10 APPENDICES

10.1 Review of Individual Study Reports

Reviews of the study reports of the three placebo-controlled raloxifene trials are available in this section:

- Note that the numbers in the parenthesis (or otherwise) associated with the section and subsection headings in the text in the study reports correspond to the numbers, sections, sub-sections, etc. in the original study reports.

TABLE OF SUMMARIES OF THE CLINICAL STUDY REPORTS

Summary 1 RUTH Clinical Study Report

Review of the Clinical Study Report: RUTH (Raloxifene Use for the Heart)

Raloxifene Hydrochloride or Placebo in Postmenopausal Women at Risk for Major Coronary Events

Clinical Study Synopsis: Study H3S-MC-GGIO

This Phase 3 multicenter study was conducted at 192 study centers in 26 countries and had 177 principal investigative sites.

Length of the study: 7 years, 5 months
Date first patient enrolled: 25 June 1998
Date last patient completed: 21 November 2005

Objectives

Primary:
To assess whether chronic oral treatment with raloxifene hydrochloride 60 mg/day, compared with placebo, reduced the incidence of:

- The **combined endpoint** of coronary death, nonfatal (including silent) myocardial infarction (MI), or hospitalized acute coronary syndrome (ACS) other than MI (coronary primary endpoint);

and

- Invasive breast cancer (breast cancer primary endpoint)

in postmenopausal women at risk for major coronary events

Secondary:
1) To assess whether chronic oral treatment with raloxifene 60 mg/day compared with placebo changed the incidence of the following endpoints:

- Cardiovascular (CV) death, nonfatal MI, hospitalized ACS other than MI, myocardial revascularization, or stroke. (Each endpoint was assessed separately and as a combined endpoint.)
- Coronary death
- Hospitalized ACS
- All-cause hospitalization
- Non-coronary arterial revascularization or non-traumatic lower extremity amputation
- All breast cancer
- Fractures
- Venous thromboembolic events (VTEs)

2) To assess the long-term safety of raloxifene 60 mg/day in postmenopausal women at risk for coronary events

Other objectives:

- To assess the effect of raloxifene compared with placebo on biochemical markers of CV risk in either the entire cohort or a subset of patients
- To assess the pharmacokinetics of raloxifene in a subset of this study population

Study Design: This was a Phase 3, multicenter, double-blind, placebo-controlled, randomized, parallel groups study. Approximately 10,000 patients were to be enrolled and randomly assigned to one of two therapy groups: raloxifene HCl 60 mg/day *or* placebo. Patients were to be followed until a minimum of

1,268 patients experienced an adjudicated coronary primary endpoint event. This was projected to occur after a minimum of 5 years follow-up on all non-discontinued patients, unless the trial was stopped early based on an interim analysis. The observed coronary primary endpoint event rate was lower than predicted.
Despite two protocol amendments to remediate this issue, the duration of study follow-up to achieve 1,268 coronary primary endpoint events would have exceeded the 5 to 7.5 years of follow-up stated in the informed consent document. Therefore, the decision was made to conclude the trial after the last randomized patient had been followed for 5 years.

Number of Patients:

Planned:	Approximately 10,000
Randomized:	5,044 raloxifene HCl 60 mg/day
	5,057 placebo
Completed:	4,060 raloxifene HCl 60 mg/day
	3,979 placebo

Diagnosis and Main Criteria for Inclusion: Patients must have met all of the following criteria to be included in the study:

- Postmenopausal (verified by no menses for 1 year) and ≥55 years of age
- Able to understand and sign the informed consent document (ICD) before entering the study
- Established coronary heart disease (CHD) *or* at increased risk for CHD. Increased risk for CHD was defined as having at least 4 points based on the following coronary conditions and risk factors:
  - established CHD, including MI, angina pectoris with documented CHD, catheter-based coronary revascularization, or coronary artery bypass grafting, or documented lower extremity arterial disease (4 points each)
  - diabetes mellitus (3 points)
  - age 70 years or greater (2 points)
  - cigarette smoking (1 point)
  - hypertension (1 point)
  - hyperlipidemia (1 point)

Test Product, Dose, and Mode of Administration: Raloxifene HCl 60 mg/day, given orally once a day as one 60 mg tablet

Reference Therapy, Dose, and Mode of Administration: Placebo tablet given orally once a day

Duration of Treatment: Until the last randomized patient had been followed for a minimum of 5 years

Variables:

Efficacy: Efficacy was evaluated based on adjudication of endpoint events (coronary death, nonfatal MI, hospitalized ACS other than MI, invasive breast cancer, fractures, all breast cancers, all deaths, all hospitalizations, VTEs, strokes, revascularizations and amputations). Supporting clinical documentation, mammograms, or electrocardiograms were reviewed during the adjudication processes. Biochemical markers of CV risk, lipid parameters and fibrinogen, were also collected.

Safety: Safety was evaluated by reporting and collection of adverse event (AE) data, vital signs, physical findings, and routine laboratory testing.

Pharmacokinetic/Pharmacodynamic: The raloxifene concentration evaluation included data from 253 patients who were subset from patients enrolled in 11 investigator sites. Two blood samples were collected from patients at Visits 5 and 7 following 12 and 24 months of 60 mg raloxifene HCl once daily, respectively. Samples were collected at least 1 hour apart during each visit.

Evaluation Methods:

Statistical: In general, all statistical tests were performed at a two-sided significance level of 0.05 with exception of the test of the primary endpoints. The conclusion of significance of the primary endpoints was adjusted for the planned efficacy assessment at three coronary and one breast cancer interim analyses, as well as for the multiplicity of the primary endpoints to preserve the trial-wide type 1 error rate at $< \delta 0.05$.
Consequently, the treatment effect on the coronary primary endpoint was tested at a two-sided significance level of 0.0423; the breast cancer primary endpoint was tested at a two-sided significance level of 0.008.
Analyses of each primary and secondary endpoint were based on the time to first event, for all randomized patients according to the intent-to-treat principle.

Pharmacokinetic: The pharmacokinetics of raloxifene in plasma were assessed for the 2-year data following the conclusion of the study. Graphical visualization and descriptive statistical analyses of the LY139481 concentration data using S-PLUS were the primary techniques of data evaluation.

Summary:
A total of 10,101 postmenopausal women at risk for major coronary events were randomly assigned to placebo (5,057) or raloxifene, 60 mg/day (5,044). The majority was Caucasian, 12% were current smokers, 46% had diabetes mellitus, 78% had hypertension, 73% had hyperlipidemia, 11% had lower extremity arterial disease, and about 50% had a history of CHD.

Median follow-up: 5.56 years

Efficacy

Primary Objectives:

- The incidence of invasive breast cancer was significantly reduced by 44% in the raloxifene group compared with the placebo group, thus meeting the primary breast cancer objective of the study. This difference was primarily due to a significant 55% reduction in the incidence of estrogen receptor positive invasive breast cancer.
- The incidence of coronary death, nonfatal MI, or hospitalized ACS other than MI *combined* did not differ between treatment groups; therefore, the primary coronary objective of the trial was not met. There was no difference between treatment groups on the incidences of coronary death, nonfatal MI, or hospitalized ACS other than MI, individually.

Secondary Objectives:

- The incidence of CV death, nonfatal (including silent) MI, hospitalized ACS other than MI, stroke, or myocardial revascularization combined did not differ between treatment groups.
- The incidence of all strokes did not differ between treatment groups.
- The incidence of *overall mortality*, including *overall CV mortality*, did not differ between treatment groups. A significant, 20% reduction in *death due to non-CV causes* was observed in patients assigned raloxifene. A significant, 49% increase in the incidence of *death due to stroke* was observed in patients assigned raloxifene; this corresponds to an absolute risk increase of 0.7 deaths due stroke per 1000 woman-years.
- Significantly fewer patients in the raloxifene group compared with the placebo group had one or more hospitalizations for any cause.
- The incidences of myocardial or non-coronary arterial revascularizations, or non-traumatic lower extremity amputations did not differ between treatment groups.
- The incidence of *all breast cancers*, regardless of invasive status, was significantly decreased by 33% in the raloxifene group compared with the placebo group.
- The incidence of clinical *vertebral fracture* was significantly reduced by 35% in the raloxifene group compared with the placebo group. There was no difference between treatment groups in the incidences of *non-vertebral fractures* or hip/femur or wrist fractures.
- The incidence of all VTEs and the combined incidence of pulmonary embolism or deep vein thromboses were each significantly increased by 44% in the raloxifene group compared with placebo.

Safety

- There were no significant differences between treatment groups in the proportions of patients reporting ≥1 treatment-emergent adverse event or reporting ≥1 serious adverse event. The proportion of patients who reported at least one AE leading to discontinuation of study drug was not significantly different between treatment groups.
- Peripheral edema, muscle spasms, hot flush, dyspepsia, cholelithiasis, arthritis, and intermittent claudication were reported by ≥2% of raloxifene-assigned patients and significantly more frequently by raloxifene-assigned patients than by placebo-assigned patients.
- An increased incidence of gallbladder disease was reported in patients assigned to raloxifene compared to placebo however, the cholecystectomy rates did not differ between treatment groups.
- There were no significant treatment group differences in the incidences of all cancers or reproductive cancers, specifically endometrial or ovarian cancers.
- No clinically relevant differences between treatment groups were observed for laboratory analytes (aspartate transaminase, total bilirubin, blood urea nitrogen, creatinine, fasting glucose, or hemoglobin A1c), vital signs (blood pressure, heart rate), or physical findings (body mass index, height, weight).

Other Objectives

- Total cholesterol, low-density lipoprotein cholesterol, and fibrinogen levels were significantly decreased in raloxifene-assigned patients compared to those assigned placebo. There was a significant increase in high-density lipoprotein cholesterol levels in the raloxifene group compared with placebo.
- The overall mean steady-state raloxifene plasma concentration in this patient population was 1.38 ng/mL. No discernible differences in plasma raloxifene concentrations were observed based on the ethnic origin of Caucasians and patients of African descent represented in this population.

Conclusions:

- Raloxifene significantly reduced the incidence of invasive breast cancer in postmenopausal women at risk for major coronary events.
- Raloxifene had no effect on the incidence of coronary death, non-fatal MI, or hospitalized ACS other than MI—combined or individually.

- There was no significant difference between treatment groups in the incidences of *all strokes* or *overall mortality*, including *CV mortality*.
- A significant reduction in *non-CV deaths* in the raloxifene group was reported; however, the clinical relevance of this observation is unknown.
- An increased incidence of *death due to stroke* was observed in women assigned to raloxifene. Since the statistical significance of this finding was relatively weak (p=0.0499), this observation may be real or due to chance.

- In the raloxifene group, there was a significant increase in the incidence of VTE, a known serious, but uncommon, AE associated with raloxifene.

- Consistent with its known skeletal effects, raloxifene significantly decreased clinical vertebral fracture incidence.

- In summary, in postmenopausal women at risk for major coronary events, the benefits of raloxifene in reducing the incidences of invasive breast cancer and clinical vertebral fracture must be weighed against the increased risk of VTE and the possible increased risk of death due to stroke. With the exceptions of gallbladder disease and death due to stroke, which are new findings not observed in previous raloxifene clinical trials; the AEs reported during the trial were consistent with the known safety profile for raloxifene.

Ethical Conduct of the Study (5.2)

- This study was conducted in accordance with the ethical principles that have their origin in the Declaration of Helsinki and that are consistent with good clinical practices (GCPs) and the applicable laws and regulations.
- This study was conducted under United States Investigational New Drug (IND) application #57,137.

Patient Information and Consent (5.3)

- Lilly provided a study-specific ICD to each investigative site. The PI at each site was allowed to modify this ICD in order to meet applicable ERB guidelines at that site. All ICDs were compliant with the International Conference on Harmonization (ICH) guideline on GCP.
- The PI was responsible for obtaining informed consent and the appropriate signatures on the ICD from each patient or legal representative prior to the performance of any protocol procedures and administration of study drug. A properly executed, signed ICD was obtained from each patient or their legal representative. The PI provided a copy of the signed ICD to the patient, and a copy was maintained at the investigative site.

Investigators and Study Administrative Structure (6)

- RUTH (Study H3S-MC-GGIO or GGIO) was a multicenter study conducted at 177 investigative sites. Appendix 16.1.4 contains a list of all investigators who were principal or co-principal investigators at any time during the study.
- Table XX below lists the study committees and describes the role each had in the supervision or conduct of the study. Appendix 16.1.16 lists the members of each of these committees.
- **Members of all committees except the Publications Committee were external to Lilly** (not employees of Lilly). The Publications Committee included individuals who were external to Lilly as well as Lilly employees associated with GGIO.

Table xxx. Committees Involved in Study Supervision or Conduct

Committee	Role
Executive Committee	Proposed decisions for all major scientific issues related to the conduct of the trial
Steering Committee	Provided strategic input related to maintaining study timeline and proposed addenda or amendments
National Coordinators	Acted as a forum contributing to the scientific and operational success of the study and as a liaison between local investigators and the appropriate decision making body
Coronary Primary Endpoint Committee	Adjudicated coronary primary endpoint events
Breast Cancer Endpoint Committee	Adjudicated breast cancer events
Secondary Endpoint Committee	Adjudicated cardiovascular non-coronary deaths, non-cardiovascular deaths, and non-coronary endpoint events, except breast cancer events
Substudy Committee	Reviewed ancillary study proposals for scientific merit and feasibility within the context of the study protocol and then made recommendations to the Executive Committee and the sponsor regarding implementation of such ancillary studies[a]
Publications Committee	Facilitated high quality, accurate publications and presentations conveying study results, reviewed and approved all proposals, and encouraged broad participation by all investigators
Central ECG Laboratory	Interpreted all scheduled ECGs obtained during the study

[a] Six ancillary substudies were conducted by various study investigators.

Table xxx lists all vendors, laboratories, contract research organizations (CROs), and individuals contracted by Lilly during the conduct of this study.

Table xxx Study Management Services

Organization	Role

b(4)

- All statistical analyses conducted prior to unblinding were performed by [redacted], an external data analysis group (DAG) responsible for administrative data management and statistical analysis support for the data safety monitoring board (DSMB). b(4)
- All statistical analyses conducted after unblinding were performed by Lilly.

Introduction (7)

- Raloxifene is classified as a selective estrogen receptor modulator (SERM) based on its ability to act as an estrogen receptor agonist in bone and on lipid metabolism, while acting as an estrogen receptor antagonist in tissues such as breast and uterus. Several observations led to the hypothesis that estrogen may play a key role as a cardio-protective agent in women. This effect was believed to be mediated by estrogen's effects on cardiovascular (CV) risk factors, such as lipids (high-density lipoprotein cholesterol [HDL-C] and low-density lipoprotein cholesterol [LDL-C]) and non-lipid factors such as vasomotor tone and coagulation.
- In preclinical studies, raloxifene administration resulted in a significant reduction in serum cholesterol, regression of atherosclerosis, and endothelial-dependent relaxation by an estrogen receptor-dependent and nitric oxide-dependent mechanism.
- In clinical studies, raloxifene significantly reduced serum total cholesterol and LDL-C, while having little or no effect on triglycerides or HDL-C.
- In a post-hoc analysis of the Multiple Outcomes of Raloxifene Evaluation (MORE) trial, the osteoporosis treatment trial, raloxifene therapy for 4 years did not significantly affect

the incidence of serious CV adverse events (AEs) in the overall cohort, but significantly reduced the incidence of serious CV AEs in a subset of women who were retrospectively determined to be at increased CV risk.

- A secondary endpoint of the MORE trial was to determine the effect of raloxifene on breast cancer. Raloxifene significantly reduced the incidence of newly diagnosed invasive breast cancer compared to placebo at 40 months and 48 months. This reduction was primarily attributable to a reduction in incidence of estrogen receptor-positive invasive breast cancer with no apparent effect of raloxifene on estrogen receptor-negative invasive breast cancer.
- Accordingly, RUTH study was designed to evaluate the effect of raloxifene on the combined coronary endpoint (coronary death, nonfatal (including silent) MI, and hospitalized acute coronary syndrome (ACS) other than MI) in postmenopausal women at risk for major coronary events. A second primary objective was to determine the effect of raloxifene in reducing the incidence of invasive breast cancer.

Objectives (8)

Primary Objectives (8.1)

The two primary objectives of GGIO were to assess whether chronic oral treatment with raloxifene HCl 60 mg/day, compared with placebo, reduces the incidence of:
1. The combined endpoint of coronary death, nonfatal (including silent) MI, or hospitalized ACS other than MI (coronary primary endpoint); *and*
2. Invasive breast cancer (breast cancer primary endpoint),
in postmenopausal women at risk for major coronary events.

Secondary Objectives (8.2)

- To assess whether chronic oral treatment with raloxifene HCl 60 mg/day compared with placebo changes the incidence of the following endpoints:
  - CV death, nonfatal (including silent) MI, hospitalized ACS other than MI, myocardial revascularization, or stroke. (Each of these endpoints was assessed separately and as a combined endpoint.)
  - Coronary death.
  - All-cause mortality.
  - Hospitalized ACS.
  - All-cause hospitalization.
  - Non-coronary arterial revascularization or non-traumatic lower extremity amputation.
  - All breast cancer.
  - Fractures.
  - Venous thromboembolic events (VTEs).
- To assess the long-term safety of raloxifene HCl 60 mg/day in postmenopausal women at risk for coronary events.

Other Objectives (8.3)

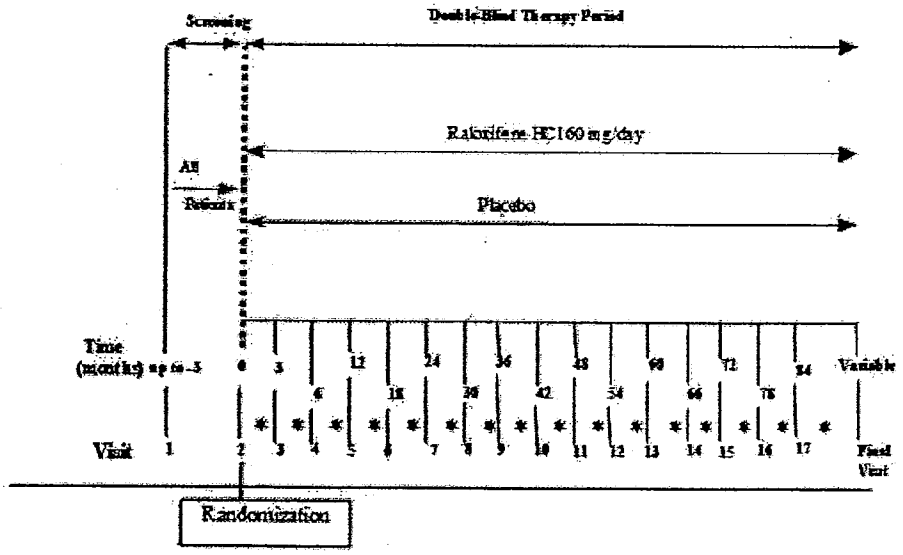
Other objectives of the study were:
- To assess the effect of raloxifene compared with placebo on biochemical markers of CV risk in either the entire cohort or a subset of patients.
- To assess the pharmacokinetics of raloxifene in a subset of this study population.
- To facilitate ancillary studies of high scientific quality and importance in conjunction with this study.

