis for the coronary primary endpoint was a log-rank test, based on the time from randomization to the occurrence of first event. Because the coronary primary endpoint was a composite endpoint composed of multiple individual events, time-to-first event was interpreted as time to whichever event occurred first. A log-rank test based on time-to-first event was performed to compare the survival functions between treatment groups. The log-rank test was conducted at the significance level of 0.0423 for the coronary primary endpoint.

Reviewer's Comments

The protocol was amended on 13 April 2000 to add invasive breast cancer event as a second, separate primary endpoint, based on data from Study H3S-MC-GGGK (GGGK) which showed a significant 76% reduction in the incidence of invasive breast cancer at 40 months in women assigned

raloxifene. Additionally, hospitalized ACS other than MI was added to the existing composite coronary primary endpoint (coronary death and MI) because the significantly lower than anticipated rate of coronary primary endpoint events reported up to that time would have resulted in more than double the study duration. The secondary objectives, study design, schedule of events, efficacy measures and analyses, data analysis methods, AEs and SAEs, ethical review, DSMB, and safety monitoring sections were updated to include events specified in the revised primary objectives. Other clarifications were made to exclusion criteria, blinding, patient assignment, laboratory tests, discontinuations, pharmacokinetic analysis, and the references.

The protocol was amended on 03 May 2002 to change the plan for the final analysis to occur when 1268 patients had experienced a primary coronary endpoint event, adjudicated as such. This change was due to the difference between the expected and observed rate of coronary primary endpoint events, even after the addition of the third coronary primary endpoint component. The observed rate of events would have led to a follow-up period that would have exceeded a decade. This change was expected to keep the length of the study similar to the one initially expected.

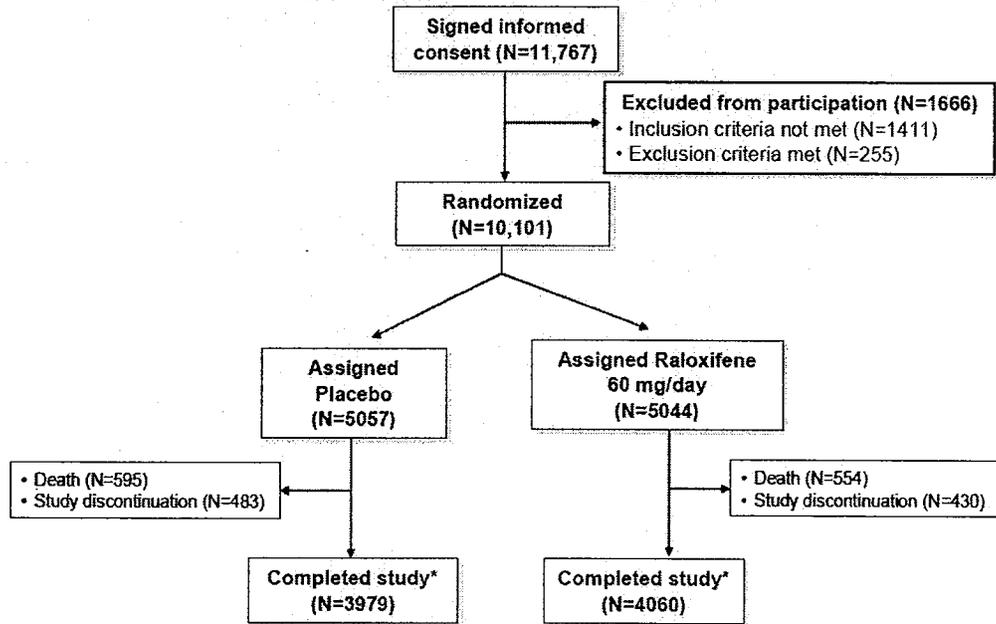
3.1.2.7 Applicant's Results and Statistical Reviewer's Findings/ Comments

3.1.2.7.1 Study Population

A total of 11,767 women signed the informed consent document with 10,101 being randomized to study groups (5044 raloxifene; 5057 placebo). A total of 4060 patients in the raloxifene group and 3979 patients in the placebo group completed the study (Figure 3.1.2.7.1.1).

Appears This Way
On Original

Figure 3.1.2.7.1.1 Patient Disposition



*Final visit on or after March 1, 2005

Table 3.1.2.7.1.1 shows the reasons for study discontinuation. A total of 2062 (20%) patients discontinued from the study before it was concluded. More raloxifene-assigned patients completed the study compared with placebo-assigned patients, though the absolute difference in the proportion of completers was small. More placebo-treated patients compared with raloxifene-treated patients discontinued from the study due to inadvertent summarization; however, this occurred in a very low proportion of patients (<2% in each treatment group).

Reviewer's Comments

P-values in tables below are for descriptive purpose only due to lack of adjustment of multiple comparisons.

Appears This Way
On Original

**Table 3.1.2.7.1.1 Subject Disposition (ITT)
Reasons for Study Discontinuation (All Randomized Patients)**

Study discontinuation reason	Placebo (N=5057)		Ralox (N=5044)		Total (N=10101)		p-Value*
	n	(%)	n	(%)	n	(%)	
Study is concluded	3979	(78.68)	4060	(80.49)	8039	(79.59)	.024
Death	595	(11.77)	554	(10.98)	1149	(11.38)	.216
Withdrawal of consent	277	(5.48)	270	(5.35)	547	(5.42)	.782
Lost to follow-up	120	(2.37)	105	(2.08)	225	(2.23)	.321
Inadvertent summarization	86	(1.70)	55	(1.09)	141	(1.40)	.009

*p-Value is obtained from a Pearson's Chi-square test.

The study was completed by 79% of women in the placebo group and 80% in the raloxifene group. Overall, 71% of patients in the placebo group and 70% in the raloxifene group took at least 70% of assigned medication. The median duration of follow-up was 5.6 years and the median study drug exposure was 5.1 years for both treatment groups.

3.1.2.7.2 Demographic and Baseline Characteristics

Table 3.1.2.7.2.1 presents patient demographic characteristics at baseline. Patient demographic characteristics were balanced between treatment groups at baseline.

Appears This Way
On Original

**Table 3.1.2.7.2.1 Patient Demographic Characteristics at Baseline
All Randomized Patients**

Characteristic	Placebo (N=5057)	Ralox (N=5044)	Total (N=10101)	p-Value*
Age (yrs) (b)				
No. patients	5057	5044	10101	.859
Mean	67.49	67.46	67.47	
Standard deviation	6.68	6.62	6.65	
Median	67.62	67.65	67.62	
Minimum	54.84	54.71	54.71	
Maximum	91.96	87.77	91.96	
Race (a)				
No. patients	5057	5044	10101	.980
Caucasian, n (%)	4247(83.98)	4234(83.94)	8481(83.96)	
Hispanic, n (%)	260(5.14)	260(5.15)	520(5.15)	
East Asian, n (%)	251(4.96)	254(5.04)	505(5.00)	
Other, n (%)	195(3.86)	196(3.89)	391(3.87)	
Afro-Caribbean, n (%)	63(1.25)	66(1.31)	129(1.28)	
West Asian, n (%)	41(0.81)	34(0.67)	75(0.74)	
Region (a)				
No. patients	5057	5044	10101	1.000
Western Europe, n (%)	2343(46.33)	2336(46.31)	4679(46.32)	
Eastern Europe, n (%)	1156(22.86)	1154(22.88)	2310(22.87)	
Latin/South America, n (%)	683(13.51)	687(13.62)	1370(13.56)	
North America, n (%)	515(10.18)	514(10.19)	1029(10.19)	
Asia Pacific, n (%)	251(4.96)	247(4.90)	498(4.93)	
Africa, n (%)	109(2.16)	106(2.10)	215(2.13)	

(a) Categorical variable. *p-Value is obtained from a Pearson's Chi-square test.

(b) Continuous variable. *p-Value is obtained from an F-test using Type III Sum of Squares from an ANOVA model: response=therapy.

Appears This Way
On Original

Characteristic	Placebo (N=5057)	Ralox (N=5044)	Total (N=10101)	p-Value*
Country (a)				
No. patients	5057	5044	10101	1.000
Poland, n (%)	366 (7.24)	363 (7.20)	729 (7.22)	
Netherlands, n (%)	303 (5.99)	307 (6.09)	610 (6.04)	
Denmark, n (%)	288 (5.70)	281 (5.57)	569 (5.63)	
Hungary, n (%)	277 (5.48)	273 (5.41)	550 (5.45)	
Brazil, n (%)	270 (5.34)	272 (5.39)	542 (5.37)	
United States, n (%)	272 (5.38)	267 (5.29)	539 (5.34)	
Czech Republic, n (%)	262 (5.18)	265 (5.25)	527 (5.22)	
Russia, n (%)	251 (4.96)	253 (5.02)	504 (4.99)	
Mexico, n (%)	251 (4.96)	252 (5.00)	503 (4.98)	
Finland, n (%)	251 (4.96)	247 (4.90)	498 (4.93)	
Canada, n (%)	243 (4.81)	247 (4.90)	490 (4.85)	
Switz/Germany, n (%)	235 (4.65)	236 (4.68)	471 (4.66)	
Norway, n (%)	227 (4.49)	226 (4.48)	453 (4.48)	
Israel, n (%)	223 (4.41)	220 (4.36)	443 (4.39)	
Spain, n (%)	212 (4.19)	215 (4.26)	427 (4.23)	
Taiwan, n (%)	201 (3.97)	199 (3.95)	400 (3.96)	
Argentina, n (%)	162 (3.20)	163 (3.23)	325 (3.22)	
Sweden, n (%)	154 (3.05)	155 (3.07)	309 (3.06)	
UK/Ireland, n (%)	149 (2.95)	146 (2.89)	295 (2.92)	
Italy, n (%)	116 (2.29)	124 (2.46)	240 (2.38)	
South Africa, n (%)	109 (2.16)	106 (2.10)	215 (2.13)	
Belgium, n (%)	98 (1.94)	100 (1.98)	198 (1.96)	
France, n (%)	87 (1.72)	79 (1.57)	166 (1.64)	
Singapore, n (%)	50 (0.99)	48 (0.95)	98 (0.97)	
Weight (kg) (b)				
No. patients	5048	5035	10083	.447
Mean	71.82	72.03	71.92	
Standard deviation	13.80	13.90	13.85	
Median	70.00	70.00	70.00	
Minimum	34.20	37.00	34.20	
Maximum	149.00	147.50	149.00	
Height (cm) (b)				
No. patients	5044	5031	10075	.481
Mean	158.09	158.00	158.05	
Standard deviation	6.84	6.83	6.83	
Median	158.00	158.00	158.00	
Minimum	126.00	130.30	126.00	
Maximum	180.50	184.00	184.00	

(a) Categorical variable. *p-Value is obtained from a Pearson's Chi-square test.

(b) Continuous variable. *p-Value is obtained from an F-test using Type III Sum of Squares from an ANOVA model: response=therapy.

Appears This Way
On Original

Characteristic	Placebo (N=5057)	Ralox (N=5044)	Total (N=10101)	p-Value*
BMI (kg/m²) (b)				
No. patients	5041	5030	10071	.270
Mean	28.72	28.83	28.77	
Standard deviation	5.11	5.18	5.14	
Median	28.06	28.20	28.13	
Minimum	15.18	16.36	15.18	
Maximum	51.95	62.19	62.19	
Obesity** (a)				
No. patients	5041	5030	10071	.771
Yes, n (%)	1919(38.07)	1929(38.35)	3848(38.21)	
No, n (%)	3122(61.93)	3101(61.65)	6223(61.79)	
Waist circumference (cm) (b)				
No. patients	5000	4973	9973	.860
Mean	93.90	93.85	93.87	
Standard deviation	13.11	13.23	13.17	
Median	93.00	93.00	93.00	
Minimum	44.00	50.00	44.00	
Maximum	143.00	160.00	160.00	
Abdominal obesity (waist circumference > 88 cm) (a)				
No. patients	5000	4973	9973	.940
Yes, n (%)	3229(64.58)	3208(64.51)	6437(64.54)	
No, n (%)	1771(35.42)	1765(35.49)	3536(35.46)	
Systolic blood pressure (mmHg) (b)				
No. patients	5057	5044	10101	.374
Mean	145.42	145.77	145.59	
Standard deviation	20.12	20.25	20.18	
Median	143.00	143.00	143.00	
Minimum	90.00	90.00	90.00	
Maximum	225.00	240.00	240.00	
Diastolic blood pressure (mmHg) (b)				
No. patients	5057	5044	10101	.984
Mean	81.96	81.96	81.96	
Standard deviation	10.35	10.51	10.43	
Median	80.00	80.00	80.00	
Minimum	0.00	27.00	0.00	
Maximum	130.00	140.00	140.00	

(a) Categorical variable. *p-Value is obtained from a Pearson's Chi-square test.

(b) Continuous variable. *p-Value is obtained from an F-test using Type III Sum of Squares from an ANOVA model: response=therapy.

**Obesity defined as BMI > 30 kg/m² except for patients in the Asia Pacific region for whom BMI > 25 kg/m² was used to define obesity.

Appears This Way
On Original

Characteristic	Placebo (N=5057)	Ralox (N=5044)	Total (N=10101)	p-Value*
Heart rate (beats/min) (b)				
No. patients	5057	5044	10101	.700
Mean	71.00	70.92	70.96	
Standard deviation	10.40	10.93	10.67	
Median	70.00	70.00	70.00	
Minimum	40.00	40.00	40.00	
Maximum	130.00	151.00	151.00	
Special diet consumed (a)				
No. patients	5055	5041	10096	.118
Yes, n (%)	3578(70.78)	3639(72.19)	7217(71.48)	
Low fat diet, n (%)	1662(52.95)	1754(55.58)	3416(54.27)	
Low cholesterol diet, n (%)	1433(49.24)	1466(51.12)	2899(50.17)	
Low salt diet, n (%)	932(38.69)	966(40.79)	1898(39.73)	
Weight reducing diet, n (%)	221(13.02)	232(14.20)	453(13.60)	
Vegetarian diet, n (%)	84(5.38)	77(5.21)	161(5.30)	
Diabetic diet, n (%)	1902(56.29)	1905(57.61)	3807(56.94)	
High fiber diet, n (%)	187(11.24)	195(12.21)	382(11.71)	
Other diet, n (%)	45(2.96)	42(2.91)	87(2.93)	
No, n (%)	1477(29.22)	1402(27.81)	2879(28.52)	
Physical activity at work (a)				
No. patients	5053	5040	10093	.484
Very physically demanding	346(6.85)	319(6.33)	665(6.59)	
Moderately physically demanding	2401(47.52)	2437(48.35)	4838(47.93)	
Minimally physically demanding	2306(45.64)	2284(45.32)	4590(45.48)	
Physical activity at leisure (a)				
No. patients	5053	5041	10094	.589
Very physically demanding	175(3.46)	192(3.81)	367(3.64)	
Moderately physically demanding	1942(38.43)	1952(38.72)	3894(38.58)	
Minimally physically demanding	2936(58.10)	2897(57.47)	5833(57.79)	
Participation in vigorous physical activity (a)				
No. patients	5054	5041	10095	.902
Not on a regular basis, n (%)	3523(69.71)	3487(69.17)	7010(69.44)	
Once a week, n (%)	307(6.07)	301(5.97)	608(6.02)	
Twice a week, n (%)	298(5.90)	307(6.09)	605(5.99)	
3 or more times a week, n (%)	926(18.32)	946(18.77)	1872(18.54)	
History of cardiac rehabilitation (a)				
No. patients	5055	5041	10096	.193
Yes, n (%)	709(14.03)	753(14.94)	1462(14.48)	
No, n (%)	4346(85.97)	4288(85.06)	8634(85.52)	

(a) Categorical variable. *p-Value is obtained from a Pearson's Chi-square test.

(b) Continuous variable. *p-Value is obtained from an F-test using Type III Sum of Squares from an ANOVA model: response=therapy.

Appears This Way
On Original

Characteristic	Placebo (N=5057)	Ralox (N=5044)	Total (N=10101)	p-Value*
Alcohol consumption (a)				
No. patients	5056	5041	10097	.679
Yes, n (%)	2177 (43.06)	2150 (42.65)	4327 (42.85)	
Less than 1 drink per week, n (%)	1313 (60.31)	1268 (58.98)	2581 (59.65)	
1 or more drinks per week, n (%)	864 (39.69)	882 (41.02)	1746 (40.35)	
No, n (%)	2879 (56.94)	2891 (57.35)	5770 (57.15)	
Current smoker** (a)				
No. patients	5057	5044	10101	.223
Yes, n (%)	649 (12.83)	607 (12.03)	1256 (12.43)	
No, n (%)	4408 (87.17)	4437 (87.97)	8845 (87.57)	
Exposure to secondary smoke (a)				
No. patients	5057	5042	10099	.171
Yes, n (%)	1331 (26.32)	1267 (25.13)	2598 (25.73)	
No, n (%)	3726 (73.68)	3775 (74.87)	7501 (74.27)	
Prior smoker (a)				
No. patients	5057	5043	10100	.268
Yes, n (%)	2180 (43.11)	2119 (42.02)	4299 (42.56)	
No, n (%)	2877 (56.89)	2924 (57.98)	5801 (57.44)	
Abnormal electrocardiogram*** (a)				
No. patients	4966	4943	9909	.231
Yes, n (%)	2052 (41.32)	1984 (40.14)	4036 (40.73)	
No, n (%)	2914 (58.68)	2959 (59.86)	5873 (59.27)	
Electrocardiographic Q-wave MI (a)				
No. patients	2050	1981	4031	.808
Yes, n (%)	571 (27.85)	545 (27.51)	1116 (27.69)	
No, n (%)	1479 (72.15)	1436 (72.49)	2915 (72.31)	
First degree female relative with heart disease diagnosed before age 65 (a)				
No. patients	4891	4855	9746	.285
Yes, n (%)	1209 (24.72)	1155 (23.79)	2364 (24.26)	
No, n (%)	3682 (75.28)	3700 (76.21)	7382 (75.74)	
First degree male relative with heart disease diagnosed before age 55 (a)				
No. patients	4824	4782	9606	.217
Yes, n (%)	1200 (24.88)	1242 (25.97)	2442 (25.42)	
No, n (%)	3624 (75.12)	3540 (74.03)	7164 (74.58)	

(a) Categorical variable. *p-Value is obtained from a Pearson's Chi-square test.

(b) Continuous variable. *p-Value is obtained from an F-test using Type III Sum of Squares from an ANOVA model: response=therapy.