The ancillary sub-studies were conducted by various study investigators. The results of these studies are not included in this clinical study report. Results for all other objectives stated above are provided in this clinical study report.

Overall Study Design and Plan (9.1)

GGIO was a Phase 3, multicenter, double-blind, placebo-controlled, randomized, parallel study enrolling 10,101 patients to one of two therapy groups: raloxifene HCl 60 mg/day or placebo. Randomization was stratified by investigative site.
The study design and occurrence of visits are illustrated below (Figure xxx).
A detailed list and a schedule of events for each visit are found in the study schedule (Table GGIO.9.1).
Notably, at baseline, a medical history was collected and a physical examination (including vital signs) was performed. A mammogram was performed at baseline but a mammogram performed within 1 year of randomization was acceptable, if the report was received by the time of the randomization visit (Visit 2).

Office visits occurred every 6 months (±4 weeks) for the duration of the study.
Follow-up mammograms and clinical breast examinations were performed every 2 years from the time of the baseline assessments.
Follow-up ECGs were performed at the Year 2, Year 4, and final visits (if not performed within the 3 months prior to the final visit).
Blood chemistry analyses were performed annually and at final visit.
Patients who developed breast cancer during the trial were discontinued from study drug and were asked questions specific to breast cancer follow-up on a regular basis.



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

Figure GGIO.9.1. Study design for H3S-MC-GGIO.

Table GGIO.9.1. Study Schedule

Activity / Visit / Month[a] (Year)	1 up to -3	2 0	3[b] 3	4 6	5 12 (1)	6 18	7 24 (2)	8 30	9 36 (3)	10 42	11 48 (4)	12 54	13 60 (5)	14 66	15 72 (6)	16 78	17 84 (7)	FV[c]
Sign informed consent	✓																	
Patient number assigned	✓																	
Obtain medical release	✓																	
Clinical Assessments																		
Medical history		✓																
Cardiovascular, breast cancer, and VTE risk assessment		✓																
Physical examination		✓			✓		✓		✓		✓		✓		✓		✓	✓
Blood pressure and pulse	✓	✓			✓		✓		✓		✓		✓		✓		✓	✓
Weight and height measurement		✓			✓		✓		✓		✓		✓		✓		✓	✓
Mammogram[d]		✓					✓				✓				✓			✓
Breast exam[e]		✓					✓				✓				✓			
ECG[f]		✓					✓				✓							✓
Review of patient data[g]	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Record of event reporting		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Breast cancer follow-up information[h]				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Record of concomitant medications		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Laboratory Assessments																		
Review of local laboratory results[i]	✓																	
Chemistry tests[j]		✓			✓		✓		✓		✓		✓					✓
Lipids[j]		✓			✓								✓					✓
Hemoglobin A_{1C}		✓					✓											✓
Fibrinogen (subset)		✓			✓								✓					✓
Study drug plasma (subset)[k]					✓		✓											
Plasma and serum for storage (frozen)[l]		✓																
Study Drug																		
Randomization		✓																
Study drug administration		<----------																---------->
Study drug return				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Miscellaneous																		
Retention-related telephone contacts		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	

[a] Timing of visits were calculated based on Visit 2 (Month 0) and may have varied by ±4 weeks.
[b] Visit 3 was an optional visit scheduled for the purpose of encouraging compliance and developing the relationship between the staff and patient.
[c] Final visit (FV) occurred: 1) when a patient died, 2) when a patient refused any further contact with the site, 3) when the RUTH trial had completed.
[d] If a mammogram had been performed no more than 1 year before beginning therapy, it did not have to be repeated, provided the investigational site reviewed a record of the results by Visit 2. Timing of the follow-up mammograms was to have been calculated in reference to the baseline mammogram (2, 4, and 6 years after the baseline mammogram). A mammogram was performed at FV only if 2 years or more have elapsed since the previous one.
[e] The breast exam may have been performed by a physician or nurse.
[f] Investigators were to use ECG machines (with 3 channels or more), with the 12 leads in the standard sequence and mode, at an amplification of 10 mm/mV. At least 3 beats per lead were to have been recorded and a paper speed of 25 mm/sec or 50 mm/sec was preferred. Calibration was to have always been marked on the paper. Two originals were to have been done: an original ECG was to be retained at the investigative site and a second original sent to Lilly. The ECG was performed at FV only if a study ECG had not been performed in the previous 3 months.

g The investigator or his/her designee reviewed all patient data collected at each office visit.
h For all patients who had breast cancer reported during the study, follow-up information was to have been obtained at all visits.
i These tests included fasting serum glucose, LDL-C, HDL-C, triglycerides, bilirubin, AST, alkaline phosphatase, and serum creatinine. The results of these tests should have been dated no more than 3 months prior to Visit 1 and should have been reviewed before the patient was randomly assigned.
j Chemistry tests included: total bilirubin, AST, BUN, serum creatinine, fasting glucose. Lipid tests included: total cholesterol, HDL-C, LDL-C (calculated and direct measured), and triglycerides. Patients were to have been instructed to fast at least 12 hours before the visits when laboratory samples were to be collected. Regular medications may have been taken during the fasting period.
k Two blood samples were to have been drawn at least 1 hour apart. The actual dates and times of the last dose and the dose before the last dose of study drug were to have been recorded plus the dates and times of the blood draws were to have been recorded.
l An additional collection of serum and plasma to be frozen could have been planned after the first year in the study to explore any changes in the selected cardiovascular markers evaluated at baseline; this additional serum and plasma sample was not collected.
Note: Additional 6-month follow-up visits could have been scheduled beyond the 7th year in the study depending on the number of cumulative CO primary endpoints centrally adjudicated at that time. If that had occurred, the same procedures planned for Visit 17 were to have been performed. The study concluded before 6-month follow-up visits beyond the 7th year became necessary.

Planed Duration of the Trial

- Participants were to be followed until a minimum of 1,268 patients experienced a coronary primary endpoint event, centrally adjudicated as such. This was projected to occur after a minimum of 5 years of follow-up on all non-discontinued patients, unless the trial was stopped early on the basis of an interim analysis.
- The final breast cancer analysis was to be performed after the maximum time allowed to perform the 4-year mammograms in all patients had elapsed.
- As the trial duration was endpoint-driven, maximum duration of patient follow-up was dependent on the coronary primary endpoint event rate. Due to lower than expected coronary primary endpoint event rate, 1,268 coronary endpoint events would not have been accrued within the follow-up duration of 5 to 7.5 years specified in the ICD. Therefore, in December 2004, Lilly and the Raloxifene Use for The Heart (RUTH) Executive Committee (EC) decided that the study would be concluded after all women were followed for a minimum of 5 years. Section 9.8.1 contains complete details surrounding the decision to conclude the trial after the last woman randomized had been followed for 5 years.

Reviewer Comments: (Discussion of Study Design and Choice of Control Groups)
The study was randomized and double-blind to minimize bias. The patient population was chosen to test the efficacy of raloxifene for primary and secondary prevention of CHD. Choice of placebo as the comparator arm was appropriate as estrogen therapy was not indicated for prevention of CVD in postmenopausal women. The addition of raloxifene to the patient's regimen did not necessitate discontinuation of drugs required for the usual treatment of the patient's coronary disease or CV risk factors.

Selection of Study Population (9.3)

Eligible women were ≥ 55 years old and ≥ 1 year postmenopausal. They either had established CHD or were at increased risk for a major coronary event

Inclusion Criteria (9.3.1)

Women must have met all of the following criteria to be included in the study:

1. Postmenopausal (verified by no menses for ≥1 year) and ≥55 years of age.

2. Able to understand and sign the ICD before entering the study.
3. At risk for a MI. this was defined as having at least four points based on the following coronary conditions and risk factors: MI, angina pectoris with documented CHD, catheter based coronary revascularization, coronary artery bypass grafting, lower extremity arterial disease, diabetes mellitus, hypertension, hyperlipidemia, smoking, or age. See the study synopsis for the points assigned for these coronary conditions and risk factors.

Exclusion Criteria (9.3.2)

Patients were excluded from entering the trial for any of the following reasons:

- An MI or coronary artery bypass grafting within 3 months prior to randomization because major coronary events in that time frame would likely be unaffected by treatment.
- A catheter-based coronary revascularization procedure less than 6 months prior to randomization because the development of restenosis could occur before the conclusion of the 9-month lag phase anticipated with raloxifene treatment.
- History of cancer or suspected to have cancer such as breast carcinoma, endometrial carcinoma, or any other forms of cancer (other than excised superficial lesions such as basal or squamous cell carcinomas of the skin) within 5 years prior to randomization because of possible safety concerns. However, patients who had a history of endometrial carcinoma but who had had a hysterectomy for the disease within at least 5 years prior to randomization were allowed to participate.
- A life expectancy of less than 5 years due to factors other than heart disease, or considered to be poor medical or psychiatric risks for treatment, in the opinion of the investigator.
- Estrogen replacement therapy required within the 6 months prior to randomization, or current treatment with any estrogen or progestin-containing compounds (with the exception of vaginal formulations applied ≤3 times/week), hormonal agents, or an approved SERM because of possible safety concerns.
- Unexplained uterine bleeding within 6 months prior to randomization, any known or probable history of deep vein thrombosis (DVT), pulmonary embolism (PE) or retinal vein thrombosis, New York Heart Association (NYHA) Class III or IV heart failure, chronic liver disease, or chronic renal failure because of possible safety concerns.

Removal of Patients from Therapy or Assessment (9.3.3)

- Patients who required discontinuing study therapy permanently were not discontinued from the trial (in keeping with an intent-to-treat (ITT) analysis).
- Any patient who was inadvertently randomized but had a history of VTE, breast cancer, or endometrial cancer (except if a hysterectomy had been performed for the disease within at least 5 years prior to randomization) was required to permanently discontinue study drug.
- Study drug was permanently discontinued if a patient was diagnosed with either a breast cancer or a venous thromboembolic event during the trial.

- Patients who discontinued study drug were encouraged to attend study visits, undergo all measurements, and were followed for primary and secondary endpoints.
- Study drug may have been temporarily discontinued in the presence of an illness or a condition which led to a prolonged period of immobilization. Study drug could be resumed when the patient's illness resolved and/or the patient had resumed her previous level of mobility. If study drug was temporarily discontinued, patients continued to be followed for the primary and secondary endpoints.

Treatments (9.4)

Treatments Administered (9.4.1)

- Raloxifene HCl 60 mg, equivalent to 55.71 mg of raloxifene, was provided as oral tablets.
- Placebo was provided in tablets identical in appearance to those of raloxifene.
- Both raloxifene HCl 60 mg and placebo were manufactured by Eli Lilly and Company as tablets.
- Study drug was packaged by [redacted] in separate kits containing 240 tablets each. Each kit was numbered with a randomly generated number. b(4)

Method of Assigning Patients to Treatment Groups (9.4.3)

- At Visit 1, an interactive voice response system (IVRS) was used to assign a patient number to each participant who signed the ICD.
- At Visit 2, patients were randomly assigned by IVRS to one of two treatment groups: either raloxifene HCl 60 mg/day or placebo.
- Randomization was stratified by investigative site.

Selection of Doses in the Study (9.4.4)

- Enrolled patients received either raloxifene HCl 60 mg/day or placebo.
- Raloxifene HCl 60-mg/day dosage is approved and marketed for prevention and treatment of osteoporosis in postmenopausal women.

Selection and Timing of Dose for Each Patient (9.4.5)

- Patients were instructed to take one study drug tablet at the same time each day.

Blinding (9.4.6)

- Patients, investigators, and all other investigative site personnel involved in conducting the study were blinded to individual treatment assignments for the duration of the study.

- The investigator was unblinded to treatment only if it was needed for reasons of patient safety.
- If a patient's treatment assignment was unblinded by the investigator, the patient was permanently discontinued from study treatment, but was followed for primary and secondary endpoints.
    - A total of 26 patients were unblinded during the study.
- Lilly personnel were blinded to treatment assignments.
- To maintain the blinding of the study to Lilly personnel and to permit interim analyses to be conducted without affecting study integrity, a statistician from an external statistical services organization was contracted by Lilly to prepare interim reports and present interim results to the DSMB.
- Treatment codes were kept by [redacted] At the conclusion of the study, after final data inconsistencies were resolved and data lock had occurred, treatment codes were provided to Lilly. b(4)

Prior and Concomitant Therapy (9.4.7)

- Patients may have received standard **CV medication** during the study at the discretion of their primary physician(s), who provided all medical care.
- The addition of raloxifene to the patient's regimen did not necessitate discontinuation of any drugs required for the usual treatment of the patient's coronary disease or CV risk factors.
- If the patient was taking **hormone replacement therapy (HRT)** (ie, estrogen or estrogen and progestin, excluding limited use of vaginal estrogen-containing formulations), other hormonal-type agents such as tibolone, or any SERM, she was required to discontinue study drug. If she later stopped HRT, study drug was allowed to be restarted as long as approximately 1 month had elapsed since the last dose of HRT, and she had not been experiencing severe hot flashes. If she stopped other hormonal-type agents, such as tibolone, or any SERM, she was allowed to restart study drug immediately.
- **Oral and transdermal estrogen-containing products** were contraindicated because the combination of oral raloxifene and oral or transdermal estrogen had not been studied. Limited use of vaginal estrogen-containing formulations was allowed, as it was assumed that use of these products would result in less systemic absorption compared to oral or transdermal products.
- **Cholestyramine therapy** (or treatment with any lipophilic anion-exchange resin) initiated after randomization was strongly discouraged; use of an alternative agent, if needed, was strongly encouraged. If, however, cholestyramine was used, discontinuation of study drug was not required. Cholestyramine (or other lipophilic anion-exchange resins) was not advised because it interrupts the enterohepatic cycling of raloxifene and results in a significant decrease in systemic exposure to raloxifene.
- While the simultaneous administration of **warfarin** and raloxifene is not contraindicated, co-administration of raloxifene and warfarin resulted in a 10% decrease in prothrombin time in a single-dose study (Evista package insert 2003). Therefore, patients in this trial who were taking warfarin (or coumadin derivatives) were asked to advise their primary

care physician that they may be taking raloxifene and that this may alter their prothrombin time. As raloxifene is more than 95% protein-bound, there is a potential for a drug-drug interaction with similar protein-bound drugs. However, to date, no such interactions have been identified.

- **Prior therapy** with estrogen, estrogen plus progestin, oral contraceptives, raloxifene, tamoxifen, other SERMs, or tibolone was recorded on the case report form (CRF) at Visit 2.
- The use of **lipid-lowering agents** was queried at each visit. All medications (other than study medication) taken during the study were recorded on the concomitant medication CRF.

Treatment Compliance (9.4.8)

- At each clinic visit, patients were to return all unused study tablets so that the remaining tablets could be counted and recorded for compliance calculation. All unused study medication was returned to [redacted] or to the local Lilly affiliate for destruction. b(4)
- Section 9.7.1.9 describes the calculation of treatment compliance and the definition of a compliant patient.

Efficacy and Safety Variables (9.5)

Endpoint Processing and Adjudication Procedures (9.5.1)

- The processing of all study endpoints and serious adverse events (SAEs) was managed by the Serious Adverse Event Endpoint Coordination Team (SECT), a team comprised of Lilly personnel who were blinded to treatment assignment.
- Study endpoint events were not solicited in this trial but were to have been reported to the sponsor within 24 hours from the time the site became aware of an event.
- The investigator-reported primary and secondary endpoints of the study were adjudicated by committees blinded to treatment assignment. After a study endpoint event was reported by the investigator, the investigative site was requested to provide to the sponsor an endpoint package containing the appropriate Investigator Summary Form and supporting documents necessary for adjudication of the study endpoints. The sponsor then forwarded the endpoint package to members of the appropriate adjudication committee for their independent assessment.
- The adjudication committee returned the completed adjudication form to Lilly.
- Disagreements between adjudication committee members were reconciled using predefined processes. Final adjudicated case decisions were recorded in the clinical trial database.

Invasive Breast Cancer [Breast Cancer Primary Endpoint] (9.5.1.1)

- **Baseline mammograms** were obtained at randomization (or within 12 months prior to randomization)
- **Follow-up mammograms** were obtained every 2 years 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.

- 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 (immunocytochemical assay).

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 Lilly.

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 (ie, evident on the baseline mammogram) or was new (ie, 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 sponsor notified the investigator of any

information included in the endpoint package which could be pertinent to the care of the patient.

All Breast Cancer [Secondary Endpoint] (9.5.1.1.1)

All breast cancer cases included invasive and 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.

Coronary Death, Nonfatal (Including Silent) MI, and Hospitalized Acute Coronary Syndrome Other Than MI (Coronary Primary Endpoint) (9.5.1.2)

The **Coronary Primary Endpoint Committee (CPEC)** was comprised of 10 cardiologists, including the Chair, none of whom were employees of Lilly. 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 was not measured, then 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 1 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 (ie, family, friends, neighbors, physicians),

- 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,
- A CV cause not otherwise specified on the CRF.

ECG Evaluation for Silent MI (9.5.1.2.1)

- Electrocardiograms were performed at baseline, at the 2- and 4-year visits, and at the final visit (if not performed within the 3 months prior to the final visit) for the purpose of identifying silent MIs.
- The sponsor sent the original ECG paper tracings to the central ECG laboratory. The central ECG medical reviewer (ie, cardiologist), who was not an employee of Lilly, interpreted each tracing, compared it to prior tracings, and completed the ECG Evaluation Form (ECGEF).
- For each ECG, the reader interpreted the ECG as normal or abnormal.
- If abnormal, the reader further classified the findings into the following major categories:
  - Definite Q-wave MI,
  - Pathologic ST-T depression
  - Conduction disturbances,
  - Atrial fibrillation or flutter,
  - Ventricular hypertrophy

- For post-baseline ECGs, only abnormal parameters which changed from the previous scheduled ECG were recorded. For example:
  - If a patient had a **normal ECG at baseline** and the next scheduled ECG was **abnormal**, then the recorder checked abnormal and the corresponding abnormal finding(s) on the ECGEF.
  - If a patient had an **abnormal ECG at baseline** and the next scheduled ECG was also **abnormal**, then the recorder checked abnormal on the ECGEF.
    - If the **same abnormality** was present as before, the reader did not tick the abnormal finding again (ie, only abnormal was ticked on the ECGEF).
    - If a **new abnormality** was identified, the reader ticked only the new abnormality on the ECGEF (ie, only the parameter(s) which changed from the previous tracing).
  - If a patient had either a normal or abnormal ECG at baseline and the subsequent ECG was normal, then the recorder checked normal on the ECGEF.

- If a new, definite Q-wave MI was identified on an ECG tracing, in comparison with prior ECG(s), the medical reviewer denoted this accordingly on the ECGEF. SECT then reviewed the clinical trial database to determine if an MI had been reported by the respective investigative site during the 2 years between the current and prior ECGs. If no event had been reported, SECT submitted a Feedback Form to the investigator inquiring if the investigator wanted to report this finding as an MI. If the investigator elected to report the event as a new MI, an endpoint package was assembled and forwarded to the CPEC for adjudication (Section 9.5.1.2). If the investigator elected not to report the event as a new MI, SECT assembled an endpoint package containing the ECG tracings and the Feedback Form and submitted this only to the CPEC Chair. The Chair's final decision on the presence or absence of a nonfatal (including silent) MI was documented in the clinical trial database.
- If a new MI was identified by the medical reviewer and the investigator had reported an event of MI in the 2 years preceding the current ECG, and the event had been adjudicated to be an MI, SECT assembled an endpoint package inclusive of this information, the recent ECG tracings, and the ECG Evaluation Form, and sent this to the CPEC Chair for reconciliation. The Chair's final adjudication decision was documented in the clinical trial database.
- Criteria for the definition of the ECG components are described in detail in Schroeder R and Schuren KP, *Praktische EKG-Auswertungg, DifferentialdiagnostischesTabellarioum.*

Secondary Endpoints (9.5.1.3)

The secondary endpoint of all breast cancer was adjudicated by the BCEC (Section 9.5.1.1.1). All other secondary endpoints were adjudicated by the **Secondary Endpoint Committee (SEC)**, comprised of the SEC Chair, the VTE Endpoint Committee Chair and members, and the Stroke Endpoint Committee members, none of whom were employees of Lilly.

Death

- All investigator-reported ***coronary deaths*** were adjudicated by the CPEC (Section 9.5.1.2).
- All investigator-reported **CV *non-coronary deaths*** (Section 9.5.1.2) and **non-CV deaths** were adjudicated by the SEC Chair.
- For each investigator-reported **CV *non-coronary death*** and **non-CV death**, an endpoint package was sent to the SEC Chair for review. Based on available clinical information (eg, discharge summary, death certificate), the cause of death was attributed to a non-CV etiology when evidence surrounding death was secondary to one of the following: any cancer, breast cancer, accidental/suicide/homicide, or a non-CV cause not otherwise specified on the CRF.
- If a death was deemed "not adjudicatible" it was (by default) classified as a non-cardiovascular death. However, an adjudicator could denote on the CRF that the death was "not adjudicatible" and still broadly classify the death as a coronary, non-coronary,

or non-cardiovascular death on the CRF, if the adjudicator felt there was enough information to do so. If a "not adjudicatable" death was broadly classified as a coronary death, that coronary death was not included in the coronary primary endpoint time to event analysis but was included in the all-cause mortality analysis.

Stroke

All investigator-reported strokes were adjudicated by the Stroke Endpoint Committee, composed of two neurologists. For each investigator-reported stroke, an endpoint package was sent to the committee for review. Based on available clinical information, the committee classified the type of stroke as hemorrhagic, ischemic, or undetermined with clinical features suggestive of a stroke and, if possible, classified the type of stroke according to pathogenesis. A consensus decision was reached between the two neurologists, and their mutual decision was documented in the clinical trial database.

- **Stroke** was defined as the rapid onset of a persistent neurologic deficit, attributed to an obstruction in cerebral blood flow and/or cerebral hemorrhage that is not due to trauma, tumor, infection, or other certain etiology. The deficit must have lasted more than 24 hours, unless death occurred or there was a demonstrable lesion compatible with an acute stroke on head computed tomography or magnetic resonance imaging scan.
- **Transient ischemic attacks (TIAs)** were not a study endpoint and were collected as AEs.

VTE

All investigator-reported VTEs were adjudicated by the VTE Endpoint Committee, composed of 3 members and a Chair, all with expertise in venous thromboembolism. For each investigator-reported VTE, an endpoint package was sent to the committee members for review.

- Based on available information, the committee classified the type of VTE as deep vein thrombosis (DVT), pulmonary embolism (PE), intracranial thrombosis, or other as specified.
- If two adjudication committee members agreed that a VTE had occurred, the case decision was documented in the clinical trial database.
- If there was no agreement between at least two members, the endpoint package was forwarded to the Chair for review and final decision. The Chair's final decision was documented in the clinical trial database.

A DVT or PE was diagnosed using Modified World Health Organization (WHO) criteria (WHO 1995).

- For a **definitive diagnosis of DVT**, supporting documents were to include either:
  - a report of a Doppler study showing venous obstruction, or
  - A report of a venogram showing venous obstruction.
- For a **definitive diagnosis of PE**, supporting documents were to include either
  - a report of a pulmonary angiogram showing pulmonary embolus, or
  - a report of a ventilation-perfusion scan showing high probability for pulmonary embolism.

When diagnostic testing results were unavailable, a probable diagnosis of VTE could be based on the clinical diagnosis alone.

Fracture, myocardial revascularization, non-coronary arterial revascularization and lower extremity amputation, and all-cause hospitalization

A Lilly designee (physician or non-physician) completed the "Adjudication Form" for these respective secondary endpoints if appropriate documentation was available, if there were no uncertainties in interpreting the document, and if the report provided supporting evidence that the event of interest occurred or was performed. If the Lilly designee was unable to adjudicate the event, the case was sent to the SEC Chair for final decision.

- **Fractures** were to have been documented with a radiology report or equivalent document.
- **Myocardial revascularization**, defined as either coronary artery bypass surgery or catheter-based coronary revascularization, was to have been documented by a procedure report or equivalent document.
- **Non-coronary arterial revascularization** and **lower extremity amputation** were to have been documented by a procedure report or equivalent document.
- **Hospitalization** was to have been documented by a discharge summary or equivalent document. The patient must have been in the hospital at least 24 hours. The primary reason for hospitalization was to have been clearly specified by the investigator.

Biochemical Markers of CV Risk (9.5.1.4)

Biochemical markers of CV risk included serum lipids and fibrinogen.

- **Serum lipids** (fasting total cholesterol, LDL-C, HDL-C, and triglycerides) were assessed for all patients at baseline, Year 1, Year 5, and at final visit.
- **Fibrinogen** was assessed in a subset of patients at baseline, Year 1, Year 5, and at final visit.

Safety Measures Assessed (9.5.2)

The following safety measurements were collected at the times shown in the study schedule (Table xxx):

Physical Examination

- A physical examination was performed at baseline, at each annual visit, and at final visit.

Vital Signs and Physical Findings

- Vital signs (heart rate, systolic and diastolic blood pressure) and physical findings (height and weight) were collected at patient screening, randomization, at each annual visit, and at final visit.

Adverse Events

- Adverse events were collected at every visit—regardless of relationship to the study drug.

- These events were captured as actual terms and coded to Medical Dictionary for Regulatory Activities (MedDRA) terms by blinded Lilly personnel.

Concomitant Therapies

- Concomitant therapies taken during the study were recorded on the CRF.

Chemistry Tests

- Total bilirubin, AST, blood urea nitrogen (BUN), serum creatinine, and fasting glucose were measured at randomization and annually thereafter, and at final visit.
- Hemoglobin A1c (HbgA1c) was assessed at randomization, Year 2, and final visit.

Study Drug Plasma Samples

- Two blood samples were drawn at least 1 hour apart at the Year 1 and Year 2 visits for pharmacokinetic analyses in patients enrolled at a subset of investigative sites.

Endpoint measures in this study

ECGs

- ECGs were performed at randomization, Year 2, Year 4, and final visit (if not performed within the 3 months prior to the final visit).

Mammograms

- Mammograms were obtained at randomization, or within 12 months prior to randomization, and every 2 years thereafter throughout the study; the maximum number of mammograms performed for each patient during the study was not to exceed 4.

Clinical Breast Exams

- Clinical breast exams were performed at baseline and every 2 years thereafter throughout the study except at final visit.

Breast Cancer Follow-up Information

- For patients who developed breast cancer during the study, follow-up information was obtained at all subsequent visits.

Appropriateness of Measurements (9.5.3)

- Efficacy assessment methods used in this protocol have been described in the literature and are generally regarded as reliable, accurate, and relevant.

Reviewer Comments: *the reviewer agrees with the above comments.*

Primary Efficacy Variable(s) (9.5.4)

- The primary efficacy variables for the coronary primary endpoint were clinical symptoms, changes in cardiac enzymes or markers, and ECG findings.
- The primary efficacy variable for the breast cancer primary endpoint was the pathology report.

Bioanalytical and Pharmacokinetic Methods (9.5.5)

Bioanalytical Methods (9.5.5.1)

Study drug plasma samples for determination of raloxifene concentration were collected from patients enrolled at a subset of investigative sites in the US and Canada and were analyzed at [redacted] located in [redacted] b(4)

- The samples were analyzed for raloxifene using a validated liquid chromatography with tandem mass spectrometry (LC/MS/MS) method.
- The lower limit of quantification was 0.050 ng/mL, and the upper limit of quantification was 5.00 ng/mL. Samples above the limit of quantification were diluted and reanalyzed to yield results within the calibrated range.
- The inter-assay accuracy (% absolute relative error) during cross validation was δ15.1%. The inter-assay precision (% relative standard deviation) during cross validation was δ3.59%.
- Raloxifene was stable for at least 1098 days when stored at approximately -20°C.

Pharmacokinetic Methods (9.5.5.2)

Following the conclusion of GGIO, the pharmacokinetics of raloxifene in plasma were assessed for the 2-year data.

- Graphical visualization and descriptive statistical analyses of the LY139481 concentration data using S-PLUS were the primary techniques of data evaluation.
- Population pharmacokinetic modeling using the nonlinear mixed effects modeling program NONMEM was not performed as outlined in the GGIO protocol since the descriptive statistical analysis results are consistent with those from prior raloxifene Phase 3 studies (H3S-MC-GGGF, H3S-MC-GGGG, H3S-MC-GGGH, and H3S-MCGGGK).

Data Quality Assurance (9.6)

To ensure accurate, complete, and reliable data, Lilly and/or its representatives did the following:

- Provided instructional material to the study sites, as appropriate. Specific details relating to study procedures were outlined.
- Conducted start-up training sessions as needed to instruct the investigators and study coordinators; these sessions gave instructions on the protocol, the completion of the CRFs, GCP/ICH guidelines, and study procedures.
- Made periodic monitoring visits to the study site.
- Were available for consultation and remained in contact with the study site personnel by mail, e-mail, telephone, and/or fax.
- Reviewed and evaluated CRF data, and used standard computer edits to detect errors in data collection.
- Conducted a quality review of the database.

- A central laboratory [redacted] was used to maintain consistency of methods and to combine laboratory data across study sites and/or across studies (See Appendix 16.1.10 for reference ranges). b(4)

- To ensure the safety of participants in the study and to ensure accurate, complete, and reliable data, the investigator kept records of laboratory tests, clinical notes, and patient medical records in the patient files as original source documents for the study. If requested, the investigator provided Lilly, applicable regulatory agencies, and/or applicable ERBs, direct access to original source documents.
- Lilly or its representative periodically checked a sample of the patient data recorded against source documents at the study site.
- The study was subject to independent audit by staff of the Medical Quality Assurance Department or its contractors. Investigators were given notice before an audit occurred. The audit certificate listing all audit types for the study is included in Appendix 16.1.8 of the study report.
- The following measures were taken for drug accountability:
  - Drug accountability was emphasized at start-up meetings, in EC communications to the investigators, and at investigator meetings conducted during the course of the trial.
  - A drug accountability form was provided in the clinical trial records binder or similar file.
  - The monitoring plan specified requirements for checking drug accountability.

Statistical Methods and Determination of Sample Size (9.7)

Statistical and Analytical Plans (9.7.1)

- This Statistical Analysis Plan (SAP) served to define all planned statistical analyses for the Raloxifene Use for The Heart (RUTH; Study H3S-MC-GGIO [GGIO]).
- Original CSR Section 9.8.2 addresses the changes made to the planned statistical analyses after unblinding of the reporting databases.

9.7.1.1. General Considerations

All analyses were performed according to the intent-to-treat (ITT) principle, unless otherwise specified.

- In an ITT analysis, data is analyzed by the treatment group to which a patient is randomized, even if the patient is inadvertently randomized, does not take the assigned treatment, does not receive the correct treatment, or otherwise does not follow the protocol or procedures. In other words, the analysis population, which was referred to as the ITT population in this document, included all randomized patients.
- The ITT population was labeled as "All Randomized Patients" in the report tables, figures, and patient data listings.

A patient was considered to have completed the protocol if either of the following criteria was met:

- the patient participated until the study was concluded (ie, a visit was conducted for the patient on or after 01 March 2005); *or*
- the patient died during the study.

The primary and secondary endpoints of the study consisted of adjudicated events that met the criteria set forth for these endpoints. The primary endpoints included the following events:

- Coronary primary endpoint: the combined endpoint of coronary death, nonfatal (including silent) MI, and hospitalized ACS other than MI.
- Breast cancer primary endpoint: invasive breast cancer.

Significance levels for the final analysis of the primary endpoints

- the treatment effect on the coronary primary endpoint was tested at a two-sided significance level of 0.0423
- the breast cancer primary endpoint was tested at a two-sided significance level of 0.008
- For all other analyses, the treatment effects were tested at a two-sided significance level of 0.05; interaction effects were tested at a significance level of 0.10.

The following tests were used for treatment group comparisons of categorical variables, and will be referred to as a "Chi-square test" in this document.

- Pearson's Chi-square test was performed if the overall number of patients in each category was equal to or greater than 10;
- Fisher's Exact test was conducted when the number of patients in each category was equal to or greater than 5, with at least one category having less than 10 patients;
- No statistical comparisons were made and only counts and proportions were reported when the total number of patients in any category was less than 5.

The following rules were used to handle ties:

- In a Cox proportional hazards regression model, an exact method (Kalbfleisch and Prentice 1980; DeLong et al. 1994) was used to handle tied time-to-event;
- In a ranked one-way analysis of variance (ANOVA) model, a mean score method was used to handle tied ranks (SAS PROC ANOVA default option, SAS 1999).
- In all analyses, "baseline" was defined as the last non-missing observation at or before the randomization visit (Visit 2). "Post-baseline" was defined as any observation recorded at or following Visit 3.

Exploratory subgroup analyses were performed to assess consistency of the treatment effects:

- Subgroup analyses were performed for the primary endpoints and for the secondary endpoint of stroke.
  - The subgroup variables included clinically relevant risk factors.
- Additional subgroup analyses were performed for baseline characteristics that were found to be unbalanced between treatment groups.

Sponsor comments: *There were no adjustments for multiplicity in these subgroup analyses and caution is necessary in the interpretation of subgroup analyses since subgroup analyses were generally underpowered.*
Reviewer Comments: *Agree with the above comments.*

Censoring Date (9.7.1.1.1)

For the **time-to-event analyses** (except for the mortality analyses), a censoring date was defined for each patient.

- If a patient completed the protocol, the censoring date was the final visit date or death date.
- If a patient failed to complete the protocol, the censoring date was the date of the last visit at which study information was obtained (for example, AE, laboratory measurements, mammogram, ECG, blood pressure, or weight). This censoring date was used in all analyses (except the mortality analyses) to censor patients who had not experienced the event of interest.

For the **time-to-event analyses for mortality**

- the censoring date for a patient who completed the protocol was as defined above.
- the censoring date for a patient who failed to complete the protocol was defined differently than as stated above:
  - Vital status may have been ascertained at a visit in the absence of collection of any other study information (for example, a family member reported that the patient is alive, but no further information is obtained). Therefore, if a patient failed to complete the protocol, the censoring date for mortality analyses was defined as:
    - the date of the final visit if she was known to be alive at the time of study discontinuation, or
    - the date of the last reported contact with the patient if her mortality status was unknown at the final visit.

Time to Event (9.7.1.1.2)

- For each patient, time-to-event for an event of interest was the number of days between the date of randomization and the onset date of the event plus one day if she experienced the event *or* the number of days between the date of randomization and the censoring date plus one day if she did not experience the event.
- If a patient experienced multiple events (for example, multiple strokes), the date of the first event was used, unless otherwise specified.

Patient-Years of Follow up (9.7.1.1.3)

- **Patient-years of follow up for an event** of interest were calculated for each patient as the time to event divided by 365.25.
- The **total patient-years of study follow up** were calculated for each patient as the number of days between the randomization date and the censoring date plus one day divided by 365.25.

Adjustments for Covariates (9.7.1.2)

- A sensitivity analysis for each of the primary endpoints was performed, with adjustment for clinically relevant risk factors and unbalanced baseline characteristics. Details are addressed in Sensitivity Analyses for the breast cancer primary endpoint and in Sensitivity Analyses for the coronary primary endpoint.

Handling of Missing Data (9.7.1.3)

- All analyses followed the ITT principle unless otherwise specified.
- When change from baseline to endpoint was assessed, only patients with a baseline and at least one post-baseline measurement were included in the analysis.
- The last-observation-carried-forward (LOCF) principle was applied when a measurement was missing. LOCF was applied independently within the baseline period and post-baseline period.