**Patient has smoked an average of ≥ 10 cigarettes per day during the 6 months prior to baseline.

***Definite Q-wave MI; pathologic ST-T depression; conduction disturbance excluding 1st-degree atrioventricular block; atrial fibrillation or flutter; ventricular hypertrophy.

Appears This Way
On Original

Characteristic	Placebo (N=5057)	Ralox (N=5044)	Total (N=10101)	p-Value*
Years postmenopausal (b)				
No. patients	5057	5042	10099	.325
Mean	19.49	19.32	19.40	
Standard deviation	8.80	8.82	8.81	
Median	20.00	19.00	19.00	
Minimum	1.00	1.00	1.00	
Maximum	78.00	63.00	78.00	
Prior hysterectomy (a)				
No. patients	5046	5040	10086	.483
Yes, n (%)	1175(23.29)	1144(22.70)	2319(22.99)	
No, n (%)	3871(76.71)	3896(77.30)	7767(77.01)	
Prior ovariectomy (a)				
No. patients	4981	4985	9966	.486
Yes, n (%)	774(15.54)	800(16.05)	1574(15.79)	
No, n (%)	4207(84.46)	4185(83.95)	8392(84.21)	
Prior use of estrogen only (a)				
No. patients	5002	4989	9991	.927
Yes, n (%)	702(14.03)	697(13.97)	1399(14.00)	
No, n (%)	4300(85.97)	4292(86.03)	8592(86.00)	
Prior use of estrogen plus progestin (a)				
No. patients	4964	4940	9904	.097
Yes, n (%)	323(6.51)	282(5.71)	605(6.11)	
No, n (%)	4641(93.49)	4658(94.29)	9299(93.89)	
Years of prior estrogen or estrogen plus progestin use (b)				
No. patients	982	927	1909	.153
Mean	4.24	3.90	4.07	
Standard deviation	5.81	4.75	5.32	
Median	2.00	2.00	2.00	
Minimum	1.00	1.00	1.00	
Maximum	50.00	38.00	50.00	
Prior use of oral contraceptives (a)				
No. patients	5047	5032	10079	.897
Yes, n (%)	969(19.20)	961(19.10)	1930(19.15)	
No, n (%)	4078(80.80)	4071(80.90)	8149(80.85)	
Prior use of raloxifene, tamoxifen, other SERM**, or tibolone (a)				
No. patients	5041	5032	10073	.440
Yes, n (%)	9(0.18)	6(0.12)	15(0.15)	
No, n (%)	5032(99.82)	5026(99.88)	10058(99.85)	

(a) Categorical variable. *p-Value is obtained from a Pearson's Chi-square test.

(b) Continuous variable. *p-Value is obtained from an F-test using Type III Sum of Squares from an ANOVA model: response=therapy.

**SERM=selective estrogen receptor modulator

Appears This Way
On Original

Characteristic	Placebo (N=5057)	Ralox (N=5044)	Total (N=10101)	p-Value*
Prior leg fracture (a)				
No. patients	5057	5039	10096	.519
Yes, n (%)	452 (8.94)	469 (9.31)	921 (9.12)	
No, n (%)	4605 (91.06)	4570 (90.69)	9175 (90.88)	
First degree relative with a history of a hip fracture (a)				
No. patients	4920	4894	9814	.931
Yes, n (%)	477 (9.70)	477 (9.75)	954 (9.72)	
No, n (%)	4443 (90.30)	4417 (90.25)	8860 (90.28)	
Prior cholecystectomy (a)				
No. patients	5057	5039	10096	.346
Yes, n (%)	947 (18.73)	907 (18.00)	1854 (18.36)	
No, n (%)	4110 (81.27)	4132 (82.00)	8242 (81.64)	

(a) Categorical variable. *p-Value is obtained from a Pearson's Chi-square test.

(b) Continuous variable. *p-Value is obtained from an F-test using Type III Sum of Squares from an ANOVA model: response=therapy.

Table 3.1.2.7.2.2 presents breast cancer risk assessment characteristics at baseline. Breast cancer risk assessment characteristics were balanced between treatment groups at baseline. The median 5-year predicted invasive breast cancer risk was 1.55% and approximately 41% of patients in each treatment group had a 5-year predicted invasive breast cancer risk of $\geq 1.66\%$.

Appears This Way
On Original

**Table 3.1.2.7.2.2 Breast Cancer Risk Assessment Characteristics at Baseline
All Randomized Patients**

Characteristic	Placebo (N=5057)	Ralox (N=5044)	Total (N=10101)	p-Value*
5-year predicted invasive breast cancer risk, % (b)				
No. patients	5056	5044	10100	.853
Mean	1.73	1.73	1.73	
Standard deviation	0.77	0.76	0.76	
Median	1.54	1.55	1.55	
Minimum	0.52	0.50	0.50	
Maximum	9.57	14.15	14.15	
5-year predicted invasive breast cancer risk >= 1.66% (a)				
No. patients	5056	5044	10100	.614
Yes, n (%)	2081 (41.16)	2101 (41.65)	4182 (41.41)	
No, n (%)	2975 (58.84)	2943 (58.35)	5918 (58.59)	
Age (yrs)** (a)				
No. patients	5057	5044	10101	.787
<=60 , n (%)	844 (16.69)	826 (16.38)	1670 (16.53)	
>60 and <=65 , n (%)	1033 (20.43)	1028 (20.38)	2061 (20.40)	
>65 and <=70 , n (%)	1213 (23.99)	1260 (24.98)	2473 (24.48)	
>70 and <=75 , n (%)	1291 (25.53)	1251 (24.80)	2542 (25.17)	
>75 , n (%)	676 (13.37)	679 (13.46)	1355 (13.41)	
Age at menarche (yrs) (b)				
No. patients	5039	5025	10064	.247
Mean	13.47	13.51	13.49	
Standard deviation	1.75	1.79	1.77	
Median	13.00	13.00	13.00	
Minimum	8.00	6.00	6.00	
Maximum	20.00	23.00	23.00	
Age at first live birth (yrs) (b)				
No. patients	4520	4500	9020	.313
Mean	23.34	23.43	23.38	
Standard deviation	4.53	4.37	4.45	
Median	23.00	23.00	23.00	
Minimum	12.00	13.00	12.00	
Maximum	54.00	44.00	54.00	
Number of live births (a)				
No. patients	5056	5043	10099	.535
0 , n (%)	521 (10.30)	529 (10.49)	1050 (10.40)	
1 , n (%)	800 (15.82)	816 (16.18)	1616 (16.00)	
2 , n (%)	1396 (27.61)	1438 (28.51)	2834 (28.06)	
>=3 , n (%)	2339 (46.26)	2260 (44.81)	4599 (45.54)	

(a) Categorical variable. *p-Value is obtained from a Pearson's Chi-square test.

(b) Continuous variable. *p-Value is obtained from an F-test using Type III Sum of Squares from an ANOVA model: response=therapy.

**Age categories presented are those which are used in the calculation of predicted risk.

Appears This Way
On Original

Characteristic	Placebo (N=5057)	Ralox (N=5044)	Total (N=10101)	p-Value*
Number of female first degree relatives with breast cancer (a)				
No. patients	4584	4600	9184	.712
0, n (%)	4139 (90.29)	4148 (90.17)	8287 (90.23)	
1, n (%)	402 (8.77)	418 (9.09)	820 (8.93)	
2, n (%)	36 (0.79)	28 (0.61)	64 (0.70)	
>=3, n (%)	7 (0.15)	6 (0.13)	13 (0.14)	
Number of prior breast biopsies (a)				
No. patients	5041	5027	10068	.082
0, n (%)	4574 (90.74)	4611 (91.72)	9185 (91.23)	
1, n (%)	372 (7.38)	343 (6.82)	715 (7.10)	
2, n (%)	65 (1.29)	58 (1.15)	123 (1.22)	
>=3, n (%)	30 (0.60)	15 (0.30)	45 (0.45)	
Prior breast biopsy with diagnosis of invasive cancer (a)				
No. patients	380	345	725	.340
Yes, n (%)	1 (0.26)	0 (0.00)	1 (0.14)	
No, n (%)	379 (99.74)	345 (100.00)	724 (99.86)	
Prior breast biopsy with diagnosis of ductal carcinoma in situ (a)				
No. patients	380	345	725	.137
Yes, n (%)	0 (0.00)	2 (0.58)	2 (0.28)	
No, n (%)	380 (100.00)	343 (99.42)	723 (99.72)	
Prior breast biopsy with diagnosis of lobular carcinoma in situ (a)				
No. patients	380	345	725	N/A
Yes, n (%)	0 (0.00)	0 (0.00)	0 (0.00)	
No, n (%)	380 (100.00)	345 (100.00)	725 (100.00)	
Prior breast biopsy with diagnosis of atypical hyperplasia (a)				
No. patients	380	345	725	.319
Yes, n (%)	8 (2.11)	4 (1.16)	12 (1.66)	
No, n (%)	372 (97.89)	341 (98.84)	713 (98.34)	
Prior breast biopsy with diagnosis of other breast condition (a)				
No. patients	386	349	735	.923
Yes, n (%)	379 (98.19)	343 (98.28)	722 (98.23)	
No, n (%)	7 (1.81)	6 (1.72)	13 (1.77)	

(a) Categorical variable. *p-Value is obtained from a Pearson's Chi-square test.

(b) Continuous variable. *p-Value is obtained from an F-test using Type III Sum of Squares from an ANOVA model: response=therapy.

Table 3.1.2.7.2.3 presents CV risk assessment characteristics at baseline. CV risk assessment characteristics were balanced between treatment groups at baseline except for a greater CV risk score in patients assigned to raloxifene compared with patients assigned to placebo. This difference was due to a greater proportion of patients in the raloxifene group reporting a history of CABG. A total of 12% were current smokers, 46% had diabetes mellitus, 78% had hypertension, 73% had hyperlipidemia, 11% had lower extremity arterial disease, and approximately 50% had a history of CHD.

Appears This Way
On Original

**Table 3.1.2.7.2.3 Cardiovascular Risk Assessment Characteristics at Baseline
All Randomized Patients**

Characteristic	Placebo (N=5057)	Ralox (N=5044)	Total (N=10101)	p-Value*
Cardiovascular risk score (b)				
No. patients	5057	5044	10101	.029
Mean	7.75	7.92	7.83	
Standard deviation	3.74	3.96	3.85	
Median	6.00	6.00	6.00	
Minimum	0.00	1.00	0.00	
Maximum	25.00	27.00	27.00	
Cardiovascular risk score category (a)				
No. patients	5057	5044	10101	.064
<4 , n (%)	14 (0.28)	22 (0.44)	36 (0.36)	
4-6 , n (%)	2558 (50.58)	2527 (50.10)	5085 (50.34)	
7-8 , n (%)	761 (15.05)	717 (14.21)	1478 (14.63)	
9-10 , n (%)	569 (11.25)	575 (11.40)	1144 (11.33)	
11-12 , n (%)	479 (9.47)	439 (8.70)	918 (9.09)	
>12 , n (%)	676 (13.37)	764 (15.15)	1440 (14.26)	
Age >65 and <70 (a)				
No. patients	5057	5044	10101	.402
Yes, n (%)	1219 (24.11)	1252 (24.82)	2471 (24.46)	
No, n (%)	3838 (75.89)	3792 (75.18)	7630 (75.54)	
Age >=70 (a)				
No. patients	5057	5044	10101	.625
Yes, n (%)	1980 (39.15)	1951 (38.68)	3931 (38.92)	
No, n (%)	3077 (60.85)	3093 (61.32)	6170 (61.08)	
Current smoker** (a)				
No. patients	5057	5044	10101	.223
Yes, n (%)	649 (12.83)	607 (12.03)	1256 (12.43)	
No, n (%)	4408 (87.17)	4437 (87.97)	8845 (87.57)	
Diabetes mellitus*** (a)				
No. patients	5043	5034	10077	.890
Yes, n (%)	2309 (45.79)	2298 (45.65)	4607 (45.72)	
No, n (%)	2734 (54.21)	2736 (54.35)	5470 (54.28)	
Lower extremity arterial disease (a)				
No. patients	5056	5044	10100	.890
Yes, n (%)	540 (10.68)	543 (10.77)	1083 (10.72)	
No, n (%)	4516 (89.32)	4501 (89.23)	9017 (89.28)	

(a) Categorical variable. *p-Value is obtained from a Pearson's Chi-square test.

(b) Continuous variable. *p-Value is obtained from an F-test using Type III Sum of Squares from an ANOVA model: response=therapy.

**Patient has smoked an average of >= 10 cigarettes per day during the 6 months prior to baseline.

***Patient reports diabetes mellitus and is taking oral hypoglycemic medications or insulin, or patient's fasting serum glucose > 7.8mmol/L.

Appears This Way
On Original

Characteristic	Placebo (N=5057)	Ralox (N=5044)	Total (N=10101)	p-Value*
Hyperlipidemia** (a)				
No. patients	5031	5022	10053	.745
Yes, n (%)	3701 (73.56)	3680 (73.28)	7381 (73.42)	
No, n (%)	1330 (26.44)	1342 (26.72)	2672 (26.58)	
Hypertension*** (a)				
No. patients	5056	5043	10099	.940
Yes, n (%)	3935 (77.83)	3928 (77.89)	7863 (77.86)	
No, n (%)	1121 (22.17)	1115 (22.11)	2236 (22.14)	
Prior myocardial infarction (a)				
No. patients	5057	5044	10101	.697
Yes, n (%)	1468 (29.03)	1482 (29.38)	2950 (29.21)	
3-36 mo. prior to baseline, n(%)	722 (49.18)	767 (51.75)	1489 (50.47)	
>36 mo. prior to baseline, n(%)	787 (53.61)	767 (51.75)	1554 (52.68)	
No, n (%)	3589 (70.97)	3562 (70.62)	7151 (70.79)	
Prior coronary artery bypass surgery (a)				
No. patients	5057	5044	10101	.015
Yes, n (%)	783 (15.48)	871 (17.27)	1654 (16.37)	
3-36 mo. prior to baseline, n(%)	406 (51.85)	486 (55.80)	892 (53.93)	
>36 mo. prior to baseline, n(%)	387 (49.43)	394 (45.24)	781 (47.22)	
No, n (%)	4274 (84.52)	4173 (82.73)	8447 (83.63)	
Prior catheter based coronary revascularization (a)				
No. patients	4443	4477	8920	.337
Yes, n (%)	824 (18.55)	866 (19.34)	1690 (18.95)	
6-36 mo. prior to baseline, n(%)	555 (67.35)	586 (67.67)	1141 (67.51)	
>36 mo. prior to baseline, n(%)	308 (37.38)	329 (37.99)	637 (37.69)	
No, n (%)	3619 (81.45)	3611 (80.66)	7230 (81.05)	
Prior angina pectoris with documented coronary disease (a)				
No. patients	5057	5044	10101	.143
Yes, n (%)	1638 (32.39)	1703 (33.76)	3341 (33.08)	
No, n (%)	3419 (67.61)	3341 (66.24)	6760 (66.92)	

(a) Categorical variable. *p-Value is obtained from a Pearson's Chi-square test.

(b) Continuous variable. *p-Value is obtained from an F-test using Type III Sum of Squares from an ANOVA model: response=therapy.

**Patient is taking lipid-lowering medications, has a fasting LDL-cholesterol >4.14mmol/L, or has a fasting HDL-cholesterol <1.16mmol/L and fasting triglycerides >2.82mmol/L.

***Patient reports having hypertension and is taking anti-hypertensive medications, or patient has systolic blood pressure >160mmHg or diastolic blood pressure >95mmHg.

3.1.2.7.3 Applicant's Efficacy Analyses

Breast Cancer Analyses

Table 3.1.2.7.3.1 presents the results of time-to-event analyses of breast cancer by invasiveness and ER status. The incidence of invasive breast cancer was significantly decreased by 44% in the raloxifene group compared with the placebo group. The log-rank test gave a p-value 0.0032. As the

protocol-specified significance level was 0.008, the breast cancer primary objective was achieved. There were 1.50 cases of invasive breast cancer per 1000 patients per year in the raloxifene group and 2.66 cases of invasive breast cancer per 1000 patients per year in the placebo group (Table 3.1.2.7.3.2) which translated to an absolute risk reduction of 1.2 cases per 1000 woman-years in the raloxifene group.

**Table 3.1.2.7.3.1 Time-to-Event Analysis of Breast Cancer by Invasiveness and ER Status
All Randomized Patients**

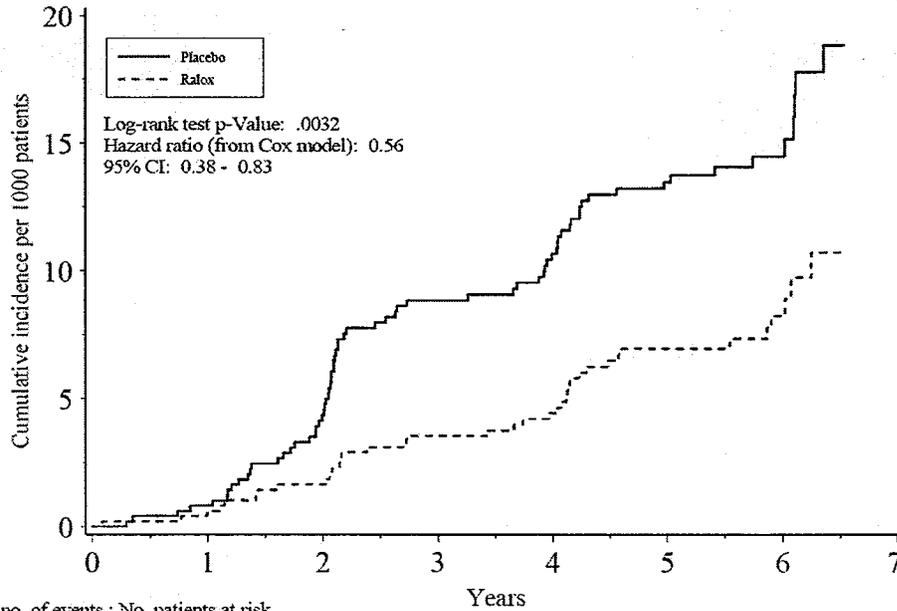
Breast cancer category	Placebo (N=5057) n (%)	Ralox (N=5044) n (%)	Hazard ratio (95% CI)	p-Value*
Invasive cases	70 (1.38)	40 (0.79)	0.56 (0.38, 0.83)	.0032
ER(+) cases	55 (1.09)	25 (0.50)	0.45 (0.28, 0.72)	.0006
ER(-) cases	9 (0.18)	13 (0.26)	1.44 (0.61, 3.36)	.3995
ER unknown cases	6 (0.12)	2 (0.04)	0.33 (0.07, 1.63)	.1507
Noninvasive cases	5 (0.10)	11 (0.22)	2.17 (0.75, 6.24)	.1414
DCIS	5 (0.10)	11 (0.22)	2.17 (0.75, 6.24)	.1414
LCIS	0 (0.00)	0 (0.00)	N/A	N/A
Invasiveness unknown cases	1 (0.02)	1 (0.02)	N/A	N/A
All cases	76 (1.50)	52 (1.03)	0.67 (0.47, 0.96)	.0270

Abbreviations: CI=confidence interval; ER=estrogen receptor; DCIS=ductal carcinoma in situ; LCIS=lobular carcinoma in situ.