An incomplete death date was imputed as described below:

- If only the day of the death date was missing, the day was imputed as follows:
  - If the date of the last reported contact for the patient fell in the same month and year as the death date where the day was missing, the day was imputed to fall halfway between the last reported contact and the end of the given month (for example, if an incomplete death date was 04- -2004, and the date of the last reported contact was 04-22-2004, the death date was imputed as 04-26-2004).
  - If the date of the last reported contact for the patient occurred before the reported month and year of the death date, the day was imputed as the 15th of the reported month (for example, if an incomplete death date was 04- -2004, and the date of the last reported contact was 03-26-2004, the death date was imputed as 04-15-2004).
- If both the month and day of the death date were missing, the month and day were imputed as follows:
  - If the date of the last reported contact for the patient fell in the same year as the incomplete death date, the death date was imputed as the 1st of the month falling halfway between the month of the last reported contact and the end of the year (for example, if an incomplete death date was - -2004, and the date of the last reported contact for the patient was 06-22-2004, the death date was imputed as 09-01-2004).
  - If the year of the last reported contact date for the patient occurred before the year of the incomplete death date, the death date was imputed as June 30th of the reported year (for example, if an incomplete death date was - -2004, and the date of the last reported contact was 06-22-2003, the death date was imputed as 06-30-2004).
- If the day, month, and year of the death date were missing, the date remained as missing.

An incomplete endpoint event (ie, primary endpoints or secondary endpoints) date was imputed as outlined below:

- If only the day of the event date was missing, the day was imputed as the 15th of the reported month.
- If both the month and day of the event date were missing, the month and day were imputed as June 30th of the reported year.
- In the case that the imputed event date fell after the patient's censoring date, the incomplete event date was imputed as the censoring date (for example, if an incomplete onset date of an event was - - 2004, and the patient's censoring date was 05-22-2004,

then the onset date of the event was imputed as 05-22-2004 rather than as the date of 06-30-2004 as imputed by following the procedure stated in the second bullet above).

- If the day, month, and year of the event date were missing, the date remained missing.

Multicenter Studies (9.7.1.4)

A total of **177 investigative sites** from **26 countries** participated in this study. Among these 26 countries, sites in Switzerland randomized 6 patients, and sites in Ireland randomized 22 patients. Due to the low number of patients in these two countries,
Switzerland was pooled with Germany, and Ireland was pooled with the United Kingdom, forming **a total of 24 countries for the purpose of analyses.** As shown in the Table xxx below the 24 countries were further pooled by **geographic region**. The pooled countries and regions were used in the sensitivity analyses for the primary endpoints.

Region	Country
North America	Canada, USA
Latin/South America	Argentina, Brazil, Mexico
Western Europe	Belgium, Denmark, Finland, Germany/Switzerland, France, Israel, Italy, Netherlands, Norway, Spain, Sweden, UK/Ireland
Eastern Europe	Czech Republic, Hungary, Poland, Russia
Africa	South Africa
Asia Pacific	Singapore, Taiwan

Multiple Comparisons/Multiplicity (9.7.1.5)

For the two **primary endpoints**, the p-value for conclusion of statistical significance was adjusted to account for the planned interim analyses and for the multiplicity of the primary endpoints.

- At each of the three **coronary interim analyses**, the statistical guideline for a conclusion of significant evidence of efficacy was a two-sided test with significance level ≤ 0.0001. The final analysis was conducted at the ≤ 0.0423 level. This ensured an overall type I error rate less than 0.04234 for the coronary primary endpoint.
- The statistical guideline for a conclusion of significant evidence of efficacy for the single interim analysis for the **breast cancer primary endpoint** was a two-sided test with significance level ≤ 0.000001. The final analysis was conducted at the ≤ 0.008 significance level. These adjustments maintained an overall type I error rate of less than 0.05 for the study.

No adjustments were made for multiple comparisons of **secondary endpoints.**

Efficacy Subsets (9.7.1.6)

Sensitivity analyses were performed for the primary endpoints, the secondary endpoint of VTE, and the secondary endpoint of stroke **for a Per-Protocol (PP) population.**

- The **PP population** was a subset of the ITT population. It consisted of patients who met all inclusion criteria, did not fulfill any exclusion criteria, and were compliant with study drug treatment.
- The **PP population for breast cancer analyses** was based on the overall treatment compliance for breast cancer analyses.
    - A **breast cancer efficacy subset** was defined as randomized patients who were **60 years *or* older** at baseline. (Age has been identified as a risk factor for breast cancer. Women aged 60 years or older were eligible to participate in the P-1 study (Fisher et al. 1998))
- The **PP population for VTE analyses** was based on the overall treatment compliance for VTE events.
- The **PP population for coronary primary endpoint *and* its individual components**, as well as other endpoints, was based on the overall treatment compliance.
    - Two **coronary efficacy subsets** were defined based on the baseline CV risk assessment.
        - The **primary prevention population:** randomized patients with no documented CHD, but at increased risk for CHD at baseline.
        - The **secondary prevention population:** randomized patients with documented CHD at baseline.

Patient Disposition (9.7.1.7)

According to the study design, patients were followed for the duration of the study, regardless of their compliance with study drug or protocol procedures. If a visit was not conducted, the reason for a non-conducted visit was collected as either "patient decision," "patient moved," "lost to follow-up," or "other."

Counts and proportions of each **reason for study discontinuation** were summarized for each treatment group. Treatment group differences were compared using a Chi-square test. The reasons for study discontinuation were defined as follows:

- Death.
- Withdrawal of consent: a patient was considered to have withdrawn consent if her final visit occurred prior to the conclusion of the study and was either a) a conducted visit, or b) a non-conducted visit due to reason of "patient decision."
- Inadvertent summarization: a patient was considered to be inadvertently summarized if her final visit occurred prior to conclusion of the study, and the visit was not conducted for reason other than "patient decision" (in other words, a patient's final visit was erroneously completed before the conclusion of the study).
- Lost to follow-up: a patient was considered to be lost to follow-up if her final visit occurred at conclusion of the study, but no study information was obtained from the patient.
- Study was concluded: a patient completed the protocol, ie, her final visit occurred on or after 01 March 2005.

Analyses of study drug discontinuations: as per the protocol, investigators may have either temporarily or permanently discontinued study drug in a given patient. It would have been informative to assess reasons for temporary or permanent study drug discontinuation. However, data was not collected to ascertain whether the discontinuation was intended to be permanent or temporary. Therefore, the following analyses were performed for study drug discontinuation:

- Reasons for study drug discontinuation were compared between treatment groups on an overall basis. In the case where a patient discontinued study drug and then resumed at a later visit, only the reason given at her final discontinuation from study drug was used in this analysis. Counts and proportions were summarized for each treatment group. Treatment group differences were assessed using a Chi-square test.
- Reasons for study drug discontinuation, which may have represented either a temporary or permanent discontinuation, were also summarized for each treatment group on a by-visit basis. Counts and proportions of patients participating in the study until the last visit at which study information was obtained were reported for each treatment group on a by-visit basis.

The total **patient-years of study follow up** were summarized by treatment group for the ITT population, the primary prevention population and the secondary prevention population. The mean, standard deviation, median, minimum, and maximum were reported within each treatment group.

Change from the protocol

According to the protocol, the Cochran-Mantel-Haenszel (CMH) test, adjusted for the country of the investigative sites, was conducted in the comparison of study drug discontinuation. Due to the possibility that many country-by-treatment combinations may have had fewer than 5 study drug discontinuations; the CMH test result may have been misleading in this situation. Therefore, the Chi-square test replaced the CMH test in this statistical analysis.

Patient Characteristics (9.7.1.8)

The following baseline patient characteristics were summarized by treatment group:

- Demographic characteristics
- Breast cancer risk assessment characteristics
- CV risk assessment characteristics
- VTE risk assessment characteristics
- Biochemical markers of CV risk.

For **categorical variables**, counts and proportions were reported by treatment group. Treatment group differences were compared using a Chi-square test.
For **continuous variables**, descriptive statistics, including mean, standard deviation, median, minimum, and maximum were reported. Mean differences between treatment groups were assessed using an F-test from a one-way ANOVA.
For **lipids and fibrinogen**, a ranked one-way ANOVA was used to compare treatment group differences.

Change from the Protocol

According to the protocol, the CMH test, adjusted for country, was conducted in the comparison of treatment differences for categorical baseline characteristic variables.

- Because the number of countries involved was large (24), and the minimum number of patients randomized in a country was 98, using a Chi-square test to compare categorical baseline characteristics was similar to using the CMH test. Therefore, a **Chi-square test replaced the CMH test** in this statistical analysis.
- The protocol also specified using an *F*-test from a two-way ANOVA model, with adjustment for country, to compare treatment differences for continuous baseline variables. To maintain consistency with the comparison for categorical variables, and the proposed one-way ranked ANOVA for serum markers, safety analytes, and vital signs, a **one-way ANOVA model replaced the two-way ANOVA model** in this statistical analysis.

Treatment Compliance (9.7.1.9)

Patients may have stopped taking study drug for various reasons, but, in accordance with the study design, remained in the study for follow up. Study drug was dispensed at Visit 2, Visit 4, and every 6-month visit thereafter. No tablets were dispensed at Visit 3, which was an optional visit. Unless a patient discontinued from the study at Visit 3, Visit 4 was the first visit at which patients were required to return tablets; therefore, Visit 4 was the first visit for which treatment compliance could have been calculated.

Treatment compliance through Visit 4 was calculated using formula A, below:

$$\text{A)}\quad \frac{[\#\text{ of tablets dispensed at Visit 2} - \#\text{ of tablets returned at Visit 4}] * 100}{[\text{number of days between Visit 2 and Visit 4}]}$$

Treatment compliance for any visit beginning at Visit 5 and continuing through the last visit at which study information is obtained was calculated using formula B, below:

$$\text{B)}\quad \frac{[\#\text{ of tablets dispensed at Visit (N-1)} - \#\text{ of tablets returned Visit N}] * 100}{[\text{number of days between Visit (N-1) and Visit N}]}$$

If a patient discontinued from the study at Visit 3, treatment compliance was calculated at Visit 3 using formula B, above.

Treatment compliance at Visit N was 0% if no study drug was dispensed at Visit N-1 (for Visit 4, Visit N-1 is Visit 2). Treatment compliance was considered to be missing when the number of tablets returned at Visit N, following Visit N-1, was unknown (for Visit 4, Visit N-1 was Visit 2). Treatment compliance for the duration of the study, which was termed overall treatment compliance, was calculated as the average of all non-missing visit treatment compliance values from Visit 4 through the last visit at which patient study information was

obtained. For patients who discontinued from the study at Visit 3, overall treatment compliance was the same as that calculated at Visit 3. Based on this scheme, a patient's compliance may be >100%.

- Descriptive statistics (mean, standard deviation, median, minimum, and maximum) of treatment compliance were summarized for each treatment group. Mean compliance was compared between treatment groups using an *F*-test from a one-way ANOVA, with treatment in the model.
- A patient was considered to be treatment-compliant if her overall treatment compliance value was ≥ 70% and ≤ 120%. Counts and proportions of compliant patients were reported by treatment group; treatment group differences were assessed using a Chi square test.

Treatment Compliance for the Breast Cancer Analyses (9.7.1.9.1)

Treatment compliance calculations for the breast cancer PP sensitivity analyses took into account the protocol requirement that patients who were diagnosed with breast cancer must have immediately and permanently discontinued study drug.
Overall treatment compliance for patients diagnosed with breast cancer and adjudicated as such was calculated as the average of visit treatment compliance values up to and including the visit preceding the breast cancer onset visit (defined as the visit at which the breast cancer was reported). Treatment compliance calculated in this manner more accurately reflected treatment compliance for patients who were diagnosed with breast cancer. Descriptive statistics and statistical tests were performed.

Treatment Compliance for the VTE Analyses (9.7.1.9.2)

Treatment compliance calculations for the VTE PP sensitivity analyses took into account the protocol requirement that patients who were diagnosed with a VTE must have immediately and permanently discontinued study drug. Overall treatment compliance for patients diagnosed with a VTE and adjudicated as such was calculated as the average of visit treatment compliance values up to and including the visit preceding the VTE onset visit (defined as the visit at which the VTE was reported). Treatment compliance calculated in this manner more accurately reflected treatment compliance for patients who were diagnosed with a VTE. Descriptive statistics and statistical tests were performed.

Compliance with Study Procedures (9.7.1.10)

Electrocardiograms, mammograms, and breast examinations were performed according to the study schedule of events. Compliance with each of these procedures were calculated at every visit at which the procedure is scheduled by dividing the number of patients who completed the procedure by the total number of patients eligible for the procedure at that visit. Compliance with these study procedures was compared between treatment groups at the visits at which the procedure was scheduled using a Chi-square test. Counts and proportions of patients who completed scheduled procedures were reported for each treatment group.

Concomitant Medications (9.7.1.11)

Prior to data lock, a list of concomitant medication categories was created. Anatomical Therapeutic Chemicals (ATC) codes were used to classify concomitant medications into categories (for example, lipid lowering agents, anti-hypertensive). Appendix 16.1.9 of the study report presents category definitions. Counts and proportions of concomitant medication used within each category were summarized by treatment group, and compared using a Chi square test. Comparisons were made within the baseline period, and within the post-baseline period. Change from baseline in the proportion of usage within each concomitant medication category was reported for each treatment group. A CMH test was used to assess the trend of concomitant medication use from baseline to post-baseline.

Efficacy Analyses (9.7.1.12)

This study had two primary endpoints. All breast cancer analyses were presented first, followed by the coronary primary endpoint analyses, and then related CV analyses and other secondary endpoint analyses.

Statistical Analyses for the Breast Cancer Endpoints (9.7.1.12.1)

The breast cancer statistical analyses were based on adjudicated events, unless otherwise specified, and are presented in the following order:

- Breast cancer primary endpoint
- Breast cancer secondary endpoint and breast cancer invasive status
- Clinical breast examination and mammogram findings
- Breast cancer characteristics
- Subgroup analyses of the breast cancer primary endpoint
- Investigator-reported breast cancer
- Follow up treatment after breast cancer diagnosis

Statistical Analyses for the Breast Cancer Primary Endpoint (9.7.1.12.1.1)

Primary Analysis

A log-rank test, based on time-to-first invasive breast cancer, was performed to compare the survival functions between treatment groups; the p-value from the log-rank test will be reported. The log-rank test was conducted at the two-sided significance level of 0.008.

Secondary Analysis

A Cox proportional hazards regression model, based on time-to-first invasive breast cancer event, was fitted to estimate the hazard ratio and 95% confidence interval (CI) comparing treatment groups. Kaplan-Meier curves based on time-to-first invasive breast cancer event were generated for each treatment group.

Sensitivity Analyses

(1) The primary and secondary analyses were repeated for the **PP population** and for a **subset of patients who were 60 years *or* older**. Since it was expected that there was a small proportion of patients in this study who were under 60 years of age at baseline, no statistical comparison was made for this subset. For patients under 60 years of age at baseline, only counts and proportions of those diagnosed with invasive breast cancer were reported for each treatment group.
(2) According to the study design, randomization was stratified by investigative site. To evaluate the consistency of the treatment effect across geographical regions, the following analyses were conducted on the ITT population.

- Due to the anticipated low incidence of invasive breast cancer, a stratified analysis with stratification by investigative site was not feasible. Instead, a stratified analysis with stratification by region was performed. The analysis evaluated the treatment effect when baseline hazards may have differed between regions. Specifically, a stratified Cox proportional hazards regression model was fitted with treatment as a fixed effect, and region as a stratum. The hazard ratio and 95% CI from this model was reported.
- To evaluate the consistency of the treatment effect across regions, the number of patients diagnosed with invasive breast cancer was evaluated in each region. For regions with less than 5 events, only counts and proportions were reported for each treatment group. The treatment-region interaction was tested based on a likelihood ratio test. The likelihood ratio is the difference of the log-likelihoods from a full Cox model with treatment, region, and treatment-region interaction, and a reduced Cox model with treatment and region. If the interaction p-value was significant ($p<0.10$), a Cox proportional hazards regression model was fitted for each region. The hazard ratio and 95% CI were reported for each region, along with the interaction p-value. In addition, the treatment-region interaction was tested again for regions with at least 5 events to see if the sparse data caused the significant interaction.

(3) A Cox proportional hazards regression model was used to adjust for the clinically relevant risk factors and unbalanced baseline characteristics that were significant at 0.05 level. The hazard ratio of invasive breast cancer, comparing raloxifene to placebo and its 95% CI from this multivariate adjusted model, was reported. The analysis used the ITT population. Table xxx presents the list of clinically relevant risk factors. A final, multivariate adjusted model was determined by the following model-fitting procedure:
a) A univariate Cox model was fitted for each of the clinically relevant risk factors and unbalanced baseline characteristics.
b) A multivariate model, with covariates, which are significant in the univariate model, was fitted. A stepwise model selection method was used to determine the final model. In the stepwise model selection process, the significance level for entry in the model was 0.10, and for remaining in the model was 0.05.
c) Treatment was added to the final model, as determined by procedure (b), above.

Table xxx. Clinically relevant risk factors

Subgroup	Categories
Age (years)	≤65, >65
Race	Caucasian, All other races
Body mass index (kg/m²)	≤25, >25 and ≤30, >30
5-year predicted invasive breast cancer risk ≥1.66% (as estimated by modified Gail model)	Yes/No
Family history of breast cancer (mother/sister/daughter)	Yes/No
Prior use of estrogen only	Yes/No
Prior use of estrogen plus progestin	Yes/No
Prior hysterectomy	Yes/No
Prior ovariectomy	Yes/No

Summary Statistics

The following summary statistics are reported for each treatment group:

- Counts and proportions of patients who were diagnosed with breast cancer
- Patient-years of follow up for breast cancer events
- Incidence rate, calculated by dividing the number of patients who were diagnosed with breast cancer by the patient-years of follow up for breast cancer events

Absolute risk reduction (ARR) was calculated as the difference in cumulative incidence of invasive breast cancer between the two treatment groups at the end of the study period.

Statistical Analyses for the Breast Cancer Secondary Endpoint and Breast Cancer Invasive Status (9.7.1.12.1.2)

Statistical analyses were performed to determine the treatment effect on the breast cancer secondary endpoint, eg, all breast cancers—invasive and non-invasive, hormone receptor positive or negative, as listed below:

- All breast cancer
- Invasive breast cancer
    - Estrogen receptor positive (ER-positive)
    - Estrogen receptor negative (ER-negative)
    - Estrogen receptor (ER) status unknown
- Noninvasive breast cancer
    - Ductal carcinoma in situ (DCIS)
    - Lobular carcinoma in situ (LCIS)
- Invasiveness unknown

Summary Statistics were reported for all of the above breast cancer categories. Statistical analyses included:

1) Counts and proportions of patients who were diagnosed with the event of interest, summarized by treatment group,
2) Comparison of treatment groups using a log-rank test based on time-to-first event, and
3) Estimation of the hazard ratio and 95% CI comparing raloxifene to placebo using a Cox proportional hazards regression model, based on time-to-first event. Kaplan-Meier curves

were generated for each treatment group for all breast cancer and invasive ER-positive breast cancer. If the total number of events for a specific endpoint was less than 5, no statistical analysis was performed and only the counts and proportions were reported.

Statistical Analyses for the Clinical Breast Examination and Mammogram Findings (9.7.1.12.1.3)

Statistical Analysis for the Breast Examination Findings

Clinical breast examinations were performed at scheduled visits. Investigators reported their findings as either normal or abnormal. If the examination was judged abnormal, investigators reported the abnormal finding(s) as either clinically significant or not clinically significant.

- The number of breast examinations reported by investigators as normal or abnormal (clinically significant or not clinically significant) was summarized for each treatment group by visit at which breast examinations were scheduled.
- Between treatment group comparisons for abnormal breast examinations were made using a Chi square test.

Statistical Analyses for the Mammogram Findings

Mammograms were performed at scheduled visits. Based on the mammogram report, investigators classified mammogram findings as showing either no significant abnormality (ie, the mammogram was read as normal, or revealed a non-clinically significant abnormality that required either no follow-up, or follow-up at a normal screening interval), or a clinically significant abnormality that required follow-up (ie, an abnormality suggestive of a malignancy, requiring prompt follow-up). For those mammogram reports indicating a clinically significant abnormality requiring follow-up, except at the final visit, the investigator reported whether or not a malignancy was diagnosed (the specific diagnosis was to be reported as an AE).

- The number of mammogram reports indicating no significant abnormality and the number indicating a clinically significant abnormality (no malignancy or malignancy diagnosed) were summarized for each treatment group by visit when the mammograms were scheduled.
- Between-treatment group comparisons for clinically significant, abnormal mammogram findings were made using a Chi-square test.

Statistical Analysis for Mammogram Findings Change from Baseline

Only women who had a **baseline mammogram** indicating no significant abnormality were included in this analysis of **change from baseline**. For the purpose of this analysis, the most serious post-baseline mammogram finding was used. An abnormal mammogram resulting in the diagnosis of a malignancy was considered the most serious finding, followed by an abnormal mammogram resulting in a diagnosis other than malignancy, and finally a mammogram with no significant abnormality.

- Counts and proportions of patients with post-baseline mammogram findings of "no significant abnormality" or "clinically significant abnormality requiring follow-up" were

summarized by treatment group. Treatment group differences were assessed using a Chi square test.

- Among patients with a post-baseline mammogram finding of a "clinically significant abnormality requiring follow-up," counts and proportions of patients whose findings are classified as either "no malignancy" or "malignancy diagnosed" were summarized by treatment group.

Statistical Analyses of Breast Cancer Characteristics (9.7.1.12.1.4)

Patients who had at least one breast cancer event were included in this analysis. If a patient was diagnosed with recurrent breast cancer, the first diagnosed event was used.

- Counts and proportions of breast cancer events by breast cancer characteristics, ie, ER status, tumor type, tumor grade, tumor stage, tumor size, and pre-existing status, were reported for each treatment group.
- Incidence rates were also calculated by dividing the number of breast cancer events of each characteristic by person-years of follow up of all breast cancer events.

Subgroup Analyses of the Breast Cancer Primary Endpoint (9.7.1.12.1.5)

Subgroup analyses of the breast cancer primary endpoint used the ITT population to maintain the statistical power and the efficiency of the estimates of the variance and analyses were performed for each of the clinically relevant risk factors (Table xxx). Unbalanced baseline characteristics were also considered as subgroup variables.

- A Cox proportional hazards regression model was fitted with treatment, subgroup indicator variable, and treatment-subgroup interaction in the model.
- The hazard ratio and 95% CI for estimating treatment effect were reported for each subgroup, along with the interaction p-value.
- If the total number of events for a specific subgroup was less than 5, no statistical analysis was performed, and only the counts and proportions were reported.

Table xxx. List of Clinically Relevant Risk Factors for the Primary Breast Cancer Endpoint

Subgroup	Categories
Age (years)	≤65, >65
Race	Caucasian, All other races
Body mass index (kg/m^2)	≤25, >25 and ≤30, >30
5-year predicted invasive breast cancer risk ≥1.66% (as estimated by modified Gail model)	Yes/No
Family history of breast cancer (mother/sister/daughter)	Yes/No
Prior use of estrogen only	Yes/No
Prior use of estrogen plus progestin	Yes/No
Prior hysterectomy	Yes/No
Prior ovariectomy	Yes/No

Statistical Analysis of Investigator-Reported Breast Cancers (9.7.1.12.1.6)

According to the study design, all investigator-reported breast cancer events were submitted for adjudication. However, some events may not have met the adjudication criteria. Therefore, a sensitivity analysis was performed for all investigator-reported breast cancer events.

- Counts and proportions of patients with at least one investigator reported breast cancer event were summarized by treatment group.
- A log-rank test, based on time-to-first event, was performed to compare the survival functions between treatment groups.
- A Cox proportional hazards regression model, based on time-to-first event, was fitted to estimate the hazard ratio and 95% CI comparing treatment groups.

Statistical Analysis of Follow up Treatment after Breast Cancer Diagnosis (9.7.1.12.1.7)

Follow up treatment information was collected every 6 months following diagnosis of breast cancer for investigator-reported breast cancers. The follow up treatment information was classified according to treatment type: tamoxifen, chemotherapy, radiation therapy, other, or without follow up treatment.

- Counts and proportions of patients ever receiving a given therapy post-diagnosis of breast cancer were summarized by treatment groups.

Statistical Analyses for the Coronary Endpoints (9.7.1.12.2)

The statistical analyses of the coronary endpoints were based on adjudicated events, and are presented in the following order:

- Coronary primary endpoint (combined endpoint of coronary death, nonfatal [including silent] MI, and hospitalized ACS other than MI)
- Components of the coronary primary endpoint
- Other analyses of the coronary primary endpoint
- ECG findings
- Subgroup analyses of the coronary primary endpoint

Statistical Analyses for the Coronary Primary Endpoint (9.7.1.12.2.1)

Primary Analysis

Analysis of the coronary primary endpoint and its individual components was based on the time from randomization to the occurrence of the 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 occurs first. For example, if a patient experienced a nonfatal MI, and later suffered a coronary death, the time-to-first event would have been the time to the nonfatal MI.

- A log-rank test based on time-to-first event was performed to compare the survival functions between treatment groups, and the p-value from the log-rank test was reported. The log-rank test was conducted at the significance level of 0.0423 for the coronary primary endpoint.

Secondary Analysis

A Cox proportional hazards regression model based on time-to-first coronary primary event and its individual components was fitted to estimate the hazard ratio and 95% CI comparing treatment groups. Kaplan-Meier curves based on time-to-first coronary primary endpoint event were generated for each treatment group.

Sensitivity Analyses

(1) The primary and secondary analyses were repeated for the **PP population** and for the **primary** *and* **secondary prevention populations** for the primary coronary endpoint and its individual components.
(2) According to the study design, randomization was stratified by investigative site. To evaluate the consistency of the **treatment effect across countries**, the following analyses were conducted for the coronary primary endpoint in the ITT population.

- Due to the large number of investigative sites, a stratified analysis with stratification by investigative site was not feasible. Instead, a stratified analysis with stratification by country was performed. The analysis evaluated the treatment effect when baseline hazards may have differed among countries. Specifically, a stratified Cox proportional hazards regression model was fitted with treatment as a fixed effect and country as a stratum. The hazard ratio and 95% CI from this model were reported.
- To evaluate the consistency of the treatment effect across countries, the number of patients experiencing a coronary primary endpoint was evaluated in each country. Countries with less than 5 coronary primary endpoints were pooled with the country that had the second smallest number of events in the same region. The treatment-country interaction was tested based on a likelihood ratio test for the pooled countries. The likelihood ratio was the difference of the log-likelihoods from a full Cox model with treatment, country, and treatment-country interaction, and a reduced Cox model with treatment and country. If the interaction p-value is significant ($p<0.10$), a Cox proportional hazards regression model was fitted for each country. The hazard ratio and 95% CI were reported for each country, along with the interaction p-value.

(3) A Cox proportional hazards regression model was utilized to adjust for the clinically relevant **risk factors** and unbalanced baseline characteristics for the coronary primary endpoint. This adjusted analysis was performed for the ITT population. Table xxx lists the clinically relevant risk factors. A final multivariate adjusted model was determined by the following model fitting procedure:

(a) A univariate Cox model was fitted for each of the clinically relevant risk factors and unbalanced baseline characteristics.
(b) A multivariate model with covariates, which were found significant in the univariate model, was fitted. A stepwise model selection method was used to determine the final model. In the stepwise model selection process, the significance level for entry in the model was 0.10 and 0.05 for remaining in the model.
(c) Treatment was added to the final model, as determined by procedure (b) above.

- The hazard ratio and 95% CI from this multivariate adjusted model comparing raloxifene to placebo for the coronary primary endpoint were reported.

Summary Statistics

For the coronary primary endpoint and its individual components, the following summary statistics was reported for each treatment group:

- **Counts and proportions** of patients who experienced a coronary primary endpoint event, listed as the composite endpoint and as individual components;
- Patient-years of **follow up** for coronary primary endpoint and its individual components
- The **incidence rate**, calculated by dividing the number of patients who developed the event during the study period by the patient-years of follow up of the coronary events of interest

Absolute risk reduction was calculated based on the difference in cumulative incidence of coronary primary endpoint and its individual components between the two treatment groups at the end of the study period.

Statistical Analyses of the Coronary Primary Endpoint Components (9.7.1.12.2.2)

Survival analyses using time-to-first event, performed on the individual components of the coronary primary endpoint and combinations of these components, are listed below:

- Coronary death
- Nonfatal (including silent) MI
- Hospitalized ACS other than MI
- Nonfatal (including silent) MI and hospitalized ACS other than MI
- Coronary death and nonfatal (including silent) MI

Counts and proportions of patients who developed the endpoint of interest were summarized by treatment group.

- Kaplan-Meier curves of the individual components were generated for each treatment group. Survival functions for all of the above endpoints were compared between treatment groups using a log-rank test.
- A Cox proportional hazards regression model was used to estimate the hazard ratio and its 95% CI for the raloxifene group compared with the placebo group.

Statistical Analysis for Time to Most Serious Event of the Coronary Primary Endpoint (9.7.1.12.2.3)

Time-to-event analysis was used to compare the treatment groups in terms of time to most serious event of the coronary primary endpoint. Coronary death was considered the most serious event, followed by nonfatal (including silent) MI, with hospitalized ACS other than MI being considered the least serious. The survival functions were compared between the treatment groups using a log-rank test. A Cox proportional hazards regression model was used to estimate the hazard ratio and its 95% CI for the raloxifene group compared with the placebo group. Kaplan-Meier curves were generated for each treatment group.

Statistical Analyses of Multiple Coronary Events (9.7.1.12.2.4)

A patient may have experienced more than one coronary primary endpoint event during the study—for example, a patient may have experienced a hospitalized ACS, followed by a nonfatal (including silent) MI and followed later by a coronary death, or a patient may have experienced 3 nonfatal (including silent) MIs. In each case, the patients were considered as having multiple coronary primary endpoint events. To assess whether raloxifene therapy reduced the number of coronary events for a patient, an analysis accounting for multiple coronary primary endpoint events was performed. In this analysis, a patient who experienced multiple coronary primary endpoint events was considered as having recurrent coronary events. The Prentice, Williams, and Peterson Gap Time (PWP-GT) model (Prentice et al. 1981) was used to analyze recurrent coronary events. The hazard ratio with 95% CI was reported for time to the first event, second event, etc. The mean time between events was also reported by treatment group.

Change from the Protocol

- The protocol specified that both weighted (according to event seriousness) and un-weighted models were to be used to assess multiple events per patient. Only the un-weighted model was used in this analysis.
- The protocol proposed to use a Wei, Lin, and Weissfeld (WLW) model (Wei et al. 1989) to analyze the multiple event types comprising the coronary primary endpoint. The analysis above focuses on recurrent coronary primary endpoint events irrespective of the type of event. Therefore, analysis of multiple event types was not performed.
- The protocol also proposed to use different models, such as the WLW model (Wei et al. 1989), the PWP-GT model (Prentice et al. 1981), and the Anderson and Gill (AG) model (Andersen and Gill 1982), to analyze the multiple events without regard to types. Published results in the statistical modeling literature have identified the PWP-GT model to be more suitable for analyzing recurrent events (Kelly and Lim 2000). Therefore, analyses of recurrent events using the WLW model and the AG model were not performed.

Statistical Analysis for Lag Time Effect of Raloxifene on the Coronary Primary Endpoint (9.7.1.12.2.5)

In the protocol design, **a raloxifene treatment benefit lag of 9 months** for the coronary primary endpoint was assumed. To assess this possibility, the counts of patients who have had at least one **coronary primary endpoint event during the first 9 months** after randomization (first period), and **after the first 9 months** (second period), were reported for each treatment group. Since the interest was in the occurrence of the first event for a patient, analysis for the second period excluded all patients who had an event during the first period.

- A Chi-square test was used to assess treatment group differences within the first period and the second period.

Statistical Analyses of ECG Findings (9.7.1.12.2.6)

Analyses of ECG findings used the results of an **independent central ECG assessment**. If the ECG tracing was assessable, the findings were designated as either normal or abnormal.

- Abnormal ECG findings were categorized: definite Q-wave MI (anterior, inferior, or posterior), pathologic ST-T depression, conduction disturbances (left bundle branch block [LBBB], right bundle branch block [RBBB], left anterior hemi block [LAH], left posterior hemi block [LPH], atrial-ventricular [AV] block > 1st degree, and other), atrial fibrillation or flutter, or ventricular hypertrophy (with or without strain).
- For all patients with abnormal ECG findings, the findings were analyzed independently by category. For each category, counts and proportions were summarized for each treatment group at each scheduled ECG visit, and compared using a Chi-square test.
- For all patients with normal baseline ECGs, counts and proportions of those who develop an abnormal post-baseline ECG finding were summarized for each category above.
- Treatment group differences within each category were assessed using a Chi-square test.
- The same analysis was performed for the primary and secondary prevention populations.

Subgroup Analyses of the Coronary Primary Endpoint (9.7.1.12.2.7)

Subgroup analysis of the coronary primary endpoint used the ITT population to maintain the statistical power and the efficiency of the estimates of the variance. Subgroup analyses were performed for each of the clinically relevant risk factors (Table xxx). Unbalanced baseline characteristics were also considered as subgroup variables.

- Cox proportional hazards regression models were fitted with treatment, subgroup indicator variable, and treatment-subgroup interaction in the model.
- The hazard ratio and 95% CI for estimating treatment effect were reported for each subgroup, along with the interaction p-value.
- If the total number of events for a specific subgroup was less than 5, no statistical analysis was performed, and only the counts and proportions were reported.

Table xxx. List of Clinically Relevant Risk Factors for the Coronary Primary Endpoint

Subgroup	Categories
Age (years)	≤65, >65 and ≤70, >70
Race	Caucasian, All other races
Body mass index (kg/m²)	≤25, >25 and ≤30, >30
Prior myocardial infarction	Yes/No
Prior angina pectoris with documented coronary disease	Yes/No
Prior CABG or PCI	Yes/No
Lower extremity arterial disease at baseline	Yes/No
Diabetes mellitus at baseline	Yes/No
Current smoker at baseline	Yes/No
Hypertension at baseline	Yes/No
Hyperlipidemia at baseline	Yes/No
Cardiovascular risk score at baseline	≤5, >5 and ≤9, >9
HMG-CoA reductase inhibitor use at baseline	Yes/No
Aspirin use at baseline	Yes/No
Beta-blocker use at baseline	Yes/No
Calcium channel blocker use at baseline	Yes/No
ACE inhibitor or ARB use at baseline	Yes/No
Diuretic use at baseline	Yes/No

Abbreviations: ACE = angiotensin converting enzyme; ARB=angiotensin receptor blocker; CABG = coronary artery bypass graft; HMG-CoA = hydroxymethylglutaryl-coenzyme A; PCI = percutaneous coronary intervention.

Statistical Analyses for Secondary Endpoints (9.7.1.12.3.)

Statistical analyses of secondary endpoints are presented in the following order:

- Stroke
- VTE
- Secondary CV endpoint combinations:
  - CV death, nonfatal (including silent) MI, hospitalized ACS other than MI, or stroke
  - CV death, nonfatal (including silent) MI, hospitalized ACS other than MI, stroke, or myocardial revascularization
- All-cause mortality
- Revascularization (myocardial and non-coronary arterial) and non-traumatic lower extremity amputation
- Fracture
- All-cause hospitalization, including hospitalized ACS

For the above secondary endpoints, if the total number of events for a specific endpoint was less than 5, no statistical analysis was performed, and only the counts and proportions were reported.

Statistical Analyses for the Stroke Endpoint (9.7.1.12.3.1.)

The effect of raloxifene on the incidence of all strokes, and the following categories of stroke, were determined:

- All strokes
  - Hemorrhagic stroke
  - Ischemic stroke
  - Undetermined

Statistical analyses included:
1) Counts and proportions of patients who developed the event of interest, summarized by treatment group,
2) A comparison of treatment groups, using a log-rank test based on time-to-first event, and
3) An estimate of the hazard ratio and its 95% CI for the raloxifene group compared with the placebo group, using a Cox proportional hazards regression model based on time-to-first event. Kaplan-Meier curves for all strokes were generated for each treatment group.