*p-Value is obtained from a log-rank test. Statistical test is not performed when the total number of patients in a category is less than 5.

Appears This Way
On Original

Figure 3.1.2.7.3.1 Kaplan-Meier curves of invasive breast cancer all randomized patients



	Cumulative no. of events : No. patients at risk							
	0	1	2	3	4	5	6	7
Placebo	0 : 5057	4 : 4910	21 : 4702	42 : 4488	50 : 4312	62 : 3813	65 : 1670	
Ralox	0 : 5044	3 : 4928	8 : 4775	17 : 4604	21 : 4404	32 : 3893	35 : 1724	

Table 3.1.2.7.3.2 Incidence Rate of Breast Cancer by Invasiveness and ER Status All Randomized Patients

Breast cancer category	Placebo (N=5057)			Ralox (N=5044)			ARR** per 1000 patients
	n (%)	Patient-years of follow-up	Incidence rate* per 1000 patients	n (%)	Patient-years of follow-up	Incidence rate* per 1000 patients	
Invasive cases	70 (1.38)	26290	2.66	40 (0.79)	26695	1.50	5.85
ER(+) cases	55 (1.09)	26339	2.09	25 (0.50)	26722	0.94	5.87
ER(-) cases	9 (0.18)	26451	0.34	13 (0.26)	26763	0.49	-0.80
ER unknown cases	6 (0.12)	26466	0.23	2 (0.04)	26782	0.07	0.79
Noninvasive cases	5 (0.10)	26467	0.19	11 (0.22)	26759	0.41	-1.19
DCIS	5 (0.10)	26467	0.19	11 (0.22)	26759	0.41	-1.19
LCIS	0 (0.00)	26483	0.00	0 (0.00)	26786	0.00	0.00
Invasiveness unknown cases	1 (0.02)	26482	0.04	1 (0.02)	26784	0.04	0.00
All cases	76 (1.50)	26273	2.89	52 (1.03)	26666	1.95	4.66

Abbreviations: ER-estrogen receptor; DCIS-ductal carcinoma in situ; LCIS-lobular carcinoma in situ.
 *Incidence rate is calculated as the number of patients who developed the event of interest divided by the patient-years of follow-up.
 **Absolute risk reduction (ARR) is calculated by subtracting the cumulative incidence of the raloxifene arm from that of the placebo arm, where cumulative incidence is estimated using $1 - \exp(-I \cdot T)$, I is the incidence rate, and T is the average patient-years of follow-up in each arm.

Coronary Analyses

Appears This Way
 On Original

A total of 1595 primary coronary events in 1221 patients were reported by investigators (Table 3.1.2.7.3.3). The analyses reported were based on 1086 patients with adjudicated coronary events.

Details of the primary coronary endpoint events are listed below:

- Coronary death: A total of 468 patients were reported to have died from coronary causes during the study period (243 in placebo, 225 in raloxifene); 526 deaths were adjudicated as due to coronary causes (273 in placebo, 253 in raloxifene). Analyses of coronary deaths were based on the 526 patients with adjudicated coronary deaths, unless otherwise specified.
- Nonfatal MIs: A total of 472 patients (241 in placebo, 231 in raloxifene) were reported to have had at least one nonfatal MI during the study period. A total of 391 patients (208 in placebo, 183 in raloxifene) had at least one adjudicated nonfatal MI. Analyses of nonfatal MIs were based on the 391 patients with adjudicated nonfatal MIs, unless otherwise specified.
- Hospitalized ACS other than MI: A total of 450 patients (247 in placebo, 203 in raloxifene) were reported to have had at least one hospitalized ACS during the study period. A total of 354 patients (185 in placebo, 169 in raloxifene) had a least one adjudicated hospitalized ACS. Analyses of hospitalized ACS were based on the 354 patients with adjudicated hospitalized ACS, unless otherwise specified.

There was no significant increase or decrease in incidence of the coronary primary endpoint of combined coronary death, nonfatal MI, or hospitalized ACS other than MI in the raloxifene group compared with the placebo group (Figure 3.1.2.7.3.2). Thus, the coronary primary objective was not achieved at the protocol-specified significance level of 0.0423.

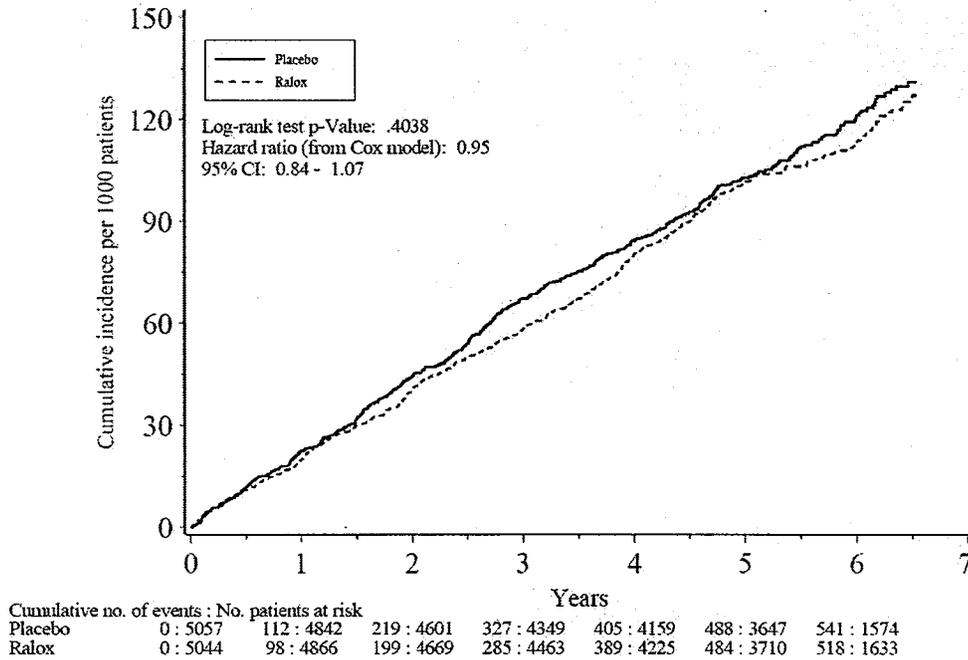
**Table 3.1.2.7.3.3 Time-to-Event Analysis of Coronary Primary Endpoint
All Randomized Patients**

Coronary endpoint	Placebo (N=5057) n (%)	Ralox (N=5044) n (%)	Hazard ratio (95% CI)	p-value*
Coronary primary endpoint	553 (10.94)	533 (10.57)	0.95 (0.84, 1.07)	.4038
Coronary death	273 (5.40)	253 (5.02)	0.92 (0.77, 1.09)	.3139
Nonfatal MI**	208 (4.11)	183 (3.63)	0.87 (0.71, 1.06)	.1639
Criterion I	40 (19.23)	40 (21.86)		
Criterion II	90 (43.27)	60 (32.79)		
Criterion III	56 (26.92)	54 (29.51)		
Criterion IV (silent MI)	9 (4.33)	24 (13.11)		
Criterion V	13 (6.25)	5 (2.73)		
Hospitalized ACS other than MI	185 (3.66)	169 (3.35)	0.90 (0.73, 1.11)	.3385
Nonfatal MI** or hospitalized ACS other than MI	360 (7.12)	326 (6.46)	0.89 (0.77, 1.04)	.1410
Coronary death or nonfatal MI**	416 (8.23)	400 (7.93)	0.95 (0.83, 1.09)	.4538

Abbreviations: CI=confidence interval; MI=myocardial infarction; ACS=acute coronary syndrome.
*p-Value is obtained from a log-rank test.
**Nonfatal MI includes silent MI.

Appears This Way
On Original

**Figure 3.1.2.7.3.2 Kaplan-Meier Curves of Coronary Primary Endpoint
All Randomized Patients**



Reviewer's Comments

1. RUTH trial failed to demonstrate the benefit of raloxifene over placebo with respect to coronary primary endpoint.
2. The co-primary endpoint of the incidence of invasive breast cancer was not originally designed as a co-primary endpoint, and was amended during the trial.
3. The analysis for invasive breast cancer events by invasiveness and estrogen receptor status are presented below. The relative risk (RR) will also be reported in summarizing the data in addition to the hazard ratio (HR) used in the applicant's report. One reason of using RR is that all medians were not reached in the trial so RR is more directly related to the annual incidence rate (IR). When the incidence rates are small, theoretically, RR and HR are approximately equal. RR is commonly used in reporting prevention trials. RR will also be used in this review for MORE and CORE trials.

Appears This Way
On Original

Breast cancer category	Raloxifene 5,044	Placebo 5,057	Raloxifene IR	Placebo IR	Absolute Risk Difference	RR (95% CI)
Invasive cases	40	70	1.50	2.66	-1.16	0.56 (0.37, 0.84)
ER(+) cases	25	55	0.94	2.09	-1.15	0.45 (0.27, 0.73)
ER(-) cases	13	9	0.49	0.34	+0.15	1.43 (0.56, 3.78)
ER unknown	2	6	0.07	0.23	-0.16	0.33 (0.03, 1.84)
Non-invasive cases	11	5	0.41	0.19	+0.22	2.18(0.70,,7.99)
DCIS	11	5	0.41	0.19	+0.22	2.18(0.70, 7.99)
LCIS	0	0	0	0	0	NA
Invasiveness unknown	1	1	0.04	0.04	+0.00	NA
All cases	52	76	1.95	2.89	- 1.04	0.67(0.46, 0.97)

Abbreviations: ER=estrogen receptor; DCIS=ductal carcinoma in situ; LCIS=lobular carcinoma in situ; RR=Relative Risk; IR= Incidence Rate (Incidence rate is calculated as the number of patients who developed the event of interest divided by the patient-years of follow-up)

4. Sensitivity analysis: there were 9 patients who had both primary coronary and invasive breast cancer events.

obs	PATIENT	THERAPY	censorco	yearco	censorbc	yearbc
1	1160	Placebo	1	1.07050	1	2.10541
2	1456	Placebo	1	4.20260	1	1.04312
3	1007	Placebo	1	5.47296	1	2.01506
4	1192	Placebo	1	4.88159	1	6.35729
5	1027	Placebo	1	0.49555	1	3.95072
6	1204	Placebo	1	3.47707	1	2.07255
7	1074	Placebo	1	0.42437	1	2.63107
8	5743	Ralox	1	4.58590	1	4.29843
9	1153	Ralox	1	6.26146	1	4.04107

A sensitivity analysis was performed as follows: if a patient had both coronary and invasive breast cancer event and the time to the invasive breast cancer was larger than the time to the coronary event, then this patient was censored at the time of the coronary event. This affected 4 patients in the placebo arm.

If a patient had only coronary event, then the patient's censoring time to the invasive breast cancer was replaced by the time to the coronary event if the time to the invasive breast cancer was larger than the time to the coronary event. This affected 345 patients in the placebo arm and 316 in the raloxifene arm.

Appears This Way
On Original

After this modification, there were 66 invasive breast cancer events in the placebo arm and 40 in the raloxifene arm. Using this definition of events, the log-rank test for the time to the invasive breast cancer was 0.0089.

5. The results of an exploratory subgroup analysis of the invasive breast cancer based on Gail score are presented below. P-values are for descriptive purpose only.

Gail Score	Invasive Breast Cancer	Raloxifene 5,044	Placebo 5,057	Absolute Risk Difference	Relative Risk (95% CI)	P-value
≥ 1.66	Subgroup	2,101	2,081	- 1.16	0.64 (0.36, 1.12)	.102
	No. Event (IR)	23 (2.09)	35 (3.25)			
< 1.66	Subgroup	2,943	2,975	- 1.11	0.49 (0.26, 0.91)	.015
	No. Event (IR)	17 (1.08)	34 (2.19)			

^a Patient 1220 had no Gail score and had invasive cancer.

6. Analyses of important safety outcomes are presented below.

Appears This Way
On Original

Events	Raloxifene 5,044	Placebo 5,057	Raloxifene IR	Placebo IR	Absolute Risk Difference	Relative Risk (95% CI)
Clinical vertebral fracture	64	97	2.40	3.70	-1.30	0.65 (0.47, 0.90)
Death	554	595	20.68	22.45	-1.77	0.92 (0.82, 1.04)
Death due to Stroke	59	39	2.20	1.47	+0.73	1.50 (0.98, 2.30)
Stroke	249	224	9.46	8.60	+0.86	1.10 (0.91, 1.32)
Deep vein thrombosis	65	47	2.44	1.78	+0.66	1.37 (0.94, 1.99)
Pulmonary embolism	36	24	1.35	0.91	+0.44	1.49 (0.89, 2.49)
Endometrial cancer ^a	21/3900	17/3882	1.01	0.83	+0.18	1.22 (0.61, 2.46)
Ovarian Cancer ^b	17/4559	10/4606	0.70	0.41	+0.29	1.71 (0.74, 4.17)
Hysterectomy ^a	58/3900	53/3882	2.79	2.60	+0.19	1.07 (0.73, 1.59)
Hot Flashes	397	241	14.82	9.09	+5.73	1.63 (1.39, 1.92)
Leg Cramps	483	334	18.03	12.60	+5.43	1.43 (1.24, 1.65)
Peripheral edema	706	583	26.36	22.00	+4.36	1.20 (1.07, 1.34)
Cholelithiasis ^c	168/4144	131/4111	7.83	6.20	+1.63	1.26 (1.00, 1.60)

Abbreviations: IR = Incidence Rate per 1000 Patient-years.

^a Only patients with an intact uterus were considered for the denominator (raloxifene denominator = 3900, placebo denominator = 3882).

^b Only patients with at least one ovary were considered for the denominator (raloxifene denominator = 4559, placebo denominator = 4606).

^c Only patients with an intact gallbladder at baseline (raloxifene n=4144, total person-years of follow-up=21467; placebo n=4111, total person-years of follow-up=21136).

Appears This Way
On Original

3.1.3 Study H3S-MC-GGG (MORE)

3.1.3.1 Objective

The primary objectives were to assess the effects of raloxifene treatment, compared with placebo, on the incidences of new vertebral fractures, lumbar spine and femoral neck bone mineral density (BMD), and safety.

Assessment of the effect of raloxifene on incidence of all breast cancer was a secondary safety endpoint.

3.1.3.2 Study Design

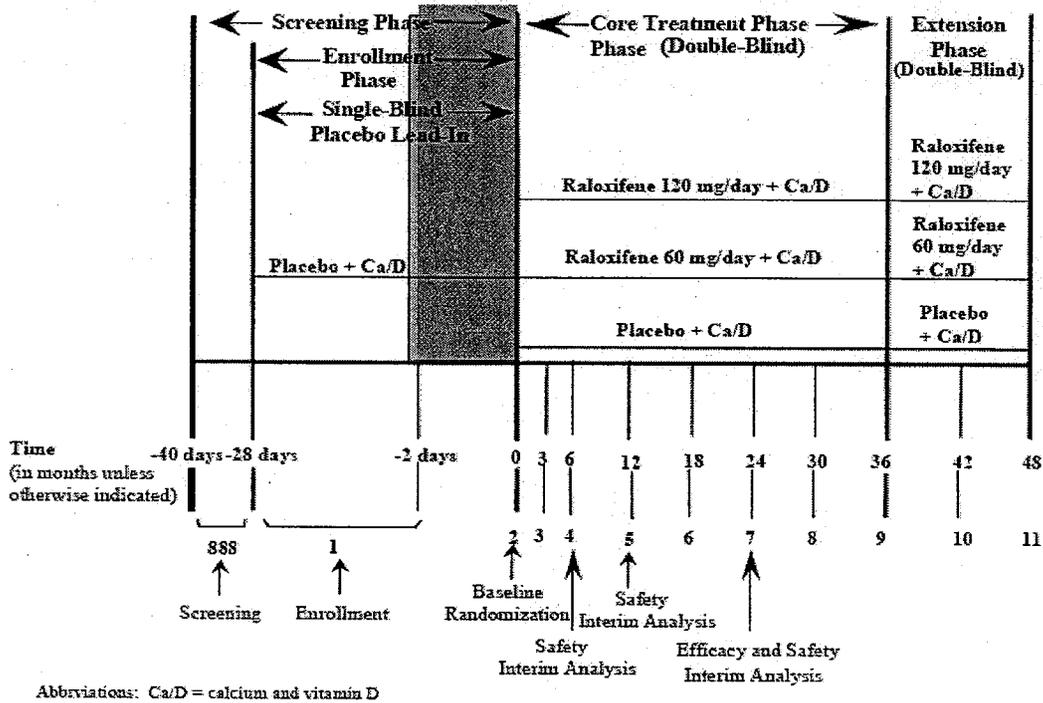
MORE was a randomized, double-blind, placebo-controlled, multinational study conducted in postmenopausal women with osteoporosis. The 7705 patients enrolled in the study were randomized to one of three treatment groups: placebo (N=2576), raloxifene HCl 60 mg/day (N=2557), or raloxifene HCl 120 mg/day (N=2572).

Women up to 80 years of age, and who were at least 2 years postmenopausal and had osteoporosis defined as lumbar spine or femoral BMD at least 2.5 standard deviations (SDs) below the mean for normal premenopausal women or at least one moderate or two mild vertebral fractures, were eligible to enroll. Patients with a known history of breast cancer were not eligible to enroll. However, patients were not enrolled based on any increased risk for developing breast cancer. Other exclusion criteria included history of deep vein thrombosis (DVT), thromboembolic disorders, or cerebrovascular accident within the past 10 years; abnormal uterine bleeding; or chronic liver disease.

The study consisted of a 36-month core treatment phase and a 12-month extension phase. All patients received supplemental calcium (500 mg/day) and vitamin D (400-600 IU/day) for the duration of the study. Concomitant use of other osteoporosis medications, including bisphosphonates, was allowed as clinically indicated during the 12-month extension phase. Study drug was permanently discontinued when a participant was unblinded or diagnosed with breast cancer or VTE. Figure 3.1.3.2.1 illustrates the study design for MORE.

Appears This Way
On Original

Figure 3.1.3.2.1 Study design for study MORE



Bilateral mammograms or ultrasound (if patient refused mammogram) were required at baseline (within 3 months before randomization) and after 2, 3, and 4 years of treatment; mammograms were optional after 1 year of treatment. All investigator-reported cases of breast cancer were reviewed and adjudicated by a board of physicians specialized in breast cancer who was blinded to patient treatment assignment and who was not employed by Applicant. For each reported case of breast cancer, the adjudicators were presented with as much of the following information as was available to the Applicant: mammographic and other relevant radiological reports, mammographic films (originals or copies), ER status, and pathologic reports from biopsy and/or surgical specimens. For each investigator-reported breast cancer event, the adjudicators determined: 1) whether the case was invasive primary breast cancer; 2) what the ER status was; and 3) whether the cancer may have been preexisting (i.e., present on the baseline mammogram) or new (occurring after the baseline visit). Non-cancer cases were included in the adjudication process.

Statistical analyses of the adjudicated breast cancer data were not prospectively defined as an efficacy endpoint in the protocol.

Appears This Way
On Original

3.1.3.3 Efficacy Measures

The primary efficacy measures were the changes in lumbar spine and femoral neck BMD and the rate of newly occurring vertebral fractures.

The incidence of breast cancer was a secondary safety endpoint.

3.1.3.4 Sample Size Considerations

The primary comparisons (which drive sample size) in these studies were the difference in vertebral fracture rates between the raloxifene treatment groups and the placebo group. The sample size was not determined based on expected breast cancer incidence.

Assuming an average age of 65 years, the rate of osteoporotic vertebral fractures was estimated to be 24 fractures/1000 patient-years for the population of Substudy I, and 65 fractures/1000 patient-years for the population of Substudy II. Calcium and vitamin D supplementation was expected to further reduce this rate by approximately 12% to 15%. Under these assumptions and an assumed 40% reduction in vertebral fracture rates in the raloxifene treatment groups compared with placebo, a sample size of 2100 patients per arm (1400 from Substudy I and 700 from Substudy II) provided over 80% power to detect a significant treatment effect in the vertebral fracture rate, pooled across the two studies, at the 24-month interim analysis. This calculation assumed a 20% drop-out rate at 24 months, and a two-tailed 0.05 significance level. This sample size also provided over 90% power to detect a significant decrease in vertebral fracture rates, pooled across the two studies, at the end of the third year. Each study was powered at over 90% to detect a 2% increase in lumbar spine and femoral neck BMD at the end of 24 months, when comparing treated groups with placebo. Each study was also powered at over 80% to show a 40% reduction in vertebral and nonvertebral fractures combined at the end of 24 months.

3.1.3.5 Interim Analysis

Two planned interim analyses were conducted under the auspices of the DMB assigned to this study. A 12-month interim analysis was conducted that focused on general safety the tests of cognitive function. A 24-month interim analysis was conducted that concentrated on the primary efficacy endpoints (incident vertebral fracture rates and BMD [total lumbar spine and femoral neck]) in addition to safety measures. All analyses were conducted at the 0.05 level for statistical significance and the 0.20 level for trends. No statistical adjustments were made because, regardless of the efficacy results of the 24-month interim analysis, the study was planned to and did continue for the third year in a double-blind fashion with the original study design.

Results of the 12-month interim safety analysis were discussed by the DMB in a conference call on December 5, 1997 which was attended by 9 DMB (5 members were employed by the Applicant). The main recommendation of the board was that there were no overriding safety concerns which

needed to be explored prior to the 24-month interim analysis which would address the safety and efficacy of raloxifene in the treatment population.

Results of the 24-month interim analysis were discussed by the DMB on March 5, 1998. This meeting was attended by 9 data monitoring board members (5 members were employed by the Applicant), and two additional study statisticians (also employed by the Applicant) who helped prepare the report. The unanimous recommendation of the board was that the 24-month data was sufficient to demonstrate efficacy in a treatment population. In addition, the board recommended that the study continued to the 36-month time point.

Three interim analyses were performed after approximately 6, 12, and 24 months of follow-up had been completed. The final analysis of the double-blind treatment phase was conducted after 36 months of follow-up and has previously been reported.

An additional interim analysis of safety data was added after the completion of the last patient's visit at 6 months. This recommendation was issued by the raloxifene data monitoring board after reviewing the 6-month interim data of three other raloxifene studies in the prevention of osteoporosis (Studies H3S-MC-GGGF, -GGGG, and -GGGH). While the data monitoring board had no safety concerns with these studies and unanimously recommended their continuation, the data monitoring board also advised making the safety review process more uniform across the large long-term studies of raloxifene. This implied the addition of a safety interim analysis at 6 months that was not provided in the previous protocol version.

3.1.3.6 Statistical Analysis Methods

The BMD measures were analyzed using an analysis of variance (ANOVA) model on the change from baseline to endpoint including terms for treatment and geographical region. The steering committee for the study defined the regions to be used in the analysis. The treatment-by-region interaction was removed from the model if it was not significant at a 0.10 level. Treatment effects were tested at a 0.05 level in each of the studies. Significance of pairwise comparisons of raloxifene groups with placebo depended on the overall significance.

The primary fracture analyses compared the fracture rates of the three treatment groups, where fracture rate was defined as the total number of new fractures divided by the total time in the studies (up to the last visit). The fracture rates were analyzed using a weighted ANOVA model incorporating the effects of treatment, region, and substudy (for the pooled analyses). The individual fracture rates may have been transformed to stabilize the variance and approximate normality (for instance, a square root transformation in the case of Poisson fractures). The method proposed by Box and Cox (1964) could find such a transformation, if necessary. Each individual's transformed rate would then be weighted by her time in a substudy to reduce variance. Because of the low fracture rate, it may have been necessary to pool sites into geographic regions to assess treatment by site interaction in the

ANOVA model. If this interaction was not significant at a 0.10 level, it would be excluded from the model. If an interaction was significant, the nature of the interaction was explored descriptively. A secondary analysis on the fracture data compared the proportion of patients with at least one new fracture in each active treatment group to placebo using Pearson's chi-square test.

There were no pre-specified statistical analysis plan for the breast cancer endpoint in the protocol and type I error rate allocated for this analysis.

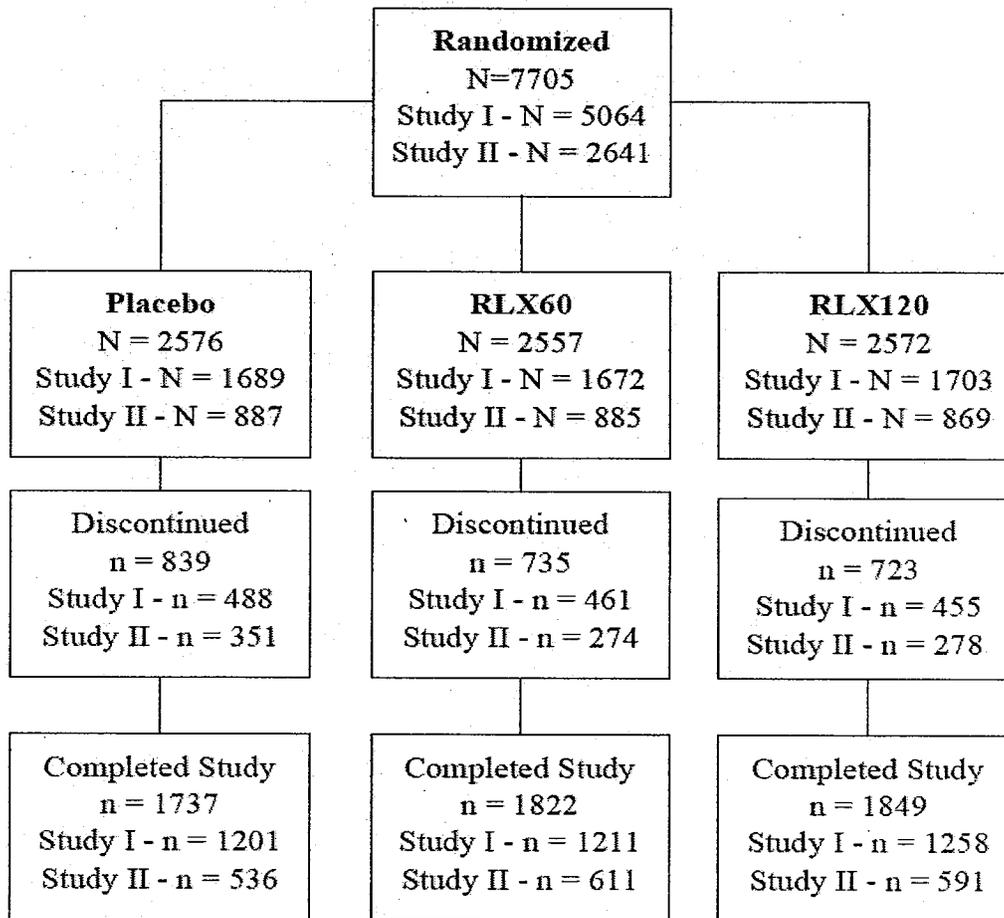
3.1.3.7 Applicant's Results and Statistical Reviewer's Findings/ Comments

3.1.3.7.1 Study Population

A total of 7705 patients were included in this final 48-month analysis. Of these 7705 patients, 2576 were randomly assigned to placebo, 2557 to raloxifene hydrochloride (HCl) 60 mg/day, and 2572 to raloxifene HCl 120 mg/day (Figure 3.1.3.7.1.1).

Appears This Way
On Original

Figure 3.1.3.7.1.1 Patient disposition



Abbreviations: N = number randomized; n = number observed.

3.1.3.7.2 Demographic and Baseline Characteristics

The demographic and other baseline characteristics of the patients did not differ significantly among the three treatment groups at baseline, with the exception of height. The maximum difference among the three treatment groups in mean height was 0.45 cm. The three groups were similar with respect to family history of breast cancer in the patient's first-generation family (mother, sisters, daughters).

Reviewer's Comments

P-values presented in tables below are descriptive purpose only which are not adjusted for multiple comparisons.

**Table 3.1.3.7.2.1 Patient Demographics Comparison of Treatment Groups at Baseline
All Randomized Patients**

Variable	PLACEBO (N=2576)	RLX060 (N=2557)	RLX120 (N=2572)	Total (N=7705)	p-Value
ORIGIN					
No. Patients	2576	2557	2572	7705	.228*
African Descent	6 (0.2)	6 (0.2)	14 (0.5)	26 (0.3)	
Western Asian	6 (0.2)	1 (0.0)	4 (0.2)	11 (0.1)	
Caucasian	2465 (95.7)	2455 (96.0)	2452 (95.3)	7372 (95.7)	
East/Southeast A	48 (1.9)	41 (1.6)	48 (1.9)	137 (1.8)	
Hispanic	44 (1.7)	48 (1.9)	41 (1.6)	133 (1.7)	
Other	7 (0.3)	6 (0.2)	13 (0.5)	26 (0.3)	
AGE: (yrs)					
No. Patients	2576	2557	2572	7705	.337**
Mean	66.60	66.48	66.31	66.47	
Median	66.92	66.86	66.73	66.85	
Standard Dev.	7.07	6.99	7.12	7.06	
Minimum	35.68	31.08	35.99	31.08	
Maximum	80.96	80.94	80.91	80.96	
HEIGHT: (cm) (VISIT: 1)					
No. Patients	2575	2557	2571	7703	.020**
Mean	158.95	158.92	159.38	159.08	
Median	159.00	159.00	159.51	159.10	
Standard Dev.	6.57	6.60	6.68	6.62	
Minimum	133.00	127.00	123.95	123.95	
Maximum	185.00	192.20	178.00	192.20	
Unspecified	0	0	1	1	
WEIGHT: (kg) (VISIT: 1)					
No. Patients	2575	2556	2572	7703	.364**
Mean	63.64	63.58	63.96	63.73	
Median	62.88	62.40	63.00	62.88	
Standard Dev.	10.52	10.35	10.73	10.53	
Minimum	33.82	34.00	35.30	33.82	
Maximum	131.21	111.00	130.75	131.21	
Unspecified	0	1	0	1	
BMI: (kg/m²) (VISIT: 2)					
No. Patients	2575	2557	2571	7703	.989**
Mean	25.24	25.23	25.22	25.23	
Median	24.82	24.66	24.78	24.77	
Standard Dev.	3.99	4.02	4.02	4.01	
Minimum	13.54	14.22	14.45	13.54	
Maximum	51.59	43.16	49.56	51.59	
Unspecified	1	0	1	2	
CURRENT SMOKER (VISIT: 2)					
No. Patients	2576	2557	2572	7705	.915*
No	2124 (83.5)	2102 (83.1)	2112 (83.2)	6338 (83.3)	
Yes	420 (16.5)	429 (16.9)	425 (16.8)	1274 (16.7)	
Unspecified	32	26	35	93	

ALCOHOL > 3 DRINKS WKLY (VISIT: 2)					
No. Patients	2576	2557	2572	7705	.606*
No	2132 (82.8)	2089 (81.7)	2134 (83.0)	6355 (82.5)	
Unknown	4 (0.2)	2 (0.1)	2 (0.1)	8 (0.1)	
Yes	440 (17.1)	466 (18.2)	436 (17.0)	1342 (17.4)	
YEARS PMP (VISIT: 1)					
No. Patients	2576	2557	2572	7705	.262**
Mean	18.89	18.76	18.51	18.72	
Median	19.00	19.00	18.00	19.00	
Standard Dev.	8.48	8.51	8.30	8.43	
Minimum	2.00	2.00	2.00	2.00	
Maximum	54.00	67.00	48.00	67.00	
FAM. HIST. OF OSTPRS (VISIT: 1)					
No. Patients	2576	2557	2572	7705	.930*
No	1595 (61.9)	1561 (61.0)	1571 (61.1)	4727 (61.3)	
Unknown	299 (11.6)	304 (11.9)	295 (11.5)	898 (11.7)	
Yes	682 (26.5)	692 (27.1)	706 (27.4)	2080 (27.0)	
FAM. HIST. OF BREAST CANCER (VISIT: 1)					
No. Patients	2576	2557	2572	7705	.814*
No	2196 (85.2)	2190 (85.6)	2183 (84.9)	6569 (85.3)	
Unknown	67 (2.6)	55 (2.2)	65 (2.5)	187 (2.4)	
Yes	313 (12.2)	312 (12.2)	324 (12.6)	949 (12.3)	
HYSTERECTOMY (VISIT: 1)					
No. Patients	2576	2557	2572	7705	.252*
No	1999 (77.6)	1950 (76.3)	2010 (78.1)	5959 (77.3)	
Yes	577 (22.4)	607 (23.7)	562 (21.9)	1746 (22.7)	
TYPE OF HYSTERECTOMY (VISIT: 1)					
No. Patients	2576	2557	2572	7705	.968*
Unknown	47 (8.1)	46 (7.6)	43 (7.7)	136 (7.8)	
Uterus, 0-1 Ovary	278 (48.2)	305 (50.2)	277 (49.3)	860 (49.3)	
Uterus, 2 Ovaries	252 (43.7)	256 (42.2)	242 (43.1)	750 (43.0)	
Unspecified	1999	1950	2010	5959	
PREV USE OF HRT (VISIT: 1)					
No. Patients	2576	2557	2572	7705	.567*
No	1833 (71.2)	1785 (69.8)	1829 (71.1)	5447 (70.7)	
Unknown	5 (0.2)	10 (0.4)	8 (0.3)	23 (0.3)	
Yes	738 (28.6)	762 (29.8)	735 (28.6)	2235 (29.0)	
PREV USE OF THIAZ DIURETICS (VISIT: 1)					
No. Patients	2576	2557	2572	7705	.174*
No	2241 (87.0)	2224 (87.0)	2249 (87.4)	6714 (87.1)	
Unknown	24 (0.9)	14 (0.5)	29 (1.1)	67 (0.9)	
Yes	311 (12.1)	319 (12.5)	294 (11.4)	924 (12.0)	
PREV USE OF SYSTEMIC FLUORIDES (VISIT: 1)					
No. Patients	2576	2557	2572	7705	.847*
No	2531 (98.3)	2506 (98.0)	2523 (98.1)	7560 (98.1)	
Unknown	4 (0.2)	4 (0.2)	2 (0.1)	10 (0.1)	
Yes	41 (1.6)	47 (1.8)	47 (1.8)	135 (1.8)	
PREV USE OF BISPHOSPHONATES (VISIT: 1)					
No. Patients	2576	2557	2572	7705	.072*
No	2522 (97.9)	2482 (97.1)	2504 (97.4)	7508 (97.4)	
Unknown	1 (0.0)	7 (0.3)	2 (0.1)	10 (0.1)	
Yes	53 (2.1)	68 (2.7)	66 (2.6)	187 (2.4)	

MARITAL STATUS (VISIT: 2)					
No. Patients	2576	2557	2572	7705	.599*
Divorced	241 (9.4)	234 (9.2)	249 (9.7)	724 (9.4)	
Married	1522 (59.3)	1543 (60.5)	1549 (60.4)	4614 (60.1)	
Never Married	137 (5.3)	139 (5.5)	125 (4.9)	401 (5.2)	
Separated	50 (1.9)	31 (1.2)	43 (1.7)	124 (1.6)	
Single but livin	0	0	1 (0.0)	1 (0.0)	
Widowed	615 (24.0)	602 (23.6)	596 (23.3)	1813 (23.6)	
Unspecified	11	8	9	28	
YEARS OF EDUCATION (VISIT: 2)					
No. Patients	2546	2530	2547	7623	.522**
Mean	11.82	11.78	11.90	11.84	
Median	12.00	12.00	12.00	12.00	
Standard Dev.	3.89	3.92	3.96	3.92	
Minimum	0.00	0.00	0.00	0.00	
Maximum	26.00	25.00	40.00	40.00	
Unspecified	30	27	25	82	
PRIOR AWARENESS OF OSTEOPOROSIS (VISIT: 2)					
No. Patients	2576	2557	2572	7705	.511*
Yes	980 (38.0)	937 (36.6)	945 (36.7)	2862 (37.1)	
No	1596 (62.0)	1620 (63.4)	1627 (63.3)	4843 (62.9)	

The MORE study randomized a total of 7705 postmenopausal women with osteoporosis (median age, 66.9 years) to treatment with placebo (N=2576), raloxifene HCl 60 mg/day (N=2557), or raloxifene HCl 120 mg/day (N=2572). MORE treatment groups appeared to be balanced with respect to baseline demographic characteristics. In MORE, 96.4% of patients were treatment compliant (defined as $\geq 70\%$ of study drug taken). Median follow-up was 47.4 months.