The statistical analyses for the stroke endpoint were repeated for the PP population

Subgroup Analyses of the Stroke Endpoints

Subgroup analyses of the stroke endpoint used the ITT population. A Cox proportional hazards regression model was fitted with treatment, subgroup indicator variable, and treatment-subgroup interaction in the model. The hazard ratio and 95% CI were reported for each subgroup, along with the interaction p-value. If the total number of events for a specific subgroup was less than 5, no statistical analysis was performed, and only the counts and proportions were reported. Subgroup analyses of all strokes were performed for each of the clinically relevant risk factors (Table GGIO.9.5). Unbalanced baseline characteristics were also considered as subgroup variables.

Table xxx List of Clinically Relevant Risk Factors for Stroke

Subgroup	Categories
Age (years)	≤65, >65 and ≤70, >70
Race	Caucasian, All other races
Region	North America, Latin/South America, Western Europe, Eastern Europe, Africa, Asia Pacific
Body mass index (kg/m^2)	≤25, >25 and ≤30, >30
Primary prevention population	Yes/No
Lower extremity arterial disease at baseline	Yes/No
Diabetes mellitus at baseline	Yes/No
Current smoker at baseline	Yes/No
Hypertension at baseline	Yes/No
Hyperlipidemia at baseline	Yes/No
Cardiovascular risk score at baseline	≤5, >5 and ≤9, >9
History of atrial fibrillation	Yes/No
HMG-CoA reductase inhibitor use at baseline	Yes/No
Warfarin use at baseline	Yes/No
Aspirin use at baseline	Yes/No
Non-aspirin antiplatelet use at baseline	Yes/No
Beta-blocker use at baseline	Yes/No
Calcium channel blocker use at baseline	Yes/No
ACE inhibitor or ARB use at baseline	Yes/No
Diuretic use at baseline	Yes/No

Abbreviations: ACE = angiotensin converting enzyme; ARB=angiotensin receptor blocker; HMG-CoA = hydroxymethylglutaryl-coenzyme A; MI = myocardial infarction.

Statistical Analyses of VTE Endpoint (9.7.1.12.3.2.)

The effect of raloxifene on the incidence of all VTEs, and the categories of VTEs listed below, were determined:

- All VTEs
  - Deep vein thrombosis (DVT)
  - Pulmonary embolism (PE)
  - PE or DVT
  - Intracranial thrombosis (IT)
  - Other VTEs

The statistical analyses included: 1) counts and proportions of patients who developed the event of interest, summarized by treatment group, 2) comparison of treatment groups using a log-rank test based on time-to-first event, and 3) an estimate of the hazard ratio and its 95% CI for the raloxifene group compared with the placebo group, using a Cox proportional hazards regression model based on time-to-first event. Kaplan-Meier curves were generated for all VTEs, DVT, PE, and PE or DVT for each treatment group. In addition, incidence rates of each VTE endpoint event were reported by treatment group.

These VTE analyses, excluding the Kaplan-Meier curves, were repeated for the PP population, the primary prevention population and secondary prevention population (Section 9.7.1.6).

Statistical Analyses of CV Endpoint Combinations (9.7.1.12.3.3.)

The effect of raloxifene on the following CV endpoint combinations was determined:

- CV death, nonfatal (including silent) MI, hospitalized ACS other than MI, or stroke.
- CV death, nonfatal (including silent) MI, hospitalized ACS other than MI, stroke, or myocardial revascularization.

The statistical analyses included: 1) counts and proportions of patients who developed the endpoint of interest, summarized by treatment group, 2) comparison of treatment groups, using a log-rank test based on time-to-first event, and 3) an estimate of the hazard ratio and its 95% CI for the raloxifene group compared with the placebo group, using a Cox proportional hazards regression model based on time-to-first event.

Statistical Analysis of All-Cause Mortality Endpoint (9.7.1.12.3.4.)

The effect of raloxifene on all deaths and on deaths due to the different causes listed below was determined:

- All Deaths
  - Cardiovascular Death
    - Coronary Death
      - Acute MI
      - Sudden death
      - Unwitnessed death
      - Heart failure with a history of coronary artery disease
      - Related to undergoing a coronary artery procedure
      - Specific cause of coronary death unavailable
    - Non-coronary Death
      - Cerebrovascular disease (stroke and other cause)
      - Aortic, mesenteric, renal, and lower limb peripheral vascular disease
      - Related to undergoing a non-coronary arterial procedure
      - Venous thromboembolic event
      - Endocarditis/myocarditis
      - Valvular disease
      - Other non-coronary death
      - Specific cause of non-coronary death unavailable
  - Non-cardiovascular Death
    - Cancer
      - Breast cancer
      - Other cancer
    - Accidental/Suicide/Homicide
    - Other non-cardiovascular death
    - Specific cause of non-cardiovascular death unavailable
  - Cause of death unavailable

The statistical analyses included: 1) counts and proportions of all deaths, summarized by treatment group, 2) counts and proportions of deaths based on cause of death, summarized by treatment group, 3) comparison of treatment groups, using a log-rank test based on time to death, and 4) an estimate of the hazard ratio and its 95% CI for the raloxifene group compared with the placebo group using a Cox proportional hazards regression model, based on time to death. Kaplan-Meier curves were generated for all deaths, cardiovascular death, and death due to any cancer for each treatment group.
For deaths categorized as "non-cardiovascular" and "other cancer," the specific cause of death identified by the adjudication process was summarized by treatment group.

Statistical Analyses of Revascularization and Non-traumatic Lower Extremity Amputation Endpoints (9.7.1.12.3.5.)
Statistical analyses were performed examining the effect of raloxifene on all revascularizations and non-traumatic lower extremity amputations.

- All revascularizations
  - Myocardial revascularization
    - Percutaneous coronary intervention (PCI)
    - Coronary artery bypass graft (CABG)
    - Other myocardial revascularization
  - Non-coronary arterial revascularization
    - Carotid district
    - Lower extremity
    - Other non-coronary arterial revascularization
- Non-traumatic lower extremity amputation
  - Above the knee
  - Below the knee
  - Foot/toe
  - Other non-traumatic lower extremity amputation

The statistical analyses included: 1) counts and proportions of patients who developed the event of interest, including subcategories, summarized by treatment group, 2) comparison of treatment groups, using a log-rank test based on time-to-first event, and 3) an estimate of the hazard ratio and its 95% CI for the raloxifene group compared with the placebo group, using a Cox proportional hazards regression model based on time-to first event. Kaplan-Meier curves were generated for all revascularizations and myocardial revascularization for each treatment group.

Statistical Analyses of Fracture Endpoint (9.7.1.12.3.6.)

Vertebral fractures:
According to the study design, there were no scheduled baseline and post-baseline radiographs to capture **prevalent vertebral fractures** at baseline and new **vertebral fractures** post-baseline. Given this limitation, it was not possible to state with certainty that a vertebral fracture was actually a new clinical vertebral fracture, since a morphometric fracture may have been present at baseline. For this reason, the interpretation of the analysis of vertebral fractures in this study report was with respect to "possible new clinical vertebral fractures." Only fractures confirmed

using radiographic evidence or other documented evidence, as available, were included in the analyses.

Non-vertebral fractures:
Non-vertebral fractures, defined by the combined fractures at sites of arm/forearm/elbow, clavicle/scapula/shoulder, wrist, ribs/sternum, pelvis/sacrum, hip/femur, and tibia/fibula/patella, were analyzed with and without inclusion of ankle fractures.

Statistical analyses were performed, examining the effect of raloxifene on the following individual or combined fracture sites:

- Non-vertebral including ankle
  - Non-vertebral excluding ankle
- Hip/femur or wrist
  - Hip/femur
  - Wrist
- Vertebral

The statistical analyses included: (1) the count and the proportion of patients who develop the endpoint of interest, calculated and summarized by treatment group, (2) the comparison of treatment groups, using a log-rank test based on time-to-first-event, and (3) an estimate of the hazard ratio and its 95% CI for the raloxifene group compared with the placebo group, using a Cox proportional hazards regression model based on time to first event. Kaplan-Meier curves were generated for each treatment group.
Note that data regarding the traumatic nature of these fractures were not captured on the case report form (CRF). Given this limitation, analyses may have included non-osteoporotic fractures which were likely not amenable to prevention by pharmacologic therapy.

Statistical Analysis of All-Cause Hospitalization Endpoint (9.7.1.12.3.7.)
The reasons for hospitalization were classified as listed below:

- MI
- VTE
- Revascularization or amputation
- Stroke
- Fracture
- Breast cancer
- Hospitalized ACS other than MI
- Other

The number of patients hospitalized for each reason was summarized for each treatment group. Chi-square tests were used to compare the reason for each cause of hospitalization between treatment groups.
In the original protocol, hospitalization due to unstable angina was a secondary endpoint. This secondary endpoint was changed to hospitalized ACS other than MI in Protocol Amendment (c) (Appendix 16.1.1). Investigators reporting hospitalization due to unstable angina were asked to re-evaluate if this hospitalization qualified as hospitalized
ACS other than MI. Prior to this protocol amendment, hospitalization due to unstable angina was collected on a separate adjudication form. Following the protocol amendment, hospitalization

due to unstable angina was no longer collected. The number of hospitalizations due to unstable angina collected prior to the protocol amendment was summarized by treatment group.

Change from the Protocol

The protocol specified that survival analyses, which included a log-rank test and a Cox proportional hazards regression model, were to be performed for all the secondary endpoints. For the all-cause hospitalization analysis, a Chi-square test replaced the survival analyses specified in the protocol. Since the criteria for hospitalization may have differed across countries, the cause of hospitalization, and not the time to hospitalization, was the focus of the analysis.

Statistical Analyses for Biochemical Markers of Cardiovascular Risk (9.7.1.12.4.)

- The biochemical markers of cardiovascular risk included fasting total cholesterol (CHOL), low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), triglycerides (TRIG), fibrinogen (collected in a subset of patients only), non-HDL-C (calculated as CHOL – HDL-C), and the ratio of total cholesterol to HDL-C (CHOL/HDL-C).
- Descriptive statistics (mean, standard deviation, standard error, median, minimum, and maximum) were summarized by treatment group for each marker at baseline, Year 1 (Visit 5), Year 5 (Visit 13), and at study conclusion (final visit).
- Mean changes and mean percent changes from baseline to endpoint were reported for each treatment group. The treatment group comparison of the mean changes from baseline was performed using a ranked one-way ANOVA. The treatment group comparison of mean percent changes was performed using a one-way ANOVA.
- Treatment group comparisons of mean change and mean percent change from baseline to Year 1 were repeated, as described above.

Change from the Protocol

Inclusion of non-HDL-C and the ratio of total cholesterol to HDL cholesterol as biochemical markers of cardiovascular risk were not specified in the protocol; however, these are recognized as a marker of cardiovascular risk, and thus were included in the analysis.

According to the protocol, biochemical markers of cardiovascular risk were analyzed using a two-way ANOVA, with treatment, country of investigative site, and their interaction in the model. The treatment group comparison of interest was the change from baseline for the biomarkers. Therefore, a ranked one-way ANOVA model, with only treatment in the model, replaced the protocol specified two-way ANOVA model.

The protocol stated that analysis of serum lipids and fibrinogen was based on the change from baseline to each visit. However, this analysis was conducted only for the change from baseline to Year 1 (Visit 5).

Safety Analyses (9.7.1.13.)

Study Exposure (9.7.1.13.1.)

Patients may have stopped taking study drug for various reasons (for example, an AE, breast cancer, venous thromboembolic event, physician decision, or patient decision), but, in accordance with the study design, remained in the study for follow-up. Therefore, to capture the study drug exposure of each patient accurately, days on therapy at each visit were calculated first. The study drug exposure in years for each patient was then defined as the sum of days at each visit on therapy divided by 365.25.

Study drug was dispensed at Visits 2, 4, and every 6-month visit thereafter. No tablets were dispensed at Visit 3, which was an optional visit. Unless a patient discontinued from the study at Visit 3, Visit 4 was the first visit at which study exposure could have been calculated. Let N represent any visit, if a patient had study drug dispensed at
Visit N-1 (for Visit 4, Visit N-1 was Visit 2) and the patient reported that she was currently taking study drug at Visit N, the days of study exposure at Visit N were calculated as the date of Visit N minus the date of Visit N-1. If a patient had study drug dispensed at Visit N-1 and reported at Visit N that she discontinued taking study drug for any reason sometime after Visit N-1, the days of study exposure at Visit N were calculated as the date of last dose minus date of Visit N-1. If the date of last dose was missing, the date of Visit N was used. If study drug was not dispensed to the patient at Visit N-1, days of study exposure at Visit N was 0.

Descriptive statistics (mean, standard deviation, median, minimum, and maximum) were summarized by treatment group for study exposure. The treatment group comparison of the mean study exposure was estimated by using a one-way ANOVA.

Counts and proportions of patients exposed to study drug for at least 1 year, 2 years, 3 years, 4 years, and 5 years or more were summarized by treatment group. Treatment groups were compared using a Chi-square test.

Adverse Event Analyses (9.7.1.13.2.)

An AE was defined as any untoward medical occurrence in a patient who was randomized in this clinical trial, without regard to the possibility of a causal relationship with study drug, and without regard to treatment group assignment, even if no study drug had been taken. Analyses of AEs were based on investigator-reported AEs, coded using MedDRA.
In this section, statistical analyses of AEs were performed for treatment-emergent adverse events (TEAEs) and adverse reactions, and serious adverse events (SAEs) and other notable AEs at the MedDRA system organ class (SOC), High-level Term, and Preferred Term levels, unless otherwise specified. Counts and proportions of patients experiencing the event of interest were reported for each treatment group. Treatment group differences were assessed using a CMH test stratified by country. If the total number of an event of interest was less than 5, no statistical test was performed.

Treatment-Emergent Adverse Events and Adverse Reactions (9.7.1.13.2.1.)

- A **TEAE** was defined as an event that first occurs or worsens (increases in severity) after baseline (Visit 2). An analysis of TEAEs occurring in at least 2% of raloxifene-treated patients was performed. Section 14.4.1 presents a detailed analysis of all TEAEs.

- **Adverse reactions** are those events which were deemed by the investigator to be reasonably, possibly related to either study drug administration or protocol procedures. Section 14.4.1 presents an analysis of adverse reactions:

Serious Adverse Events, Other Notable Adverse Events, and Adverse Events Leading to Discontinuation of Study Drug (9.7.1.13.2.2.)

Serious Adverse Events
In this study, the primary and many secondary endpoint events fulfilled the MedWatch definition of an SAE (ie, death, life-threatening, hospitalization, disability, or require intervention to prevent permanent impairment or damage). However, in the protocol, an SAE was defined as an event that meets one of the following criteria: life-threatening, severe or permanently disabling, cancer, or significant for any other reason. Only in the following instances was a primary or secondary endpoint event reported as an SAE:

- Any venous thromboembolic event
- Any death or hospitalization due to an endpoint, or a non-endpoint event deemed by an investigator to be reasonably, possibly related to either study drug administration or protocol procedures
- Any of the remaining endpoints (MI, hospitalized ACS other than MI, stroke, myocardial revascularization, non-coronary revascularization, non-traumatic lower extremity amputation, breast cancer, or fracture) which met one of the serious criteria, and were deemed by the investigator to be reasonably, possibly related to either study drug administration or protocol procedures, or
- Any hospitalization for a non-endpoint event that met one of the serious criteria, regardless of relationship to study drug

Other Notable Adverse Events
Notable AEs included events that are deemed "clinically significant" for raloxifene based on previous data or the literature. Special search categories (SSCs) were defined by grouping related MedDRA lower-level terms into clinically relevant event categories. Table GGIO.9.6 lists notable AEs identified by SSCs. Appendix 16.1.9 presents the details of the SSCs.
In addition to the analyses specified earlier, survival analyses based on time-to-first event were performed for the "all cancer" and "endometrial cancer" categories. A Cox proportional hazards regression model was fitted to estimate the hazard ratio and 95% CI comparing treatment groups, and Kaplan-Meier curves were generated for each treatment group.

Analysis of All Adverse Events Leading to Discontinuation of Study Drug
Analysis of all AEs leading to discontinuation of study drug was performed at the MedDRA SOC, High-level Term, and Preferred Term levels.

Table xxx Notable AEs Identified by Special Search Categories

Benign breast changes or diseases:

- **Fibrocystic breast disease**
- **Fibroadenoma**
- **Breast cysts**
- **Fibroses**
- **Sclerosing adenosis**
- **Dysplasia**
- **Hyperplasia**
  - **Atypical hyperplasia**
  - **Other hyperplasia**
- **Miscellaneous and breast neoplasm benign**
- **Breast conditions**
  - **Nipple discharge**
  - **Galactorrhea**
  - **Intraductal papilloma**
  - **Mastitis**
  - **Mammary duct ectasia**
  - **Breast pain or tenderness**
  - **Breast hypertrophy**
  - **Breast lump NOS**
  - **Miscellaneous breast conditions**

Benign gynecological conditions:

- Cervix neoplasm
- Ovarian neoplasm
- Vaginal neoplasm
- Vulvar neoplasm
- Postmenopausal bleeding
- Uterine neoplasm
  - Uterine polyps
  - Fibroid/leiomyoma/endometriosis
  - Uterine cysts
  - Benign uterine neoplasm
  - Uterine hyperplasia
  - Uterine hypoplasia
  - Other uterine neoplasm
- Other benign gynecological conditions

Cardiac arrhythmias:

- Supraventricular arrhythmias
  - Atrial fibrillation
  - Supraventricular arrhythmias other than atrial fibrillation
- Ventricular arrhythmias
- Cardiac conduction disorder
- Other arrhythmias

Hot flashes

Leg cramps

Influenza-like syndrome

Peripheral edema

Cataracts

Gallbladder disease:

- **Cholecystitis and cholelithiasis**
- **Gallbladder disorder/miscellaneous**

All cancer:

- **Breast cancer**
- **Endocrine cancer**
  - **Thyroid cancer**
  - **Other endocrine cancer**
- **Gastrointestinal cancer**
  - **Anal cancer**
  - **Colon cancer**
  - **Colorectal cancer**
  - **Gastric cancer**
  - **Esophageal cancer**
  - **Pancreas cancer**
  - **Rectal cancer**
  - **Small intestine cancer**
  - **Lip and oral cavity cancer**
  - **Salivary gland cancer**
  - **Other gastrointestinal cancer**
- Hematopoietic cancer
- Hepatic and biliary cancers
  - Bile duct cancer
  - Gallbladder cancer
  - Hepatic cancer
  - Other hepatic and biliary cancers
- Leukemias
  - Acute myeloid leukemia
  - Acute lymphocytic leukemia
  - Chronic myeloid leakage
  - Chronic lymphocytic leukemia
  - Myelodysplastic syndrome
  - Other leukemias
- Lymphomas
  - Hodgkin's disease
  - Non-Hodgkin's B-cell
  - Non-Hodgkin's T-cell
  - Non-Hodgkin's lymphomas
  - Other lymphomas
- Nervous system (malignant)
- Ocular cancer
- Plasma cell neoplasm malignant
  - Plasma cell cancer
  - Multiple myeloma

All cancer (continued)

- Renal and urinary tract cancer
  - Bladder cancer
  - Non-renal cell kidney cancer
  - Renal cell kidney cancer
  - Renal pelvis and ureter cancer
  - Urinary tract cancer
- Reproductive cancer
  - Cervix cancer
  - Endometrial cancer
  - Fallopian tube cancer
  - Vaginal cancer
  - Vulva cancer
  - Ovarian cancer
  - Ovarian choriocarcinoma
  - Uterine cancer
  - Uterine sarcoma
  - Other reproductive cancer
- Respiratory and mediastinal cancer
  - Mesothelioma
  - Small cell lung cancer
  - Non-small cell lung cancer
  - Other respiratory cancer
- Skeletal cancer
- Skin cancer
  - Melanoma
  - Basal cell skin cancer
  - Squamous cell skin cancer
  - Other skin cancer
- Soft tissue cancer
  - Sarcoma (other than bone and uterine)
  - Other soft tissue cancer
- Miscellaneous/site unknown cancer

Analyses of Clinical Laboratory Measurements (9.7.1.13.3.)

Clinical laboratory measurements collected in this trial included the following:

- Total bilirubin
- AST
- BUN
- Serum creatinine
- Fasting glucose
- HgbA1c

Analyses of fasting glucose and HgbA1c were performed separately on patients with and without diabetes mellitus at baseline.
For each safety analyte, summary statistics (mean, standard deviation, standard error, median, minimum, and maximum) were reported for each treatment group by year. The change from baseline to endpoint was compared between treatment groups using a ranked one-way ANOVA. Shift tables presented patient counts and percentages, with rows representing baseline lab value category, and columns representing maximum post-baseline value category.
Changes from baseline category to post-baseline category were classified as down (a decrease), up (an increase), and same (no change), and were reported by treatment group. Treatment groups were compared using a likelihood-ratio test.

Change from the Protocol
According to the protocol, safety analytes were analyzed using a two-way ANOVA, with treatment, country of investigative site, and their interaction in the model. In this analysis, the overall comparison between treatment groups for changes from baseline was of interest. Therefore, a ranked one-way ANOVA model, with only treatment in the model, replaced the protocol-specified two-way ANOVA model. In addition, shift tables were generated.

Vital Signs and Physical Findings Analyses (9.7.1.13.4.)
Summary statistics of vital signs (heart rate, systolic and diastolic blood pressure) and physical findings (height, weight, and body mass index [BMI]) were reported by treatment group. The change from baseline to endpoint for each vital sign and physical finding was compared between treatment groups using a ranked one-way ANOVA. A within-treatment group comparison of the change from baseline to endpoint for each vital sign and physical finding was performed using the Wilcoxon Signed Rank test.

Change from the Protocol
According to the protocol, vital signs were to be analyzed using a two-way ANOVA, with treatment, country of investigative site, and their interaction in the model. In this analysis, the overall comparison between treatment groups for changes from baseline was of interest. Therefore, a ranked one-way ANOVA model, with only treatment in the model, replaced the protocol-specified two-way ANOVA model. Analysis of vital signs by visit was to be performed.

Interim Analyses and Data Monitoring (9.7.1.14.)

The Data Safety Monitoring Board (DSMB) consisted of 8 individuals who were external to the sponsor, and who were not investigators in any raloxifene studies. Membership consisted of individuals with experience in clinical trials, CVD, oncology, and statistics.
The Chair of the committee oversaw the process of selecting the DSMB members.

- The DSMB functioned under operating guidelines describing their activities. Only the DSMB was authorized to review unblinded interim analyses. These analyses were prepared by an external DAG, allowing Eli Lilly and Company (Lilly) to remain blinded to the data. The DSMB reported their recommendations to the RUTH EC Chair and the Eli Lilly Senior Management Designee.

- 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 27 January 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 ≤ 0.000001. The final breast cancer analysis was conducted at the ≤ 0.008 level.

Three **interim efficacy analyses of the coronary primary endpoint** were planned 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
28 January 2002, at which time 328 patients (26% of 1268) had experienced a coronary primary endpoint event. The second interim analysis occurred on 08 September 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 ≤ 0.0001.
The final analysis was conducted at the ≤ 0.0423 level. This ensured an overall type I error rate of 0.04234 for the coronary primary endpoint.

Change in the study duration methodology

In 2004, the sponsor identified a discrepancy between the protocol and the informed consent document (ICD). The ICD stated that patients will participate in the study for between 5 and 7.5 years, while the protocol stated that patients will 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 sponsor 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 sponsor, EC, and DSMB agreed that it was not in the best interest of patients to extend their participation beyond 7.5 years. The sponsor and EC requested that the DSMB assess two other options: continuing the trial until each woman randomized and surviving has been followed for a minimum of 5 years or continuing the trial until each woman has 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 (ie, 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.

Determination of Sample Size (9.7.2.)

A sample size requirement of approximately 10,000 patients was calculated using the method of Lakatos (Lakatos 1986, 1988; Shih 1995), which provides sample size estimates for tests based on the 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 non-cardiovascular 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%.

Changes in the Conduct of the Study or Planned Analyses (9.8.)

Changes in the Conduct of the Study (9.8.1.)

The Protocol GGIO was initially approved on 18 December 1997, with subsequent amendments approved (a) 14 January 1998, (b) 12 May 1998, (c) 13 April 2000, and (d) 03 May 2002.

Amendment (a): 14 January 1998

Protocol GGIO Amendment (a) clarified: exclusion criteria (use of vaginal creams; prior participation in raloxifene studies), blinding (discontinuation due to unblinding by the investigator), and SAE reporting (cancer other than breast cancer). The Schedule of Events was updated to allow the completion of patient medical histories at Visit 1 and Visit 2.

Amendment (b): 12 May 1998

Protocol GGIO Amendment (b) was implemented to update or revise preclinical information, information on concomitant medications, and the references. Additionally, revisions were made to clarify language, definitions, or criteria in the study design; inclusion/exclusion and protocol violation criteria; compliance and discontinuation; unblinding, and efficacy and safety measures; and analyses including secondary endpoints and interim analyses. The definition of the Data Safety Monitoring Board
(DSMB) was changed to expand its membership during the study and to add 3-month reviews. The Schedule of Events was updated to clarify the definitions, procedures, and timings of mammograms, physical exams, laboratory values, ECGs, and the final visit.

Amendment (c): 13 April 2000

Protocol GGIO Amendment (c) added invasive breast cancer as a second, separate primary endpoint, based on data from Study H3S-MC-GGGK (GGGK) showing a significant 76% reduction in the incidence of invasive breast cancer at 40 months in women assigned raloxifene (Cummings et al. 1999). 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.

Amendment (d): 03 May 2002

Protocol GGIO Amendment (d) changed 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. In addition, new references were added and reference to the contract research organization, where applicable, was deleted because
Lilly had assumed related responsibilities.

Changes in the Planned Analyses (9.8.2.)

The data was locked for final analysis on 02 February 2006. Post-hoc analyses were performed on baseline characteristics, fractures, strokes, deaths due to stroke, and other safety findings. These analyses are described in the following subsections and the results are reported in the relevant sections of this report.

Changes in Patient Demographic and Other Baseline Characteristics (9.8.2.1.)

The following baseline patient characteristics were summarized by treatment group for the primary and secondary prevention populations:

- Demographic characteristics
- Cardiovascular risk assessment characteristics
- VTE risk assessment characteristics

For categorical variables, counts and proportions were reported by treatment group. Treatment group differences were assessed using a Chi-square test. For continuous variables, descriptive statistics, including mean, standard deviation, median, minimum, and maximum were reported.

Mean differences between treatment groups were assessed using an F-test from a one-way ANOVA.

Changes in Concomitant Medications (9.8.2.2.)

Baseline concomitant medication use was reported for each treatment group for the primary and secondary prevention populations. Treatment group differences were assessed using a Chi-square test.

Changes in the percentage of patients using prespecified concomitant medications from baseline to post-baseline were reported for each treatment group for the primary and secondary prevention populations. A CMH test was used to assess treatment group differences in the proportions of patients using concomitant medication from baseline to post-baseline.

Changes in the Analysis of the Secondary Endpoint of Stroke (9.8.2.3.)

Changes in the Time to Event Analysis of Stroke (9.8.2.3.1.)

The primary analysis of the stroke endpoint was for all strokes and the following stroke subtypes: hemorrhagic, ischemic, and type undetermined. The **pathogenesis** of each hemorrhagic or ischemic stroke was **further categorized** by the stroke adjudicating committee based on available information. Therefore, the primary analysis was updated to include this additional information.

Incidence Rate of Stroke (9.8.2.3.2.)

The incidence rates of all strokes, each stroke subtype, and each hemorrhagic or ischemic stroke categorized in terms of pathogenesis were reported. The incidence rate was calculated by dividing the number of patients who had an adjudicated stroke during the study period by the patient-years of follow-up. Absolute risk reduction was calculated based on the difference in cumulative incidence of the stroke endpoint of interest between the two treatment groups at the end of the study period.

9.8.2.3.3. Time to Event Analysis of the Stroke Endpoint by Year

Analysis of all strokes from randomization to the end of first, second, third, fourth, fifth, sixth, and seventh year of study follow-up was performed. For each time period, a patient was considered as having an event if she had a stroke, adjudicated as such, during the specified time period and was censored if she did not have a stroke. A Cox proportional hazards regression model was fitted to estimate the hazard ratio and 95% CI.

Additional Subgroup Analyses of the Stroke Endpoint (9.8.2.3.4.)

Additional subgroup analyses for stroke were performed for baseline characteristics (Table xxx).

Table xxx. Baseline Characteristics for Post-Hoc Subgroup Analyses of

All Strokes

Subgroup	Categories
Systolic BP at baseline (mmHg)	≤160, >160
Systolic BP at baseline (mmHg)	≤140, 140-160, ≥160
Diastolic BP at baseline (mmHg)	≤80, 80-90, ≥90
Mean Pulse Pressure at baseline (mmHg)[a]	≤60, >60
LDL-C at baseline (mg/dL)	≤100, 100-130, ≥130
Total cholesterol at baseline (mg/dL)	≤200, 200-240, ≥240
Triglycerides at baseline (mg/dL)	<150, ≥150
Congestive heart failure at baseline[b]	Yes/No
Ventricular hypertrophy on baseline ECG[c]	Yes/No
History of stroke or TIA[d]	Yes/No

Abbreviations: BP = blood pressure; ECG = Electrocardiogram; LDL-C = low-density lipoprotein cholesterol; NEC = not elsewhere classified; TIA = transient ischemic attack;.

a Pulse pressure was calculated as the difference between Systolic Blood Pressure and Diastolic Blood pressure.

b Defined as patients who reported a historical diagnosis or secondary condition that mapped to one of the following High-level Terms in MedDRA: (1) Heart failures NEC, (2) Heart failure signs and symptoms, (3) Right ventricular failures, or (4) Left ventricular failures.

c Defined as patients who were identified to have ventricular hypertrophy (with or without strain) on their baseline ECG, as interpreted by the central ECG laboratory.

d Defined as patients who reported a historical diagnosis or secondary condition that mapped to one of the following High-level Terms in MedDRA: (1) Central nervous system hemorrhages and cerebrovascular accidents, (2) Transient cerebrovascular events, or (3) Central nervous system vascular disorders NEC.

Time to Event Analysis of the Stroke Endpoint Adjusted for Baseline Risk Characteristics (9.8.2.3.5.)

A Cox proportional hazards regression model was used to adjust for clinically relevant risk factors. The analysis used the ITT population. Table GGIO.9.5 presents the list of clinically relevant risk factors. A final, multivariate adjusted model was determined by the following model-fitting procedure:
a) A univariate Cox model was fitted for each of the clinically relevant risk factors and unbalanced baseline characteristics.
b) A multivariate model, with covariates, which are significant in the univariate model, was fitted. A stepwise model selection method was used to determine the final model. In the stepwise model selection process, the significance level for entry in the model was 0.10, and for remaining in the model was 0.05.
c) Treatment was added to the final model, as determined by procedure (b), above.

The hazard ratio of stroke, comparing raloxifene to placebo and its 95% CI from this multivariate adjusted model, was reported.

Association between Stroke and Current Smoking for Patients Assigned to Placebo (9.8.2.3.6.)

A Chi-square test for association between stroke and current smoking status (defined at baseline as having smoked an average of ≥10 cigarettes per day during the last 6 months) was performed for patients assigned to placebo.

Changes in the Analysis of the Secondary Endpoint of VTE (9.8.2.4.)

Time to Event Analysis of VTE by Year (9.8.2.4.1.)

Analysis of VTE for time periods from randomization to the end of first, second, third, fourth, fifth, sixth, and seventh year of study follow-up was performed. For each time period, a patient was considered as having an event if she had a VTE adjudicated as such, during the specific time period and was censored if she did not have a VTE. A Cox proportional hazards regression model was fitted to estimate the hazard ratio and 95% CI.

Changes in the Analysis of the Secondary Endpoint of All Cause Mortality (9.8.2.5.)

Incidence Rate of All Cause Mortality (9.8.2.5.1.)

The incidence rate of all cause mortality and each cause of death was calculated by dividing the number of patients who died during the study period by the patient-years of follow up for each cause of death. Absolute risk reduction was calculated based on the difference in cumulative incidence of death between the treatment groups at the end of the study period.

Kaplan-Meier Curve for Deaths due to Stroke (9.8.2.5.2.)

Kaplan-Meier curves were generated for all deaths due to stroke for each treatment group.

Subgroup Analysis of Deaths due to Stroke (9.8.2.5.3.)

Additional subgroup analyses for death due to stroke were performed for baseline characteristics (Table xxx).

Table xxx. Baseline Characteristics for Subgroup Analysis of Deaths Due to Stroke

Subgroup	Categories
Age (years)	≤65, >65 and ≤70, >70
Race	Caucasian, All other races
Region	North America, Latin/South America, Western Europe, Eastern Europe, Africa, Asia Pacific
Body mass index (kg/m^2)	≤25, >25 and ≤30, >30
Primary prevention population	Yes/No
Lower extremity arterial disease at baseline	Yes/No
Diabetes mellitus at baseline	Yes/No
Current smoker at baseline	Yes/No
Hypertension at baseline	Yes/No
Hyperlipidemia at baseline	Yes/No
Cardiovascular risk score at baseline	≤5, >5 and ≤9, >9
History of atrial fibrillation	Yes/No
HMG-CoA reductase inhibitor use at baseline	Yes/No
Warfarin use at baseline	Yes/No
Aspirin use at baseline	Yes/No
Non-aspirin antiplatelet use at baseline	Yes/No
Beta-blocker use at baseline	Yes/No
Calcium channel blocker use at baseline	Yes/No
ACE inhibitor or ARB use at baseline	Yes/No
Diuretic use at baseline	Yes/No
Systolic BP at baseline (mmHg)	≤160, >160
Mean Pulse pressure at baseline (mmHg)[a]	≤60, >60
Congestive heart failure at baseline[b]	Yes/No
Ventricular hypertrophy on baseline ECG[c]	Yes/No
History of stroke or TIA[d]	Yes/No

Abbreviations: ACE = angiotensin converting enzyme; ARB = angiotensin receptor blocker; ECG = Electrocardiogram; HMG-CoA = hydroxymethylglutaryl-coenzyme A; MI = myocardial infarction; BP = blood pressure; LDL = low-density lipoprotein; TIA = transient ischemic attack.

[a] Pulse pressure was calculated as the difference between Systolic Blood Pressure and Diastolic Blood pressure.

[b] Defined as patients who reported a historical diagnosis or secondary condition that mapped to one of the following High-level Terms in MedDRA: (1) Heart failures not elsewhere classified (NEC), (2) Heart failure signs and symptoms, (3) Right ventricular failures, or (4) Left ventricular failures.

[c] Defined as patients who were identified to have ventricular hypertrophy (with or without strain) on their baseline ECG, as interpreted by the central ECG laboratory.

[d] Defined as patients who reported a historical diagnosis or secondary condition that mapped to one of the following High-level Terms in MedDRA: (1) Central nervous system hemorrhages and cerebrovascular accidents, (2) Transient cerebrovascular events, or (3) Central nervous system vascular disorders NEC.

Time to Event Analysis of Deaths due to Stroke by Year (9.8.2.5.4.)

Analysis of deaths due to stroke for time periods from randomization to the end of first, second, third, fourth, fifth, sixth, and seventh year of study follow-up was performed.
For each time period, a patient was considered as having an event if she died during the specific time period with an adjudicated cause of death being due to a cerebrovascular etiology and a patient was censored if she was alive. A Cox proportional hazards regression model was fitted to estimate the hazard ratio and 95% CI.

Summary of Types of Strokes in Patients who died due to a Stroke (9.8.2.5.5.)

On the death adjudication form, a death may have been classified: due to a cerebrovascular etiology (ie, stroke or other). Following data lock, all deaths due to cerebrovascular causes were reviewed and every death in this category was attributable to a stroke.

The committee adjudicating reported strokes was independent of the committee adjudicating deaths and assigning causality. An investigator may have reported *a death due to a stroke* and *a stroke* as trial endpoints. Available clinical documentation of the stroke was submitted to the stroke committee for review and determination if the prespecified criteria for stroke were met. If a reported stroke was adjudicated as such, the committee was asked to classify the type of stroke.

A death was adjudicated by a different committee and cause of death was assigned based on available clinical information, death certificate, or autopsy report. No criteria were prespecified in the protocol defining a death due to a cerebrovascular cause.

Consequently, a reported stroke may not have been adjudicated as such, but the cause of death may have been attributed to a cerebrovascular cause. Only the stroke committee prospectively classified strokes as ischemic or hemorrhagic in origin; the committee adjudicating deaths did not classify a death due to a cerebrovascular cause as ischemic or hemorrhagic.

It was of clinical interest to discern the type of stroke resulting in death. Because of the adjudication processes described above, the type of stroke resulting in death was determined based on the adjudication status of the last investigator reported stroke.