The population of women in this study was not selected based on a high risk of breast cancer. Nevertheless, baseline breast images were collected and classified as either normal or abnormal by the investigator. Abnormal breast images were then further classified by the investigator as either clinically relevant or not clinically relevant. At baseline, there appeared no differences in the proportion of patients with normal, abnormal and not clinically relevant, or abnormal and clinically relevant breast imaging among treatment arms (Table 3.1.3.7.2.2).

Appears This Way
On Original

**Table 3.1.3.7.2.2 Breast Imaging Results
All Randomly Assigned Patients**

	Placebo (N=2576)	RLX060 (N=2557)	RLX120 (N=2572)	p-value ^a
Baseline Breast Imaging Results^b				
Normal	1864 (72.4%)	1864 (72.9%)	1904 (74.0%)	0.387
Abnormal, Not Clinically Relevant	671 (26.1%)	653 (25.5%)	628 (24.4%)	0.388
Abnormal, Clinically Relevant	38 (1.5%)	40 (1.6%)	39 (1.5%)	0.966
Any Abnormal Result	709 (27.5%)	693 (27.1%)	667 (25.9%)	0.411

^a Chi-square test for total count ≥ 10 ; Fisher's Exact test for total counts 5 through 9.

^b Patients with more than one baseline breast image were classified according to their most severe result.
Abbreviations: N = number of randomly assigned patients; RLX060 = raloxifene 60 mg/day; RLX120 = raloxifene 120 mg/day.

Mammograms were required for all patients at baseline and at the 24-, 36-, and 48-month visits; mammograms were optional at the 12-month visit. If mammography was not acceptable to a patient, ultrasonography of the breast was performed instead, although patients were encouraged to undergo mammography.

Table 3.1.3.7.2.3 shows the number of eligible patients who underwent breast imaging at the baseline, 12-, 24-, 36-, and 48-month visits. For each visit interval, a patient was defined to be "eligible" for breast imaging if that patient was continuing in the study at the beginning of the visit interval (for example, a patient who had not discontinued by the 12-month visit was considered eligible for a 24-month breast image).

Appears This Way
On Original

**Table 3.1.3.7.2.3 Distribution of Breast Imaging Results
All Randomly Assigned Patients**

	Placebo (N=2576)	RLX060 (N=2557)	RLX120 (N=2572)	p-value ^a
Baseline Visit				
Eligible Patients ^b	2576	2557	2572	---
Patients With Breast Images	2574	2557	2571	0.778
Mammogram ^c	2500	2486	2496	---
Sonogram Only	74	71	75	0.951
12-Month Visit^d				
Eligible Patients ^b	2576	2557	2572	---
Patients With Breast Images	1249	1234	1244	0.986
Mammogram ^c	1198	1184	1184	---
Sonogram Only	51	50	60	0.574
24-Month Visit^d				
Eligible Patients ^b	2339	2283	2311	---
Patients With Breast Images	2176	2163	2171	0.052
Mammogram ^c	2113	2099	2121	---
Sonogram Only	63	64	50	0.328
36-Month Visit^d				
Eligible Patients ^b	2109	2124	2148	---
Patients With Breast Images	1894	1957	1986	0.004
Mammogram ^c	1839	1899	1923	---
Sonogram Only	55	58	63	0.809
48-Month Visit^d				
Eligible Patients ^b	1920	1964	1997	---
Patients With Breast Images	1779	1840	1861	0.444
Mammogram ^c	1731	1788	1811	---
Sonogram Only	48	52	50	0.961

^a The p-value for "Breast Images" compares the three treatment groups with respect to the number of eligible patients who had any breast imaging during a particular visit interval. The p-value for "Sonogram Only" compares the three treatment groups with respect to the number of patients with breast imaging who only had a breast sonogram during a particular visit interval. The p-value is calculated using Fisher's Exact test, since the proportion of patients without images is very small in some cases.

^b Eligible patients are defined as those who were continuing in the study at the beginning of the visit interval (e.g., a patient who had not discontinued by the 12-month visit was considered eligible for a 24-month breast image).

^c Patients who had multiple breast images during any visit interval were classified as having mammography if any of the images were mammograms, otherwise, they were classified as having only sonography.

^d Mammogram or sonogram results recorded at Visit 3 (3 months), Visit 4 (6 months), or Visit 5 (12 months) were considered 12-month breast images. Those results recorded at Visit 6 (18 months) or Visit 7 (24 months) were considered 24-month breast images. Those results recorded at Visit 8 (30 months) or Visit 9 (36 months) were considered 36-month breast images. Those results recorded at Visit 10 (42 months) or Visit 11 or 12 (48 months) were considered 48-month breast images.

Abbreviations: N = number of randomly assigned patients; RLX060 = raloxifene 60 mg/day, RLX120 = raloxifene 120 mg/day.

Breast imaging was performed in 99.97% of patients at baseline, and 48% of women elected to have an optional breast imaging procedure at the 12-month visit. At the 24-, 36-, and 48-month visits, 94%, 91%, and 93% of participants continuing in the study, respectively, had breast imaging procedures performed. For all randomly assigned patients, there were no differences among the three treatment groups in the number of patients who had breast imaging at baseline, 12, 24, and 48 months. Among those patients who underwent breast imaging, there was no difference among the three treatment groups in the number of patients who elected sonography instead of mammography at baseline or at any visit. Among those patients eligible for breast imaging at 36-months, there was a difference among the three treatment groups in the number of patients who had breast imaging.

For the analyses of breast imaging, patients with more than one post baseline breast image were classified according to their most "severe" result, with the result of "abnormal, clinically relevant" being classified as more severe than the result of "abnormal, not clinically relevant," which was classified as more severe than the result of "normal."

3.1.3.7.3 Applicant's Efficacy Analyses

Eighty-two cases of primary breast carcinoma were reported to the Applicant. One cancer (in a raloxifene-treated patient) was assigned a diagnosis date by the investigator that was 4 days prior to her randomization date; because this patient was randomly assigned to study drug, she was included in the analyses. Three breast carcinoma cases were reported after the end of the study and submitted for adjudication. One of these cases has since been adjudicated and is included in these analyses.

By 48 months, raloxifene use was associated with a reduction in the incidence of breast cancer. The reported incidence of invasive and noninvasive breast cancers in the pooled raloxifene group was reduced by 62% compared with placebo. This reduction in breast cancer risk was highly statistically significant (95% confidence interval 39% to 76%).

Of the 82 reported cases of breast carcinoma in MORE, 4 cases were excluded from the analyses. The following is a listing of the reasons for the exclusions:

- Case 081-6018 was adjudicated "metastatic adenocarcinoma of unknown primary";
- Case 085-6480 was adjudicated "No cancer";
- Cases 086-7631 and 068-6961 had not been adjudicated as of 22 May 2002 because required adjudication documents had not been submitted to the applicant by the investigative sites.

The estimated incidence rates of breast cancer and invasive breast cancer are presented in Table 3.1.3.7.3.1. Annual incidence rates of breast cancer and invasive breast cancer were lower in patients assigned to raloxifene than in patients assigned to placebo. Neither the incidence of breast cancer nor invasive breast cancer was significantly different between the raloxifene 60- and 120-mg treatment groups. Because treatment effects in these two groups were similar the raloxifene groups are pooled for all further analyses.

**Table 3.1.3.7.3.1 Estimated Annual Incidence Rates for
Breast Cancer and Invasive Breast Cancer
All Randomly Assigned Patients**

Population	Therapy	No. Randomized	Cases	Patient-years of Follow-up		Rate (per 1000)
Breast Cancer	Placebo	2576	44	8716		5.05
	RLX060		17	8756		1.94
	RLX120	5129	17	8868		1.92
	Pooled Ralox		34	17624		1.93
Invasive Breast Cancer	Placebo	2576	38	8718		4.36
	RLX060		11	8756		1.26
	RLX120	5129	10	8869		1.13
	Pooled Ralox		21	17625		1.19

Abbreviations: No. = number; Ralox = raloxifene.

The results shown in Table 3.1.3.7.3.2 demonstrated a 62% reduction in breast cancer incidence for raloxifene-treated women compared with the placebo group. Considering only invasive tumors, the percent reduction was 73%. For the subset of subjects who presented with invasive ER+ tumors, raloxifene demonstrated an 83% reduction in the incidence of breast cancer compared with placebo.

**Table 3.1.3.7.3.2 Breast Cancer Relative Risk Analysis of all Cases
All Randomly Assigned Patients**

Category	Number of Cases		Rate per 1000 Women		Risk Ratio (95% CI)
	Placebo	Raloxifene	Placebo	Raloxifene	
All cases	44	34	5.05	1.93	0.38 (0.24, 0.61)
Invasive cases	38	21	4.36	1.19	0.27 (0.15, 0.48)
ER-positive cases	31	13	3.56	0.74	0.21 (0.10, 0.41)
Invasive ER-positive cases	29	10	3.33	0.57	0.17 (0.07, 0.36)
ER-negative cases	4	10	0.46	0.57	1.24 (0.36, 5.40)
Cases of unknown ER status	9	11	1.03	0.62	0.60 (0.23, 1.65)

Abbreviations: ER = estrogen receptor; CI = confidence interval.

Table 3.1.3.7.3.3 presents the invasive breast cancer results for MORE for placebo and raloxifene HC1 60 mg/day. Compared with placebo, raloxifene showed a statistically significant 71% decrease in the incidence of invasive breast cancers.

**Table 3.1.3.7.3.3 Invasive Breast Cancer Relative Risk Analysis
Placebo vs. Raloxifene HCl 60 mg/day**

Breast Cancer Category	MORE (N=5133) ^a					
	Placebo N=2576		Raloxifene N=2557		HR (95% CI)	p-value
	n	IR	n	IR		
Invasive	38	4.36	11	1.26	0.29 (0.15, 0.56)	<0.001
ER- positive	29	3.33	6	0.69	0.20 (0.08, 0.49)	<0.001
ER- negative	4	0.46	5	0.57	1.23 (0.33, 4.60)	0.752
ER unknown	5	0.57	0	0	N/A	N/A

Abbreviations: CI = confidence interval; ER = estrogen receptor; HR = hazard ratio; IR = incidence rate per 1000 patient-years; n = number of patients with breast cancer events; N = number of patients analyzed; N/A = Not Applicable;

^a Patients randomized in MORE to either placebo or raloxifene HCl 60 mg/day. Breast cancers reported from randomization in MORE to end of MORE (48 months) are presented.

Of 61 total breast cancer events reported in MORE, 8 (13%) were classified as noninvasive. Of these 8 cases (all of which were DCIS), 5 and 3 occurred within the placebo and raloxifene groups, respectively. Invasiveness status could not be ascertained for 4 of the 61 adjudicated breast cancers (placebo, 1; raloxifene, 3).

Reviewer's Comments

1. For MORE trial, the incidence of invasive breast cancer was a secondary safety endpoint. The trial was not designed to demonstrate the raloxifene effect on the incidence of invasive breast cancer. Although p-value of the log-rank test on the invasive breast cancer endpoint was less than 0.05, this statistical test was not planned, especially, not adjusted for multiple secondary and safety endpoints. One should realize that for a trial having many secondary and safety endpoints, one can always find some "significant" endpoints. Such results may really provide true discoveries, but one should not rely on this type of results because the significance level of the tests associated with this type of results is completely uncontrolled.

2. The analyses for the incidence of invasive breast cancer by invasiveness and estrogen receptor status are presented below.

Appears This Way
On Original

Breast Cancer Category*	Placebo N=2576	Raloxifene 60 mg N=2557	Relative Risk (95% CI)
	n (IR)	n (IR)	
Invasive	38 (4.36)	11 (1.26)	0.29 (0.13, 0.58)
ER Positive	29 (3.33)	6 (0.69)	0.21 (0.07, 0.50)
ER Negative	4 (0.46)	5 (0.57)	1.25 (0.27, 6.28)
ER Unknown	5 (0.57)	0	N/A
Non-invasive	5 (0.57)	3 (0.34)	0.60 (0.09, 3.07)
DCIS	5 (0.57)	3 (0.34)	0.60 (0.09, 3.07)
LCIS	0 (0.00)	0 (0.00)	NA
Invasiveness unknown	1 (0.11)	3 (0.34)	2.99 (0.24, 1.56)
All	44 (5.05)	17 (1.94)	0.38 (0.21, 0.69)

*Patients randomized in MORE to either placebo or raloxifene HCl 60 mg/day. Breast cancers reported from randomizations in MORE (48 months) are presented.

3. The analyses of important safety outcomes are presented below.

Appears This Way
On Original

Events ^a	Raloxifene 2,557	Placebo 2,576	Raloxifene IR	Placebo IR	Absolute Risk Difference	RR (95% CI)
Clinical vertebral fracture	62	107	7.08	12.27	-5.19	0.58 (0.42, 0.80)
Death	64/5129	36	3.63	4.13	-0.50	0.88 (0.58, 1.36)
Death due to Stroke	9/5129	6	0.51	0.69	-0.18	0.74 (0.23, 2.52)
Stroke	91/5129	56	5.16	6.42	-1.26	0.80 (0.57, 1.14)
Deep vein thrombosis	44/5129	8	2.50	0.92	+1.58	2.72 (1.27, 6.68)
Pulmonary embolism	22/5129	4	1.25	0.46	+0.79	2.72 (0.92, 10.85)
Endometrial and uterine cancer ^b	8/3960	5/1999	0.59	0.74	+0.15	0.80 (0.23, 3.10)
Ovarian Cancer	6/5129	6/1999	0.34	0.69	-0.35	0.49 (0.13, 1.84)
Hysterectomy ^b	40/3960	22/1999	2.93	3.24	-0.31	0.90 (0.52, 1.60)
Hot Flashes	512/5129	151	29.04	17.31	+11.73	1.68 (1.40, 2.03)
Leg Cramps	443/5129	150	25.13	17.20	+7.93	1.46 (1.21, 1.77)
Peripheral edema	340/5129	134	19.29	15.36	+3.93	1.26 (1.03, 1.55)
Cholelithiasis ^c	93/5129	45	5.28	5.16	+0.12	1.02 (0.71, 1.50)

Abbreviations: IR = Incidence Rate per 1000 Patient-years; RR=Relative risk.

^a Breast cancer and clinical vertebral fracture events are for the raloxifene 60 mg/day arm only; denominator = 2557. For the safety events of death, death due to stroke, stroke, deep vein thrombosis, pulmonary embolism, and ovarian cancer, the raloxifene 60 and 120 mg/day arms were pooled to have the greatest opportunity to detect safety signals; thus, the denominator for these events is 5129.

^bOnly patients with a uterus at baseline (pooled raloxifene n=3960, total person-years of follow-up=13659.16; placebo n=1999, total person-years of follow-up=6791.41). "Hysterectomy" included MedDRA Preferred Terms of "Hysterectomy," "Hysterosalpingo-oophorectomy," and "radical hysterectomy."

^cGallbladder status at baseline was not ascertained in the MORE trial.

Bone Efficacy

New Vertebral Fractures

Table 3.1.3.7.3.4 summarizes the proportion of women and relative risk for having one or more new incident adjudicated, vertebral fractures during the trial for each dose of raloxifene and pooled raloxifene doses compared with placebo, along with 95% confidence intervals, for each substudy and for the entire study population. Each dose of raloxifene statistically significantly decreased the proportion of women with adjudicated, new incident vertebral fractures in each substudy and overall compared with the placebo group. Overall, there was a 36% reduction ($p<0.001$) in such fractures in the raloxifene 60-mg group and a 43% reduction ($p<0.001$) in the raloxifene 120-mg group

compared with the placebo group. Overall, there was not a statistically significant difference between the two raloxifene groups in the proportion of patients with at least one new vertebral fracture.

**Table 3.1.3.7.3.4 New Incident Vertebral Fracture Results Overall and by Substudy
All Randomly Assigned Patients**

	Placebo	RLX060	RLX120	Pooled RLX Doses
Substudy I	n=1521	n=1492	n=1512	n=3004
Number of patients with ≥ 1 incident fracture (%)	97 (6.4%)	51 (3.4%)	57 (3.8%)	108 (3.6%)
Relative risk (95% CI) compared with placebo		0.54 (0.38, 0.75)	0.59 (0.43, 0.81)	0.56 (0.43, 0.74)
Pairwise comparison with placebo		p<0.001	p=0.001	p<0.001
Pairwise comparison with RLX060			p=0.605	
Substudy II	n=771	n=767	n=765	n=1532
Number of patients with ≥ 1 incident fracture (%)	191 (24.8%)	130 (16.9%)	107 (14.0%)	237 (15.5%)
Relative risk (95% CI) compared with placebo		0.68 (0.56, 0.83)	0.56 (0.46, 0.70)	0.62 (0.53, 0.74)
Pairwise comparison with placebo		p<0.001	p<0.001	p<0.001
Pairwise comparison with RLX060			p=0.109	
Pooled Substudies	n=2292	n=2259	n=2277	n=4536
Number of patients with ≥ 1 incident fracture (%)	288 (12.6%)	181 (8.0%)	164 (7.2%)	345 (7.6%)
Relative risk (95% CI) compared with placebo		0.64 (0.53, 0.76)	0.57 (0.48, 0.69)	0.61 (0.52, 0.70)
Pairwise comparison with placebo		p<0.001	p<0.001	p<0.001
Pairwise comparison with RLX060			p=0.304	

Abbreviations: RLX = raloxifene; RLX060 = raloxifene 60 mg/day; RLX120 = raloxifene 120 mg/day; CI = confidence interval;
n = number of patients with evaluable radiographs at baseline and endpoint.

Bone Mineral Density (BMD)

At 48 months for every skeletal site measured, the mean percentage change in BMD from baseline to endpoint in each raloxifene group was significantly greater than in the placebo group ($p<0.001$ for the pairwise comparisons between each raloxifene group and the placebo group) (Table 3.1.3.7.3.5). These results were similar to the result observed at 36 months.

Appears This Way
On Original

**Table 3.1.3.7.3.5 Summary of Percentage Change in BMD From Baseline to Endpoint
All Randomly Assigned Patients**

Test		Treatment Group			Overall p-value*
		Placebo	RLX060	RLX120	
Lumbar Spine BMD	Mean Baseline	0.814	0.817	0.814	0.555
	Mean Change	0.006	0.027 ^c	0.026 ^c	0.000
	Mean Percentage Change	0.740	3.293 ^c	3.268 ^c	0.000
Femoral Neck BMD	Mean Baseline	0.622	0.625	0.621	0.310
	Mean Change	-0.009	0.004 ^c	0.006 ^c	0.000
	Mean Percentage Change	-1.296	0.797 ^c	0.970 ^c	0.000
Trochanter BMD	Mean Baseline	0.556	0.558	0.552 ^d	0.051
	Mean Change	-0.005	0.007 ^c	0.008 ^c	0.000
	Mean Percentage Change	-0.835	1.284 ^c	1.683 ^c	0.000
Inter-Trochanter BMD	Mean Baseline	0.837	0.838	0.836	0.921
	Mean Change	-0.010	0.006 ^c	0.008 ^c	0.000
	Mean Percentage Change	-1.148	0.746 ^c	1.060 ^c	0.000
Wards Triangle BMD	Mean Baseline	0.459	0.462	0.460	0.456
	Mean Change	-0.015	0.000 ^c	0.001 ^c	0.000
	Mean Percentage Change	-2.911	0.391 ^c	0.590 ^c	0.000
Radial Ultradistal BMD	Mean Baseline	0.309	0.309	0.306	0.441
	Mean Change	-0.006	0.002 ^c	0.000 ^c	0.000
	Mean Percentage Change	-1.474	1.211 ^c	0.655 ^c	0.000
Radial Distal 1/3 BMD	Mean Baseline	0.541	0.543	0.540	0.807
	Mean Change	-0.005	0.001 ^c	0.001 ^c	0.000
	Mean Percentage Change	-0.809	0.336 ^c	0.212 ^c	0.000
Whole Body BMD	Mean Baseline	0.893	0.891	0.888	0.606
	Mean Change	-0.004	0.006 ^c	0.005 ^c	0.000
	Mean Percentage Change	-0.474	0.762 ^c	0.616 ^c	0.000

* Using ANOVA with Unranked data
a - pairwise comparison statistically significantly (p < 0.05) different from placebo
b - pairwise comparison statistically significantly (p < 0.01) different from placebo
c - pairwise comparison statistically significantly (p < 0.001) different from placebo
d - pairwise comparison of RLX060 statistically significantly (p < 0.05) different from RLX120

Reviewer's Comments

Based on these results, raloxifene was approved for the treatment of osteoporosis.

Appears This Way
On Original

3.1.4 Study H3S-MC-GGJY (CORE)

3.1.4.1 Objective

The primary objective was to test the hypothesis that a statistically significant reduction in the incidence of invasive breast cancer would occur in postmenopausal women with osteoporosis treated with raloxifene HCl 60 mg/day compared with placebo over a long-term period of observation.

The long-term observation period was defined by relationship to the patients' enrollment in Study MORE. The period of interest began 1 January 1999 (the date of the last breast cancer data analysis to support the osteoporosis treatment indication for marketing authorizations). Thus, the beginning of this period of observation in Study CORE corresponded to at least 3 years after the randomization of patients into MORE trial; this period of observation continued for approximately 8 years after the randomization of patients into MORE trial.

3.1.4.2 Study Design

CORE was a double-blind, placebo-controlled, multinational study that enrolled postmenopausal women with osteoporosis who had been randomized in MORE for an additional 4 years of follow-up. The primary objective of CORE was to compare the long-term effect of raloxifene HCl 60 mg/day versus placebo on the reduction in incidence of invasive breast cancer in postmenopausal women with osteoporosis. The secondary objectives were to assess the long-term effect of raloxifene HCl 60 mg/day on the incidence of invasive, ER-positive breast cancer and nonvertebral fractures in postmenopausal women with osteoporosis.

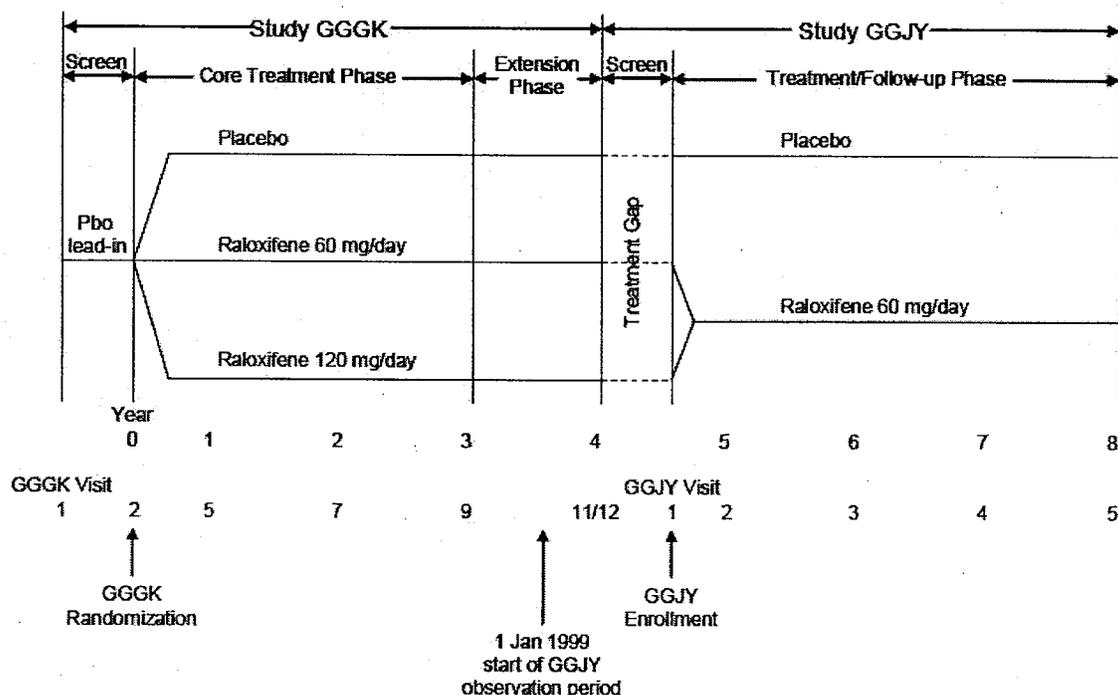
The selection of raloxifene HCl 60 mg/day as the only active treatment dose in CORE was based on the following:

- raloxifene HCl 60 mg/day had similar efficacy to raloxifene HCl 120 mg/day in MORE in terms of reduction in the incidence of breast cancer; and
- raloxifene HCl 60 mg/day had similar efficacy to raloxifene HCl 120 mg/day in MORE in terms of reduction in the incidence of new vertebral fractures.

The observation period for the primary analysis of the breast cancer endpoints was defined by relationship to the patients' enrollment in MORE. The observation period began on 01 January 1999, during the fourth year of MORE, and continued through the 4 years of CORE. The date of 01 January 1999 was chosen because that was the date of the last breast cancer data analysis to support the osteoporosis treatment indication. The start of the observation period (01 January 1999) was also the date at which the primary study endpoint changed from incidence of vertebral fractures (MORE) to incidence of invasive breast cancer (CORE). Thus, the CORE observation period began at least 3 years after the randomization of patients into MORE and continued for approximately 8 years after the randomization of patients into MORE.

Figure 3.1.4.2.1 shows the study design of CORE and its relationship to MORE. Of the 180 investigative sites that participated in MORE, 130 sites agreed to participate in CORE. Patients who were randomized in MORE and who were at the 130 investigative sites choosing to participate in CORE were invited to participate in CORE after their completion or discontinuation from MORE. All patients randomized in MORE at these 130 sites (N=6511) were eligible for CORE, and 4011 chose to enroll in CORE, comprising a population hereafter referred to as CORE enrollees.

Figure 3.1.4.2.1 Study design for Study CORE.



Here, GGGK=MORE, and GGJY=CORE.

As per CORE protocol, CORE enrollees were not re-randomized; instead, the randomization assignment from MORE was carried forward into CORE. Specifically, those CORE enrollees who had been randomized to raloxifene HCl 60 mg/day (n=1355) or 120 mg/day (n=1370) in MORE were assigned to receive raloxifene HCl 60 mg/day in CORE (n=2725); those who had been assigned to receive placebo in MORE continued on placebo in CORE (n=1286). Therefore, in CORE, approximately twice as many patients were assigned to receive raloxifene HCl 60 mg/day as were assigned to receive placebo.

Women randomized in MORE could enroll in CORE even if they were not allowed to take study medication or chose not to take study medication. CORE enrollees were not allowed to take study medication if they had a diagnosis of any malignancy considered to be estrogen-dependent (including

malignancies of the breast or uterus), had a history of VTE, or had a safety concern during MORE that necessitated unblinding of their treatment assignment. Of the CORE enrollees, 811 patients (268 [20.8%] in placebo and 543 [19.9%] in raloxifene) did not take study medication, either because they met one of the criteria above or because they chose not to.

Study drug was permanently discontinued when a participant was diagnosed with breast cancer or venous thromboembolism. Concomitant use of other osteoporosis medications, including bisphosphonates, calcitonin, or fluorides, was allowed during CORE.

Each patient's 5-year predicted risk of invasive breast cancer was calculated at baseline using the modified Gail model.

Bilateral mammograms were required at baseline (within 12 months before baseline) and every 2 years thereafter; clinical breast examinations were required at baseline and annually thereafter. All investigator-reported breast cancers were reviewed and adjudicated by a board of physicians specialized in breast cancer who were blinded to patient treatment assignment and who were not employed by Applicant. For each investigator-reported case of breast cancer, the adjudicators were provided with as much of the following information as was available to the Applicant: reports of all mammograms considered abnormal, other relevant radiological reports, ER status, and pathologic reports from biopsy and/or surgical specimens. Mammograms were defined as abnormal if the written report suggested that follow-up imaging procedures were required, if a lesion that required sampling was identified, or if the investigator deemed the mammogram to be clinically significant for other reasons. As per the protocol, noncancer cases were not included in the adjudication process; therefore, ascertainment bias was a possibility.

Breast cancer analyses for CORE were performed using time to first event methods. The primary analysis of CORE was performed using the primary analysis dataset, which included a subpopulation of MORE randomized patients. The CORE primary analysis dataset included:

- 1217 patients who did not enroll in CORE, but were still participating in MORE at the start of the CORE observation period (01 January 1999) and who contributed data for the CORE primary analysis until their completion of MORE, and
- 3996 CORE enrollees who had not been diagnosed with breast cancer as of the start of the CORE observation period (01 January 1999) for a total of 5213 patients (n=1703 for placebo and n=3510 for raloxifene).

As the start of the CORE observation period (01 January 1999) overlapped with the fourth year of MORE, breast cancers reported from 01 January 1999 to the end of the fourth year of MORE were included in the MORE analysis and also in the CORE primary analysis. In addition, 1684 patients in the CORE primary analysis dataset were assigned to raloxifene HCl 120 mg/day at MORE randomization and continued to take raloxifene HCl 120 mg/day from 01 January 1999 until the end of their participation in MORE.

Thus, to avoid double-counting of breast cancers reported from 01 January 1999 to the end of MORE in both the MORE and CORE analyses and to focus on breast cancers reported during the time that CORE enrollees were assigned to raloxifene HCl 60 mg/day, this presented the results from CORE enrollment (Visit 1) to the end of CORE (Visit 5), the period during which CORE enrollees were assigned to raloxifene HCl 60 mg/day or placebo.

From CORE enrollment (Visit 1) to end of CORE (Visit 5), CORE enrollees (N=4011) were assigned to either placebo or raloxifene HCl 60 mg/day. Of the 4011 CORE enrollees, 21 (12 in placebo and 9 in raloxifene) developed breast cancer prior to Visit 1 and, therefore, were excluded from the analysis of the breast cancer endpoints.

3.1.4.3 Efficacy Measures

The primary efficacy variable was the date of diagnosis of invasive breast cancer, as diagnosed via mammograms, breast examinations, and breast biopsies or the date the patient left the study without having experienced a diagnosis of breast cancer.

3.1.4.4 Sample Size Considerations

Study CORE was designed to provide follow-up data for as many women as possible who were enrolled in Study MORE. Power calculations were performed using various permutations of assumptions regarding the number of protocol completers, the relative risk of invasive breast cancer in raloxifene patients, and the true annual placebo rate of invasive breast cancer. It was estimated that a minimum of 2610 patients were needed to complete the study. This calculation assumed a true relative risk of invasive breast cancer of 0.24 in patients assigned to study medication, and that 3000 patients would enroll in the study, with 390 patients not on study medication. In reality, however, all patients who were willing to continue and were eligible from the MORE trial continued in this trial.

3.1.4.5 Interim Analysis

One planned interim analysis occurred after all participants had their 6-year visit and after sufficient time (approximately 3 months) had been allowed for the follow-up of suspicious mammograms. All cases reported from 1 January 1999 to the data cutoff date were included in the interim analysis. The denominator for the interim analysis was the same as for the final analysis of invasive breast cancer (primary objective). The interim analysis used ≤ 0.001 as the level of significance. Results from the interim analysis did not meet the predefined stopping criteria for outstanding efficacy (participants assigned to raloxifene did not have a statistically significantly reduced incidence of invasive breast cancer). The data monitoring board recommended the study continue as planned, with the exception that the observation period for patients in the PAD who did not enroll in Study CORE be changed to 1 January 1999 until the time of final follow-up contact in Study MORE. The final analysis took place after the 8th year of follow-up, with ≤ 0.0495 as the level of significance.

3.1.4.6 Statistical Analysis Methods

The primary analysis was a log-rank test. The primary analysis dataset (PAD) would be defined as all patients who were eligible for participation in CORE, and who had not been diagnosed with breast cancer as of January 1, 1999. This population was comprised of all surviving, breast cancer-free, former MORE patients at sites participating in CORE.

Reviewer's Comments

The primary analysis was modified in an amendment which was submitted to the Agency on March 23, 2001. In this amended plan, the log-rank test was defined as the primary analysis and replaced the original primary analysis of the Mantel Haenszel test.

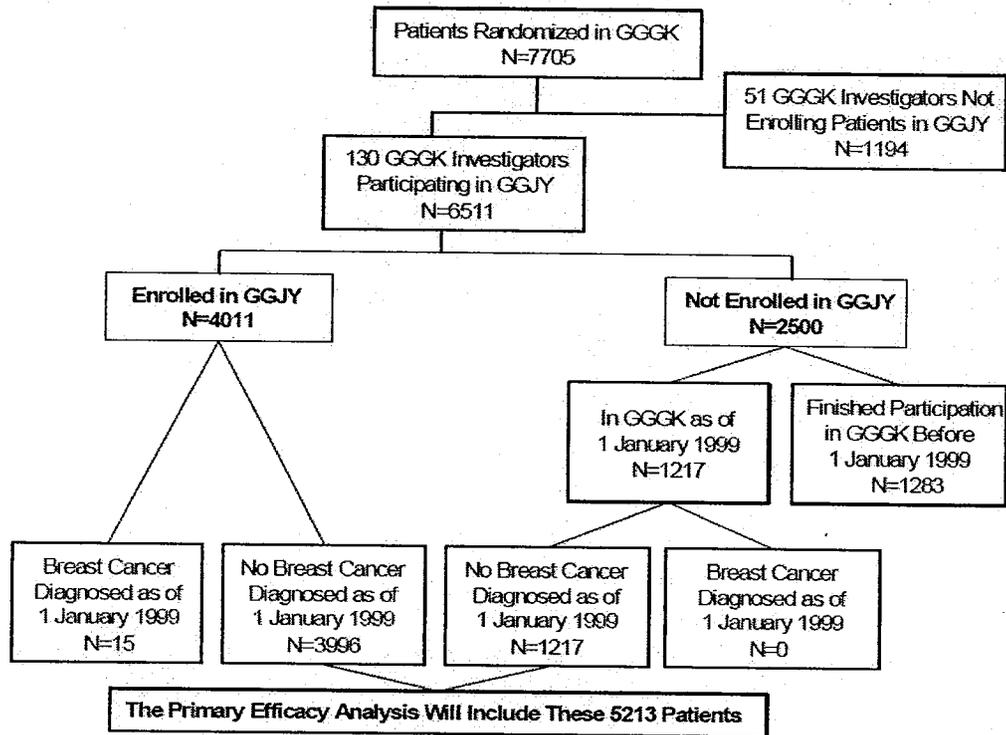
3.1.4.7 Applicant's Results and Statistical Reviewer's Findings/ Comments

3.1.4.7.1 Study Population

All patients randomly assigned in Study MORE were eligible for enrollment in Study CORE if their investigator elected to participate in CORE. Figure 3.1.4.7.1.1 illustrates the succession of patients from MORE into Study CORE. Sixty-three percent of eligible patients in Study MORE (n=6511) elected to enroll in Study CORE (n=4011).

Appears This Way
On Original

Figure 3.1.4.7.1.1 Patient succession from MORE into CORE.

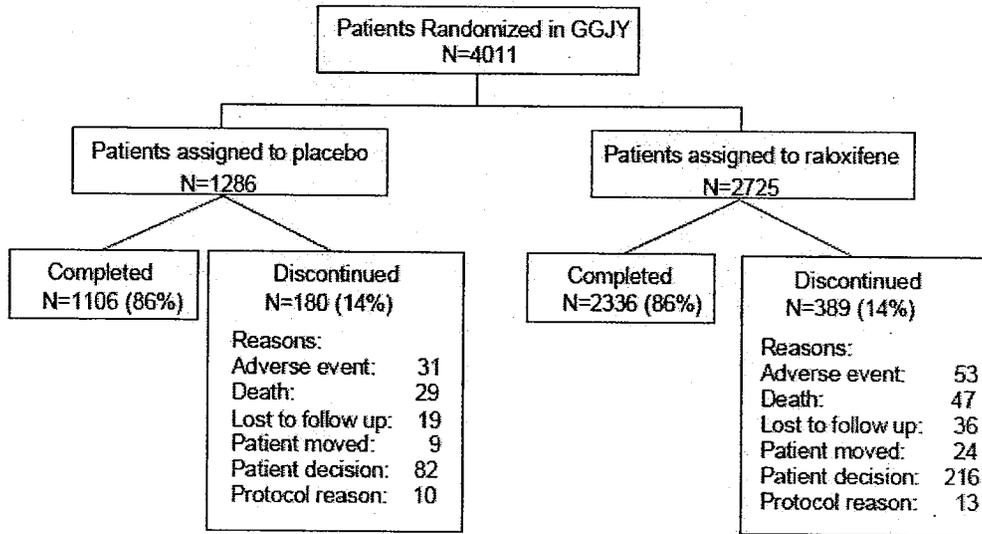


The primary analysis dataset (PAD) for the Study CORE breast cancer analyses was comprised of all Study MORE patients who were at investigative sites that participated in Study CORE. To reduce bias, data from all women at Study CORE sites, regardless of whether they enrolled in Study CORE, were included in the PAD. Patients in the PAD diagnosed with invasive breast cancer before 1 January 1999 (15 patients enrolled in Study CORE) were excluded from the primary analysis, which considered the time to first invasive breast cancer after this date. Thus, the primary analysis included 5213 patients, 3996 (77%) of who enrolled in Study CORE.

Figure 3.1.4.7.1.2 provides an overview of patient disposition for all enrolled patients in Study CORE. Of the 4011 CORE enrollees, 811 patients (268 [20.8%] in placebo and 543 [19.9%] in raloxifene) did not take study medication, either because they met one of the protocol-specified criteria (diagnosis of an estrogen-dependent malignancy, had a history of venous thromboembolic event (VTE), or had a safety concern during MORE that necessitated unblinding of their treatment assignment) or because they chose not to. All were included in the dataset for the analysis of breast cancer endpoints. In CORE, 55% of patients were treatment compliant (defined as $\geq 80\%$ of study drug taken). Of the patients enrolled in Study CORE, 86% in each treatment group completed the protocol and only 14% in each group discontinued the study. Of note, 2 patients in the placebo group

and 5 patients in the raloxifene group each had an adverse event but completed the protocol prior to discontinuation, and, thus, were included in the number of patients who completed the protocol.

Figure 3.1.4.7.1.2 Patient disposition



3.1.4.7.2 Demographic and Baseline Characteristics

Table 3.1.4.7.2.1 provides demographic and baseline characteristics.

Reviewer's Comments

P-values in tables below are for descriptive purpose only, and are not adjusted for multiple comparisons.

Appears This Way
On Original

Table 3.1.4.7.2.1 Patient Demographics at the Start of Study CORE
All Patients Enrolled in Study CORE

Variable	PLACEBO (N=1286)	RLX060 (N=2725)	Total (N=4011)	p-Value
ORIGIN				
No. Patients	1286	2725	4011	.060*
African Descent	1 (0.1)	2 (0.1)	3 (0.1)	
Western Asian	4 (0.3)	1 (0.0)	5 (0.1)	
Caucasian	1235 (96.0)	2622 (96.2)	3857 (96.2)	
East/Southeast A	26 (2.0)	39 (1.4)	65 (1.6)	
Hispanic	19 (1.5)	51 (1.9)	70 (1.7)	
Other	1 (0.1)	10 (0.4)	11 (0.3)	
AGE: (yrs)				
No. Patients	1286	2725	4011	.441**
Mean	70.91	70.73	70.79	
Median	71.01	70.95	70.97	
Standard Dev.	6.72	6.76	6.75	
Minimum	40.90	49.85	40.90	
Maximum	86.00	85.95	86.00	
HEIGHT: (cm) (VISIT: 1)				
No. Patients	1271	2708	3979	.962**
Mean	158.26	158.25	158.25	
Median	158.10	158.40	158.20	
Standard Dev.	6.66	6.68	6.67	
Minimum	138.00	137.50	137.50	
Maximum	177.40	190.90	190.90	
Unspecified	15	17	32	
WEIGHT: (kg) (VISIT: 1)				
No. Patients	1273	2710	3983	.408**
Mean	63.95	64.24	64.15	
Median	63.00	63.11	63.11	
Standard Dev.	10.61	10.55	10.57	
Minimum	29.74	34.50	29.74	
Maximum	108.40	132.00	132.00	
Unspecified	13	15	28	
BMI: (kg/m2) (VISIT: 1)				
No. Patients	1268	2707	3975	.357**
Mean	25.54	25.67	25.63	
Median	25.10	25.29	25.24	
Standard Dev.	4.10	4.04	4.06	
Minimum	14.52	15.13	14.52	
Maximum	44.29	50.42	50.42	
Unspecified	18	18	36	

Appears This Way
On Original

COUNTRY (VISIT: 1)				
No. Patients	1286	2725	4011	1.00*
Argentina	95 (7.4)	200 (7.3)	295 (7.4)	
Austria	3 (0.2)	9 (0.3)	12 (0.3)	
Australia	43 (3.3)	94 (3.4)	137 (3.4)	
Belgium	36 (2.8)	81 (3.0)	117 (2.9)	
Canada	78 (6.1)	162 (5.9)	240 (6.0)	
Czech Republic	12 (0.9)	29 (1.1)	41 (1.0)	
Denmark	59 (4.6)	119 (4.4)	178 (4.4)	
Spain	36 (2.8)	70 (2.6)	106 (2.6)	
Finland	30 (2.3)	71 (2.6)	101 (2.5)	
France	7 (0.5)	21 (0.8)	28 (0.7)	
United Kingdom	40 (3.1)	77 (2.8)	117 (2.9)	
Hungary	21 (1.6)	61 (2.2)	82 (2.0)	
Israel	15 (1.2)	28 (1.0)	43 (1.1)	
Italy	40 (3.1)	90 (3.3)	130 (3.2)	
Mexico	17 (1.3)	40 (1.5)	57 (1.4)	
The Netherlands	59 (4.6)	135 (5.0)	194 (4.8)	
Norway	198 (15.4)	396 (14.5)	594 (14.8)	
New Zealand	20 (1.6)	32 (1.2)	52 (1.3)	
Poland	40 (3.1)	91 (3.3)	131 (3.3)	
Sweden	20 (1.6)	48 (1.8)	68 (1.7)	
Singapore	10 (0.8)	18 (0.7)	28 (0.7)	
Slovenia	7 (0.5)	18 (0.7)	25 (0.6)	
Slovakia	5 (0.4)	10 (0.4)	15 (0.4)	
United States	395 (30.7)	825 (30.3)	1220 (30.4)	
HYSTERECTOMY (VISIT: 1)				
No. Patients	1286	2725	4011	.956*
Yes	278 (21.6)	587 (21.5)	865 (21.6)	
No	1008 (78.4)	2138 (78.5)	3146 (78.4)	
TYPE OF HYSTERECTOMY (VISIT: 1)				
No. Patients	1286	2725	4011	.756*
Uterus/1 ovary	138 (49.6)	295 (50.3)	433 (50.1)	
Uterus/2 ovaries	123 (44.2)	249 (42.4)	372 (43.0)	
Unknown	17 (6.1)	43 (7.3)	60 (6.9)	
Unspecified	1008	2138	3146	

Table 3.1.4.7.2.2 provides baseline breast cancer risk assessment.

Appears This Way
On Original

**Table 3.1.4.7.2.2 Study CORE Baseline Breast Cancer Risk Assessment
All Patients Enrolled in Study CORE**

Variable	PLACEBO (N=1286)	RLX060 (N=2725)	Total (N=4011)	p-Value
CORE Gail Score (VISIT: 1)				
No. Patients	1286	2725	4011	.903**
Mean	1.94	1.94	1.94	
Median	1.70	1.70	1.70	
Standard Dev.	0.93	0.98	0.96	
Minimum	0.40	0.70	0.40	
Maximum	11.10	13.10	13.10	
Age at Menarche (VISIT: 1)				
No. Patients	1286	2725	4011	.522*
6 - <12	145 (11.3)	313 (11.5)	458 (11.4)	
12 - <14	575 (44.7)	1166 (42.9)	1741 (43.5)	
14 - <99	565 (44.0)	1242 (45.6)	1807 (45.1)	
Unspecified	1	4	5	
Age at Menarche (VISIT: 1)				
No. Patients	1285	2721	4006	.631**
Mean	13.35	13.38	13.37	
Median	13.00	13.00	13.00	
Standard Dev.	1.56	1.63	1.61	
Minimum	9.00	8.00	8.00	
Maximum	19.00	19.00	19.00	
Unspecified	1	4	5	
Age of First Live Birth (VISIT: 1)				
No. Patients	1286	2725	4011	.635*
0	31 (2.8)	59 (2.5)	90 (2.6)	
>0 - <20	85 (7.6)	199 (8.3)	284 (8.1)	
20 - <25	494 (44.0)	1019 (42.5)	1513 (43.0)	
25 - <30	356 (31.7)	806 (33.7)	1162 (33.0)	
>=30	157 (14.0)	312 (13.0)	469 (13.3)	
Unspecified	163	330	493	
Age of First Live Birth (VISIT: 1)				
No. Patients	1123	2395	3518	.628**
Mean	24.53	24.40	24.44	
Median	24.00	24.00	24.00	
Standard Dev.	8.15	7.35	7.61	
Minimum	0.00	0.00	0.00	
Maximum	99.00	99.00	99.00	
Unspecified	163	330	493	

Appears This Way
On Original

CORE First Degree Relatives with BC (VISIT: 1)				
No. Patients	1286	2725	4011	.175*
1 - <2	150 (90.9)	333 (89.5)	483 (89.9)	
2 - <3	14 (8.5)	33 (8.9)	47 (8.8)	
3 - <4	0	6 (1.6)	6 (1.1)	
>=4	1 (0.6)	0	1 (0.2)	
Unspecified	1121	2353	3474	
Number of Breast Biopsies (VISIT: 1)				
No. Patients	1286	2725	4011	.079*
1 - <2	157 (68.6)	343 (74.9)	500 (72.8)	
>=2	72 (31.4)	115 (25.1)	187 (27.2)	
Unspecified	1057	2267	3324	
Number of Breast Biopsies (VISIT: 1)				
No. Patients	229	458	687	.313**
Mean	1.77	1.57	1.64	
Median	1.00	1.00	1.00	
Standard Dev.	2.95	2.17	2.46	
Minimum	1.00	1.00	1.00	
Maximum	40.00	35.00	40.00	
Unspecified	1057	2267	3324	
Any Biopsies with Atypical Hyperplasia (VISIT: 1)				
No. Patients	1286	2725	4011	***
Yes	7 (3.1)	11 (2.4)	18 (2.6)	
No	203 (88.6)	416 (90.8)	619 (90.1)	
Unknown	19 (8.3)	31 (6.8)	50 (7.3)	
Unspecified	1057	2267	3324	

The CORE treatment groups were balanced with regard to breast cancer risk assessment characteristics. The mean baseline Gail model-based 5-year predicted risk of invasive breast cancer was 1.94% in both treatment groups. Approximately 54% of patients in each treatment group had a 5-year predicted invasive breast cancer risk of greater than or equal to 1.66%.

3.1.4.7.3 Applicant's Efficacy Analyses

The incidence of invasive breast cancer was determined from baseline (Visit 1) in CORE to the end of CORE. Of the 4011 women who enrolled in CORE, 21 (12 in the placebo group and 9 in the raloxifene group) developed breast cancer during their participation in MORE prior to Visit 1 of CORE. These 21 cases were included in the MORE breast cancer analysis and, accordingly, excluded from the dataset for the analysis of breast cancer endpoints in CORE.

Table 3.1.4.7.3.1 presents the efficacy results of invasive breast cancer for CORE.

Appears This Way
On Original

Table 3.1.4.7.3.1 Invasive Breast Cancer for CORE

Breast Cancer Category	CORE (N=3990) ^a					
	Placebo N=1274		Raloxifene N=2716		HR (95% CI)	p-value
	n	IR	n	IR		
Invasive	20	5.41	19	2.43	0.44 (0.24, 0.83)	0.009
ER-positive	15	4.05	12	1.54	0.37 (0.17, 0.79)	0.007
ER-negative	3	0.81	6	0.77	0.95 (0.24, 3.79)	0.941
ER unknown	2	0.54	1	0.13	N/A	N/A

Abbreviations: CI = confidence interval; ER = estrogen receptor; HR = hazard ratio; IR = incidence rate per 1000 patient-years; n = number of patients with breast cancer events; N = number of patients analyzed; N/A = Not Applicable;

^a A total of 4011 patients enrolled in CORE. This analysis includes only those patients enrolled in CORE who had not been diagnosed with breast cancer prior to enrollment (N=3990). The raloxifene group includes 1352 patients who were originally assigned to raloxifene HCl 60 mg/day in MORE and 1364 patients who were originally assigned to raloxifene HCl 120 mg/day in MORE. Breast cancers reported from CORE baseline (Visit 1) to the end of CORE are presented.

Of 46 total breast cancers reported during the 4-year treatment period of CORE, 7 (15%) were classified as noninvasive. Of these 7 cases (all of which were DCIS), 2 occurred among 1274 placebo-assigned patients and 5 occurred among 2716 raloxifene-assigned patients.

Reviewer's Comments

- For CORE trial, the results are difficult to interpret since subjects were from a subgroup of patients from the MORE trial and were not re-randomized in the CORE trial. Although baseline characteristics between two treatment arms appeared to be balanced, however, the trial results are more appropriate to be used as exploratory due to lack of re-randomization. The key of the randomization is to control some "known" factors, and more importantly, to control many "unknown" factors. Although the CORE trial showed that there were less incidence of the invasive breast cancer in raloxifene arm than that in the placebo arm, one should be cautious in interpreting the trial results because it was not a randomized trial.
- It is not clear what impact would be due to a treatment gap between the end of their participation in MORE and the start of their participation in CORE (the median time off therapy was approximately 10.6 months). During this gap, subjects did not receive study drug but could have taken marketed raloxifene, tamoxifen, other SERMS, or a hormone.
- The analyses for the incidence of invasive breast cancer by invasiveness and estrogen receptor status are presented below.

Breast Cancer Category	Placebo N=1,274 n (IR)	Raloxifene 60 mg N=2,716 n (IR)	Relative Risk (95% CI)
Invasive	20 (5.41)	19 (2.43)	0.45 (0.23, 0.89)
ER Positive	15 (4.05)	12 (1.54)	0.38 (0.16, 0.87)
ER negative	3 (0.81)	6 (0.77)	0.95 (0.20, 5.85)
ER unknown	2 (0.54)	1 (0.13)	NA
Non-invasive	2 (0.54)	5 (0.64)	1.18 (0.19, 12.44)
DCIS	2 (0.54)	5 (0.64)	1.18 (0.19, 12.44)
LCIS	0 (0.00)	0 (0.00)	NA
Invasive unknown	0 (0.00)	0 (0.00)	NA
All	22 (5.95)	24 (3.07)	0.52 (0.28, 0.96)

4. The results of an exploratory subgroup analysis of the invasive breast cancer based on Gail score are presented below. P-values are for descriptive purpose only.

Gail Score	Invasive Breast Cancer	RLX 2,716	PLB 1,274	Absolute Risk Difference	RR (95% CI)	P-value
≥ 1.66	Subgroup	1,473	670	- 4.63	0.36 (0.15, 0.85)	.007
	No. Event (IR)	11 (2.60)	14 (7.23)			
< 1.66	Subgroup	1,243	604	- 1.19	0.65 (0.20, 2.28)	.430
	No. Event (IR)	8 (2.24)	6 (3.43)			

5. The analyses of important safety outcomes are presented below.

Appears This Way
On Original

Events ^a	RLX 2,716	PLB 1,274	RLX IR	PLB IR	Absolute Risk Difference	Relative Risk (95% CI)
Clinical vertebral fracture ^b	65/2725	32/1286	8.28	8.56	-0.28	0.97 (0.62, 1.53)
Death	47/2725	29/1286	5.99	7.76	-1.77	0.77 (0.48, 1.27)
Death due to Stroke	6/2725	1/1286	0.76	0.27	+0.49	2.81 (0.34, 129)
Stroke	49/2725	14/1286	6.24	3.75	+2.49	1.65 (0.92, 2.98)
Deep vein thrombosis	17/2725	4/1286	2.17	1.07	+1.10	2.01 (0.68, 5.95)
Pulmonary embolism	9/2725	0/1286	1.15	0.00	+1.15	NA
Endometrial and uterine cancer ^c	4/2138	3/1008	0.65	1.02	-0.37	0.64 (0.11, 4.35)
Ovarian Cancer	2/2725	2/1286	0.25	0.54	-0.29	0.46 (0.03, 6.39)
Hysterectomy ^c	13/2138	10/1008	2.11	3.40	-1.29	0.62 (0.25, 1.58)
Hot Flashes	26/2725	11/1286	3.31	2.94	+0.37	1.13 (0.54, 2.52)
Leg Cramps	90/2725	36/1286	11.46	9.63	+1.83	1.19 (0.80, 1.80)
Peripheral edema	61/2725	30/1286	7.77	8.03	-0.26	0.97 (0.62, 1.55)
Cholelithiasis ^d	35/2725	12/1286	4.46	3.21	+1.25	1.39 (0.70, 2.94)

Abbreviations: IR = Incidence Rate per 1000 Patient-years; PLB = Placebo; RLX = Raloxifene.

^a Breast cancer events were for the patients who enrolled in CORE and had not been diagnosed with breast cancer prior to Visit 1.

^b Vertebral fractures were collected as adverse events.

^c Only patients with an intact uterus were considered for denominator (raloxifene denominator = 2138, placebo denominator = 1008).

^d Gallbladder status at baseline was not ascertained in the CORE trial.

3.2 Evaluation of Safety

Please refer to efficacy analyses section and Clinical Review of this application for complete safety evaluation.

4. Findings in Special/Subgroup Populations

4.1 Gender, Race, and Age

Since all patients were female, statistics by gender are irrelevant.

Table 4.1.1 presents the descriptive statistics of the incidence of invasive breast cancer by trial and age group.

Table 4.1.1 Invasive Breast Cancer by Age

Trial	Subgroup	Raloxifene			Control		
		N	n	(%)	N	n	(%)
STAR	≤ 65	7926	136	(.017)	7943	123	(.015)
	> 65	1825	37	(.020)	1793	45	(.025)
RUTH	≤ 65	1854	12	(0.65)	1877	26	(1.39)
	> 65	3190	28	(0.88)	3180	44	(1.39)
MORE	≤ 65	1008	3	(.003)	1026	12	(.012)
	> 65	1549	8	(.005)	1550	26	(.017)
CORE	≤ 65	904	7	(.008)	404	6	(.015)
	> 65	1812	12	(.007)	870	14	(.016)

Table 4.1.2 presents the descriptive statistics of the incidence of invasive breast cancer by trial and race.

Table 4.1.2 Invasive Breast Cancer by Race

Trial	Subgroup	Raloxifene			Control		
		N	n	(%)	N	n	(%)
STAR	caucasian	9112	162	(.018)	9105	160	(.018)
	other	639	11	(.017)	631	8	(.013)
RUTH	caucasian	4234	38	(0.90)	4247	66	(1.55)
	other	810	2	(0.25)	810	4	(0.49)
MORE	caucasian	2455	11	(.004)	2465	37	(.015)
	other	102	0	(.000)	111	1	(.009)
CORE	caucasian	2189	15	(.007)	1034	15	(.015)
	other	527	4	(.008)	240	5	(.020)

4.2 Other Special/Subgroup Populations

There was no analysis performed on other subgroups.

5. Summary and Conclusions

5.1 Statistical Issues and Collective Evidence

The analyses and results of four trials, STAR, RUTH, MORE and CORE, were submitted in this supplemental application to support two indications.

Table 5.1.1 presents the analyses for the incidence of invasive breast cancer for four trials. The statistical test was a stratified log-rank test for STAR, and a log-rank test for RUTH, MORE and CORE.

Table 5.1.1 Analyses for the Incidence of Invasive Breast Cancer

Trial	Treatment	Number Of Subjects	Invasive Breast Cancer Event	Incidence Rate per 1000 patient-years	Relative Risk (95% CI)	P-value
STAR	Raloxifene	9,751	173	4.40	1.02 (0.82, 1.27)	0.9868
	Tamoxifen	9,736	168	4.30		
RUTH	Raloxifene	5,044	40	1.50	0.56 (0.37, 0.84)	0.0032
	Placebo	5,057	70	2.66		
MORE	Raloxifene	2,557	11	1.26	0.29 (0.13, 0.58)	<0.0001
	Placebo	2,576	38	4.36		
CORE	Raloxifene	2,716	19	2.43	0.45 (0.23, 0.89)	0.0092
	Placebo	1,274	20	5.41		

For STAR trial, the stratified log-rank test for the primary analysis was not statistically significant. A post-hoc non-inferiority analysis, which used the sub-population from the NSABP-P1 trial, was performed to estimate tamoxifen effect size. The NSABP-P1 trial was conducted to compare the incidence of invasive breast cancer between tamoxifen and placebo. The data of women 50 years of age or older from the NSABP-P1 trial showed that tamoxifen decreased the incidence of invasive breast cancer by 53%. If one assumes that tamoxifen would have the same effect in STAR trial if a placebo arm would be included in STAR trial, the results of the non-inferiority analysis indicated that raloxifene may lose up to 35% of tamoxifen effect.

Since STAR trial failed to demonstrate superiority of raloxifene over tamoxifen which was the primary goal of the trial, any additional analyses used to support any claims violate the statistical principle. One may argue that the STAR trial was a large trial which should not be ignored and it will be almost impossible to repeat such a trial. However, a large trial, in many aspects, is not necessarily a better trial. The reason for a large trial in this case is that the effect is too small and events are rare. When the effect size is small and sample size is large, it is often difficult to control many confounding factors.

In a non-inferiority analysis setting, a percent retention needs to be pre-specified, the control effect used needs to be well estimated which is often derived from several similar historical trials, variability among trials and within trials needs to be considered, and the constancy assumption which assumes that tamoxifen would have the same effect over placebo between NSABP-P1 trial and STAR trial if a placebo arm would be included in STAR trial needs to be satisfied. In the current non-inferiority analysis, the percent retention was not pre-specified, the control effect was derived from a subpopulation of one trial, variability within the trial was not discussed, and the validity of the constancy assumption is very difficult or impossible to verify.

For RUTH trial, the co-primary endpoint of the incidence of invasive breast cancer was amended after initiation of the trail. The log-rank test for the primary analysis was statistically significant.

For MORE trial, the incidence of invasive breast cancer was a secondary safety endpoint. The trial was not designed to demonstrate the raloxifene effect on the incidence of invasive breast cancer. Although p-value of the log-rank test on the invasive breast cancer endpoint was less than 0.05, this statistical test was not planned, especially, not adjusted for multiple secondary and safety endpoints. One should realize that for a trial having many secondary and safety endpoints, one can always find some “significant” endpoints. Such results may really provide true discoveries, but one should not relay on this type of results because the significance level of the tests associated with this type of results is completely uncontrolled.

For CORE trial, the results are difficult to interpret since subjects were from a subgroup of patients from the MORE trial and were not re-randomized in the CORE trial. Although baseline characteristics between two treatment arms appeared to be balanced, however, the trial results are more appropriate to be used as exploratory due to lack of re-randomization. The key of the randomization is to control some “known” factors, and more importantly, to control many “unknown” factors. Although the CORE trial showed that there were less incidence of the invasive breast cancer in raloxifene arm than that in the placebo arm, one should be cautious in interpreting the trial results because it was not a randomized trial.

Table 5.1.2 presents the efficacy and important safety outcomes for STAR trial.

Appears This Way
On Original

Table 5.1.2 STAR: Efficacy and Important Safety Outcomes

Events	# Events (%)		IR ^a		RR (95% CI) ^b
	Tamoxifen N=9736	Raloxifene N=9751	Tamoxifen	Raloxifene	
All breast cancers	228 (2.3)	256 (2.6)	5.85	6.54	1.12(0.93,1.34)
Invasive	168 (1.7)	173 (1.8)	4.30	4.40	1.02(0.82,1.27)
Non-invasive	60 (0.6)	83 (0.9)	1.54	2.12	1.38(0.98,1.95)
Clinical vertebral fracture	58	58	1.47	1.46	0.99(0.68,1.46)
Death	109	104	2.76	2.62	0.95(0.72,1.25)
Death due to stroke	7	5	0.18	0.13	0.71(0.18,2.60)
Stroke	56	54	1.42	1.36	0.96(0.65,1.42)
Deep Vein Thrombosis	92	67	2.35	1.69	0.72(0.52,1.00)
Pulmonary Embolism	58	38	1.47	0.96	0.65(0.42,1.00)
Endometrial Cancer ^c	37/4739	23/4715	1.99	1.21	0.61(0.34,1.05)
Ovarian Cancer	14	18	0.52	0.66	1.27(0.60,2.76)
Cataracts	435	343	13.19	10.34	0.78(0.68,0.91)
Hysterectomy	246/4739	92/4715	13.25	4.84	0.37(0.28,0.47)
Hot Flashes	7170	6748	181.71	169.91	0.94(0.90,0.97)
Leg Cramps	5999	5373	152.03	135.29	0.89(0.86,0.92)
Edema ^d	664	741	16.83	18.66	1.11(1.00,1.23)
Cholelithiasis ^e	NA	NA	NA	NA	NA

^aIR=incidence rate per 1000 patient-years

^bRelative risk for raloxifene compared to tamoxifen.

Relative Risk >1 indicates higher incidence for raloxifene compared to tamoxifen

Relative Risk < 1 indicates lower incidence for raloxifene compared to tamoxifen

^c Only patients with a uterus at baseline (tamoxifen n = 4739; raloxifene n = 4715)

^c Hysterectomy was calculated as a risk ratio.

^d Peripheral edema is not a coding term in CTC v2.0.

^e Cholelithiasis is not a coding term in CTC v2.0.

Table 5.1.3 presents the efficacy and important safety outcomes for RUTH trial.

Appears This Way
On Original

Table 5.1.3 RUTH: Efficacy and Important Safety Outcomes

Events	Raloxifene 5,044	Placebo 5,057	Raloxifene IR	Placebo IR	Absolute Risk Difference	Relative Risk (95% CI)
Invasive breast cancer	40	70	1.50	2.66	-1.16	0.56 (0.37, 0.84)
Noninvasive breast cancer	11	5	0.41	0.19	+0.22	2.18 (0.70, 7.99)
Invasiveness unknown	1	1	0.04	0.04	+0.00	NA
All breast cancers	52	76	1.95	2.89	- 1.04	0.67 (0.46, 0.97)
Clinical vertebral fracture	64	97	2.40	3.70	-1.30	0.65 (0.47, 0.90)
Death	554	595	20.68	22.45	-1.77	0.92 (0.82, 1.04)
Death due to Stroke	59	39	2.20	1.47	+0.73	1.50 (0.98, 2.30)
Stroke	249	224	9.46	8.60	+0.86	1.10 (0.91, 1.32)
Deep vein thrombosis	65	47	2.44	1.78	+0.66	1.37 (0.94, 1.99)
Pulmonary embolism	36	24	1.35	0.91	+0.44	1.49 (0.89, 2.49)
Endometrial cancer ^a	21/3900	17/3882	1.01	0.83	+0.18	1.22 (0.61, 2.46)
Ovarian Cancer ^b	17/4559	10/4606	0.70	0.41	+0.29	1.71 (0.74, 4.17)
Hysterectomy ^a	58/3900	53/3882	2.79	2.60	+0.19	1.07 (0.73, 1.59)
Hot Flashes	397	241	14.82	9.09	+5.73	1.63 (1.39, 1.92)
Leg Cramps	483	334	18.03	12.60	+5.43	1.43 (1.24, 1.65)
Peripheral edema	706	583	26.36	22.00	+4.36	1.20 (1.07, 1.34)
Cholelithiasis ^c	168/4144	131/4111	7.83	6.20	+1.63	1.26 (1.00, 1.60)

Abbreviations: IR = Incidence Rate per 1000 Patient-years.

^a Only patients with an intact uterus were considered for the denominator (raloxifene denominator = 3900, placebo denominator = 3882).

^b Only patients with at least one ovary were considered for the denominator (raloxifene denominator = 4559, placebo denominator = 4606).

^c Only patients with an intact gallbladder at baseline (raloxifene n=4144, total person-years of follow-up=21467; placebo n=4111, total person-years of follow-up=21136).

Table 5.1.4 presents the efficacy and important safety outcomes for MORE trial.

Table 5.1.4 MORE: Efficacy and Important Safety Outcomes

Events ^a	Raloxifene 2,557	Placebo 2,576	Raloxifene IR	Placebo IR	Absolute Risk Difference	RR (95% CI)
Invasive breast cancer	11	38	1.26	4.36	-3.10	0.29 (0.13, 0.58)
Noninvasive breast cancer	3	5	0.34	0.57	-0.23	0.60 (0.09, 3.07)
Invasiveness unknown	3	1	0.34	0.11	+0.23	2.99 (0.24, 156)
All breast cancers	17	44	1.94	5.05	-3.11	0.38 (0.21, 0.69)
Clinical vertebral fracture	62	107	7.08	12.27	-5.19	0.58 (0.42, 0.80)
Death	64/5129	36	3.63	4.13	-0.50	0.88 (0.58, 1.36)
Death due to Stroke	9/5129	6	0.51	0.69	-0.18	0.74 (0.23, 2.52)
Stroke	91/5129	56	5.16	6.42	-1.26	0.80 (0.57, 1.14)
Deep vein thrombosis	44/5129	8	2.50	0.92	+1.58	2.72 (1.27, 6.68)
Pulmonary embolism	22/5129	4	1.25	0.46	+0.79	2.72 (0.92, 10.85)
Endometrial and uterine cancer ^b	8/3960	5/1999	0.59	0.74	+0.15	0.80 (0.23, 3.10)
Ovarian Cancer	6/5129	6/1999	0.34	0.69	-0.35	0.49 (0.13, 1.84)
Hysterectomy ^b	40/3960	22/1999	2.93	3.24	-0.31	0.90 (0.52, 1.60)
Hot Flashes	512/5129	151	29.04	17.31	+11.73	1.68 (1.40, 2.03)
Leg Cramps	443/5129	150	25.13	17.20	+7.93	1.46 (1.21, 1.77)
Peripheral edema	340/5129	134	19.29	15.36	+3.93	1.26 (1.03, 1.55)
Cholelithiasis ^c	93/5129	45	5.28	5.16	+0.12	1.02 (0.71, 1.50)

Abbreviations: IR = Incidence Rate per 1000 Patient-years; RR=Relative risk.

^a Breast cancer and clinical vertebral fracture events are for the raloxifene 60 mg/day arm only; denominator = 2557. For the safety events of death, death due to stroke, stroke, deep vein thrombosis, pulmonary embolism, and ovarian cancer, the raloxifene 60 and 120 mg/day arms were pooled to have the greatest opportunity to detect safety signals; thus, the denominator for these events is 5129.

^bOnly patients with a uterus at baseline (pooled raloxifene n=3960, total person-years of follow-up=13659.16; placebo n=1999, total person-years of follow-up=6791.41). "Hysterectomy" included MedDRA Preferred Terms of "Hysterectomy," "Hysterosalpingo-oophorectomy," and "radical hysterectomy."

^cGallbladder status at baseline was not ascertained in the MORE trial.

Table 5.1.5 presents the efficacy and important safety outcomes for CORE trial.

Table 5.1.5 CORE: Efficacy and Important Safety Outcomes

Events ^a	RLX 2,716	PLB 1,274	RLX IR	PLB IR	Absolute Risk Difference	Relative Risk (95% CI)
Invasive breast cancer	19	20	2.43	5.41	-2.98	0.45 (0.23, 0.89)
Noninvasive breast cancer	5	2	0.64	0.54	+0.10	1.18 (0.19, 12.44)
Invasiveness unknown	0	0	0.00	0.00	0.00	NA
All breast cancers	24	22	3.07	5.95	-2.88	0.52 (0.28, 0.96)
Clinical vertebral fracture ^b	65/2725	32/1286	8.28	8.56	-0.28	0.97 (0.62, 1.53)
Death	47/2725	29/1286	5.99	7.76	-1.77	0.77 (0.48, 1.27)
Death due to Stroke	6/2725	1/1286	0.76	0.27	+0.49	2.81 (0.34, 129)
Stroke	49/2725	14/1286	6.24	3.75	+2.49	1.65 (0.92, 2.98)
Deep vein thrombosis	17/2725	4/1286	2.17	1.07	+1.10	2.01 (0.68, 5.95)
Pulmonary embolism	9/2725	0/1286	1.15	0.00	+1.15	NA
Endometrial and uterine cancer ^c	4/2138	3/1008	0.65	1.02	-0.37	0.64 (0.11, 4.35)
Ovarian Cancer	2/2725	2/1286	0.25	0.54	-0.29	0.46 (0.03, 6.39)
Hysterectomy ^c	13/2138	10/1008	2.11	3.40	-1.29	0.62 (0.25, 1.58)
Hot Flashes	26/2725	11/1286	3.31	2.94	+0.37	1.13 (0.54, 2.52)
Leg Cramps	90/2725	36/1286	11.46	9.63	+1.83	1.19 (0.80, 1.80)
Peripheral edema	61/2725	30/1286	7.77	8.03	-0.26	0.97 (0.62, 1.55)
Cholelithiasis ^d	35/2725	12/1286	4.46	3.21	+1.25	1.39 (0.70, 2.94)

Abbreviations: IR = Incidence Rate per 1000 Patient-years; PLB = Placebo; RLX = Raloxifene.

^a Breast cancer events were for the patients who enrolled in CORE and had not been diagnosed with breast cancer prior to Visit 1.

^b Vertebral fractures were collected as adverse events.

^c Only patients with an intact uterus were considered for denominator (raloxifene denominator = 2138, placebo denominator = 1008).

^d Gallbladder status at baseline was not ascertained in the CORE trial.

Summary

The STAR trial failed to demonstrate superiority of raloxifene over tamoxifen, and had many problems in the post-hoc non-inferiority analysis. The RUTH, MORE and CORE trials showed that

there were fewer invasive breast cancer events in raloxifene-treated subjects than that placebo-treated subjects. However, one should be cautious in interpreting of the results of MORE and CORE trials. In addition, raloxifene-treated subjects had more exposure to thromboembolic adverse events than those placebo-treated subjects numerically.

5.2 Conclusions and Recommendations

The applicant submitted the analyses and results of four trials, STAR, RUTH, MORE and CORE, to seek registration of raloxifene for two indications: “reduction in the risk of invasive breast cancer in postmenopausal women at high risk for breast cancer”, and “reduction in the risk of invasive breast cancer in postmenopausal women with osteoporosis”. Raloxifene is currently approved for the treatment of osteoporosis and prevention of osteoporosis.

The data and analyses from STAR trial, which supports the indication “reduction in the risk of invasive breast cancer in postmenopausal women at high risk for breast cancer”, failed to demonstrate the superiority of raloxifene over tamoxifen. The applicant performed a non-inferiority analysis, which was not planned in the original design, to compare raloxifene to tamoxifen. The results of the non-inferiority analysis indicated that raloxifene may lose up to 35% tamoxifen effect, but there were many problems involved in this non-inferiority analysis. Raloxifene had more events in several safety categories while had fewer events in other safety categories compared to tamoxifen.

The data and analyses from RUTH, MORE and CORE trials which support the indication “reduction in the risk of invasive breast cancer in postmenopausal women with osteoporosis” showed that there were fewer invasive breast cancer events in raloxifene-treated subjects than that in placebo-treated subjects. However, one should note that in RUTH trial the incidence of the invasive breast cancer was amended as a co-primary endpoint, in MORE trial the incidence of the invasive breast cancer was a secondary safety endpoint, and in CORE trial subjects were not randomized. In addition, raloxifene-treated subjects had more exposure to thromboembolic adverse events than those placebo-treated subjects numerically.

This supplemental application was discussed at the Oncology Drugs Advisory Committee (ODAC) on July 24, 2007. The committee recommends approval for both indications.

The final regulatory action should be based on clinical judgment and acceptability of risk-benefit profile.

Appears This Way
On Original

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

/s/

Kun He
9/7/2007 01:42:26 PM
BIOMETRICS

Rajeshwari Sridhara
9/9/2007 10:12:24 PM
BIOMETRICS

Please also see secondary review comments by the tatistical
team leader.

Aloka Chakravarty
9/10/2007 10:30:15 AM
BIOMETRICS

Appears This Way
On Original