A Cox proportional hazards regression model was fitted to estimate the hazard ratio and 95% CI comparing treatment groups for stroke death by each stroke type as determined above.

Time to Event Analysis of Death due to Stroke Adjusted for Baseline Risk Characteristics (9.8.2.5.6.)

A Cox proportional hazards regression model was used to adjust for clinically relevant risk factors. The analysis used the ITT population. Table GGIO.9.5 presents the list of clinically relevant risk factors. A final, multivariate adjusted model was determined by the following model-fitting procedure:
a) A univariate Cox model was fitted for each of the clinically relevant risk factors and unbalanced baseline characteristics.

b) A multivariate model, with covariates, which are significant in the univariate model, was fitted. A stepwise model selection method was used to determine the final model. In the stepwise model selection process, the significance level for entry in the model was 0.10, and for remaining in the model was 0.05.
c) Treatment was added to the final model, as determined by procedure (b) above.
The hazard ratio of stroke death, comparing raloxifene to placebo and its 95% CI from this multivariate adjusted model, was reported.

Changes in the Analysis of Fractures (9.8.2.6.)

The **incidence rates of fractures** were calculated by dividing the number of patients who had fractures during the study period by patient-years of follow-up.
Absolute risk reduction was calculated based on the difference in the cumulative incidence of a given fracture event between treatment groups.

Changes in the Special Search Categories or Analyses of Notable Adverse Events (9.8.2.7.)

Changes in the Special Search Categories for Benign Breast Changes or Diseases (9.8.2.7.1.)

Benign breast changes or diseases were identified as notable adverse events. Appendix 16.1.9 details the lower level terms included in the special search categories for each of the adverse events categorized under this header (ie, Benign Breast Changes or Diseases). After data lock, it was recognized that several lower level terms specific to breast pain or tenderness were inadvertently mapped to the adverse event "miscellaneous and breast neoplasm benign" and were excluded from the SSC for the adverse event "breast pain or tenderness." Therefore, post hoc, the following LLTCODE/LLTCLASTs were removed from the SSC for "miscellaneous and breast neoplasm benign" and added to the SSC for "breast pain or tenderness": 10000426/ache breast, 10006298/breast pain, 10006299/breast pain female, 10026876/mastalgia, 10026892/mastodynia, 10033504/painful breasts, 10041354/sore breasts, and 10041370/soreness breast.

Analyses of Gallbladder Disease (9.8.2.7.2.)

Kaplan Meier curves were generated for gallbladder disease for each treatment group.
Patients who either had intact gallbladder at randomization, ie, patients who either did not report a historical diagnosis of cholecystectomy, or were reported to have gallbladder disease post-baseline were included in this analysis. Treatment group comparisons of the incidence of gallbladder disease were performed using a CMH test, stratified by country.
The following event categories and combinations were analyzed: (1) gallbladder disease; (2) cholecystectomy (all AEs that mapped to the MedDRA Version 8.0 Preferred Term of "cholecystectomy"); (3) gallbladder disease and cholecystectomy (all AEs that mapped to the MedDRA Version 8.0 Preferred Term of "cholecystectomy"); (4) gallbladder procedures (all AEs that mapped to the MedDRA Version 8.0 High-level Term of "biliary tract and gallbladder

therapeutic procedures"); and (5) gallbladder disease and gallbladder procedures (all AEs that mapped to the MedDRA Version 8.0 High-level
Term of "biliary tract and gallbladder therapeutic procedures").
Kaplan-Meier curves were generated for gallbladder disease in patients with an intact gallbladder at randomization for each treatment group.

Sensitivity Analyses of All Cancers and Endometrial Cancer (9.8.2.7.3.)

All cancers were identified using a pre-specified SSC. After approval of the SAP but before data unblinding, the LLT terms, "metastatic colorectal cancer" and "colorectal cancer stage IV," were removed from the "colon cancer" SSC and added to the "colorectal cancer" SSC.
One sarcoma was listed in the soft tissue cancer category for a patient assigned to raloxifene. Because a sarcoma is a rare cancer, all available source documents for this event were reviewed after data lock. It was subsequently identified that this was actually a uterine sarcoma. Originally the investigator reported the event as "sarcoma uteri" but later deleted this term and reported the event as a "low malignant leiosarcoma." Using the later AE terminology, the coding of this event in MedDRA led to it being mapped to a soft tissue sarcoma. However, retrospective review of the biopsy report confirmed the diagnosis as "leiomyosarcoma uterine." Therefore, this event actually was a uterine sarcoma. Sensitivity analyses were performed for all reproductive cancers and endometrial and uterine cancers combined, including this event as a "uterine sarcoma."

Incidence Rate and Kaplan-Meier Curves of Endometrial or Uterine Cancer Combined (9.8.2.7.4.)

Pathology reports from all reported uterine cancers were reviewed; it was determined that all uterine cancers were, in fact, endometrial cancers. Therefore, endometrial and uterine cancer adverse events were combined for post-hoc analyses. The incidence rate of endometrial cancer or uterine cancers combined was calculated by dividing the number of patients who reported an endometrial or uterine cancer by the total patient-years of study follow-up. Absolute risk reduction was calculated based on the difference in cumulative incidence of endometrial cancer or uterine cancer events combined between treatment groups. Only patients with an intact uterus at the time of randomization were included in this analysis. Kaplan-Meier curves for endometrial cancer or uterine cancers combined were generated for each treatment group.

Analyses of Ovarian Cancer (9.8.2.7.5.)

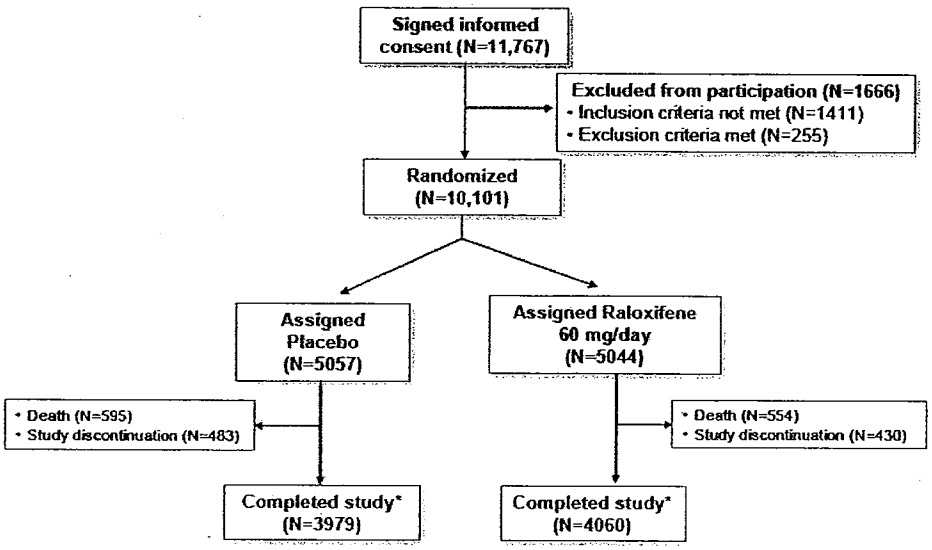
Patients with at least one ovary at randomization were included in the following analyses:

- Incidence rate of ovarian cancer was calculated by dividing the number of patients who reported an ovarian cancer by the total patient-years of study follow-up. Absolute risk reduction was calculated based on difference in cumulative incidence of ovarian cancer events between treatment groups.
- Kaplan Meier curves were generated for ovarian cancer for each treatment group.
- For patients who were diagnosed with ovarian cancer, mean age by treatment group was calculated.
- For patients who were diagnosed with ovarian cancer, mean time to the diagnosis of ovarian cancer by treatment group was calculated.

Patients (10)

Disposition of Patients (10.1.)

GGIO was a Phase 3, multicenter, double-blind, placebo-controlled, randomized, parallel study with two treatment groups: raloxifene HCl 60 mg/day and placebo. 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 xxx).



*Final visit on or after March 1, 2005

Sources: SFTSDSC, MSTINV1, TAB1PAT, TAB2PAT

Figure xxx. Patient disposition.

Reasons for study discontinuation

Table xxx below shows the reasons for study discontinuation. A total of 2062 (20%) patients discontinued from the study before it was concluded. Statistically significantly more raloxifene-assigned patients completed the study compared with placebo-assigned patients, though the absolute difference in the proportion of completers was small.
Statistically significantly 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).

Table xxx. Reasons for Study Discontinuation (All Randomized Patients)

Study discontinuation reason	Placebo (N=5057) n (%)	Ralox (N=5044) n (%)	Total (N=10101) n (%)	p-Value*
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.

Program: RMP.H3SSGGIO.SASPGM(SFCMDSP1)
Data: RMP.SAS.H3SM.L.MCGGIOSA.FINAL.MAIN Output: RMP.H3SO.GGIO.FINAL(SFTSDSC)

Table xxx. Total Patient-Years of Study Follow Up (All Randomized Patients)

	Placebo (N=5057)	Ralox (N=5044)	Total (N=10101)
Total patient-years of study follow up*			
Mean	5.24	5.31	5.27
Standard deviation	1.45	1.35	1.40
Median	5.55	5.57	5.56
Minimum	0.05	0.01	0.01
Maximum	7.06	7.04	7.06
Total patient-years of study follow up* for primary prevention population			
Mean	5.19	5.29	5.24
Standard deviation	1.49	1.39	1.44
Median	5.54	5.57	5.56
Minimum	0.05	0.01	0.01
Maximum	7.06	7.04	7.06
Total patient-years of study follow up* for secondary prevention population			
Mean	5.28	5.33	5.31
Standard deviation	1.41	1.32	1.36
Median	5.56	5.57	5.56
Minimum	0.09	0.01	0.01
Maximum	7.02	7.02	7.02

*{Censoring date (for all analyses except mortality) - randomization date + 1} / 365.25.

Program: RMP.H3SSGGIO.SASPGM(MSCTFLUP)
Data: RMP.SAS.H3SM.L.MCGGIOSA.FINAL.MAIN Output: RMP.H3SO.GGIO.FINAL(MSTFLUP)

Reasons for Final Study Drug Discontinuation

Table xxx shows the reasons for the **final** study drug discontinuation. A total of 5822 (57.6%) patients discontinued study drug due to study completion. The primary reason for study drug discontinuation was either an AE or patient decision in the initial 4.5 years of follow-up in the trial. The largest proportion of women discontinued study drug in the first 6 months due to an AE or patient decision. The frequency of study drug discontinuation declined thereafter until the fifth year of the trial. (Table GGIO.14.2)

- Statistically significantly more raloxifene-treated patients compared with placebo-treated patients discontinued study drug permanently due to an AE.

Table xxx. Reasons for Final Study Drug Discontinuation (All Randomized Patients)

Final study drug discontinuation reason	Placebo (N=5057) n (%)	Ralox (N=5044) n (%)	Total (N=10101) n (%)	p-Value*
Study is concluded	2915(57.64)	2907(57.63)	5822(57.64)	.992
Adverse event	1008(19.93)	1107(21.95)	2115(20.94)	.013
Patient decision	671(13.27)	619(12.27)	1290(12.77)	.133
Death	217(4.29)	197(3.91)	414(4.10)	.329
Lost to follow-up	141(2.79)	122(2.42)	263(2.60)	.244
Patient moved	68(1.34)	53(1.05)	121(1.20)	.175
Protocol violation	21(0.42)	17(0.34)	38(0.38)	.521
Protocol entry criteria not met	15(0.30)	18(0.36)	33(0.33)	.596
Unknown	1(0.02)	4(0.08)	5(0.05)	.218

*p-Value is obtained from a Pearson's Chi-square test if total>=10, Fisher's exact test if 5<=total<10, and N/A otherwise.

Program: RMP.H3SSGGIO.SASPGM(SFCMDSP3)
Data: RMP.SAS.H3SM.L.MCGGIOSA.FINAL.MAIN Output: RMP.H3SO.GGIO.FINAL(SFTDDSC)

Protocol Violations (10.2)

Significant protocol violations are those departures from the protocol that, per the International Conference on Harmonization (ICH) E3 Guideline, are "related to inclusion or exclusion criteria, conduct of the trial, and patient management or assessment that could have a direct effect on the outcome of the study."
The numbers and proportions of patients meeting each type of significant protocol violation were similar between treatment groups (Table GGIO.10.4). Appendix 16.2.2 presents a detailed by-patient listing of each important protocol violation.
A few patients did not have a baseline mammogram (0.04%) or ECG (0.12%) and a few patients did not have a post-baseline mammogram (5.07%) or ECG (4.58%). These were numerically comparable between treatment groups at baseline and post-baseline. It is unlikely these missed procedures had any direct impact on the primary endpoint analyses or conclusions presented in this report.
Similarly few patients did not have scheduled labs at baseline (0.31%) or post-baseline (2.97%). This was unlikely to directly impact the laboratory analyses, especially since labs were evaluated annually.

As per the protocol, patients were required to discontinue study drug if they were prescribed oral or transdermal hormone therapy, any SERM, or tibolone. Determination of whether hormone therapy was taken orally or vaginally could not be ascertained because investigators were not required to report the mode of administration for concomitant medications, and brand and generic names for hormone therapy do not reflect the mode of administration. Determination of concomitant use of any SERM or tibolone could not be made because it is unknown if the SERM or tibolone use occurred after study drug was permanently discontinued or was prescribed only for a short period of time during temporary study drug discontinuation, after which study drug may have been resumed per the protocol. Therefore, an accurate assessment of this protocol violation could be not conducted.

Table GGIO.10.4. Patients with Significant Protocol Violations

Important protocol violations	Placebo N=5057 n (%)	Raloxifene N=5044 n (%)	Total N=10,101 n (%)
Inclusion criteria not met or exclusion criteria met[a]	298 (5.89)	314 (6.23)	612 (6.06)
Study drug continued following a diagnosis of VTE or a breast cancer[b]	1 (0.02)	1 (0.02)	2 (0.02)
Scheduled mammograms not performed	984 (19.46)	983 (19.49)	1967 (19.47)
At baseline	2 (0.04)	2 (0.04)	4 (0.04)
At least one post-baseline	982 (19.42)	983 (19.49)	1965 (19.45)
At any post-baseline	269 (5.32)	243 (4.82)	512 (5.07)
Scheduled ECGs not performed	812 (16.06)	831 (16.48)	1643 (16.27)
At baseline	4 (0.08)	8 (0.16)	12 (0.12)
At least one post-baseline	810 (16.02)	824 (16.34)	1634 (16.18)
At any post-baseline	242 (4.79)	221 (4.38)	463 (4.58)
Scheduled labs not performed	1217 (24.07)	1242 (24.62)	2459 (24.34)
At baseline	10 (0.20)	21 (0.42)	31 (0.31)
At least one post-baseline	1213 (23.99)	1230 (24.39)	2443 (24.19)
At any post-baseline	149 (2.95)	151 (2.99)	300 (2.97)
Concomitant enrollment in other clinical trials	11 (0.22)	7 (0.14)	18 (0.18)
Received incorrect study drug kits[c]	12 (0.24)	14 (0.28)	26 (0.26)

Source: MSTPVIOL and CRF notes to file.

Abbreviations: ECG = electrocardiogram; N = number of patients assessed; n = number of patients with a protocol violation; SERM = selective estrogen receptor modulator; VTE = venous thromboembolic event.

[a] Includes violations of any criterion except for the exclusion criterion prohibiting concomitant use of study drug with oral estrogens, SERMs, and tibilone. See text above table for explanation.

[b] Within a 6-month window.

[c] Patients who received incorrect study drug kits may have received correct study drug.

Efficacy Evaluation (11)

Datasets Analyzed (11.1.)

Analyses of primary and secondary efficacy objectives were performed based on the ITT principle. The **ITT population** is referred to as "all randomized patients" throughout this report.

Sensitivity analyses were performed on subsets of patients as follows:

- **Per-protocol (PP) population:** a subset of all randomized patients, consisting of patients who met all inclusion criteria, did not fulfill any exclusion criteria, and were compliant with study drug treatment
- **Primary prevention population:** randomized patients at increased risk for CHD
- **Secondary prevention population:** randomized patients with documented CHD
- Randomized patients who were **60 years or older at baseline** (for invasive breast cancer endpoint)

On 02 February 2006 the final reporting database was validated and locked.

Demographic and Other Baseline Characteristics (11.2.)

The following baseline patient characteristics were summarized by treatment group:

- Demographic characteristics (Section 11.2.1)
- Breast cancer risk assessment characteristics (Section 11.2.2)
- CV risk assessment characteristics (Section 11.2.2)
- VTE risk assessment characteristics (Section 11.2.2)
- Biochemical markers of CV risk (Section 11.2.2)
- Concomitant medications (Section 11.2.3)

Post-hoc analyses were performed for the primary and secondary prevention populations for patient demographic characteristics, CV risk assessment characteristics, and VTE risk assessment characteristics (Section 9.8.2.1), and concomitant medications (Section 9.8.2.2). Appendix 16.2.4 presents a by-patient listing of demographic characteristics.

Patient Demographic Characteristics at Baseline (11.2.1.)

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

Table xxx. 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.

Program: RMP.H3SSGGIO.SASPGM(MSCHBL1)
Data: RMP.SAS.H3SM.L.MCGGIOSA.FINAL.MAIN Output: RMP.H3SO.GGIO.FINAL(MSTBLDEM)

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.

Program: RMP.H3SGGIO.SASPGM(MSCMBL1)
Data: RMP.SAS.H3SM.L.MCGGIOSA.FINAL.MAIN Output: RMP.H3SO.GGIO.FINAL(MSTBLDEM)

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.

Program: RMP.H3SSGGIO.SASPGM(MSCMBL1)
Data: RMP.SAS.H3SK.L.MCGGIOSA.FINAL.MAIN Output: RMP.H3SO.GGIO.FINAL(MSTBLDEM)

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

Program: RMP.H3SGGIO.SASPGM(MSCMBL1)
Data: RMP.SAS.H3SM.L.MCGGIOSA.FINAL.MAIN Output: RMP.H3SO.GGIO.FINAL(MSTBLDEM)

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.

Program: RMP.H3SGGIO.SASPGM(MSCMBL1)
Data: RMP.SAS.H3SM.L.MCGGIOSA.FINAL.MAIN Output: RMP.H3SO.GGIO.FINAL(MSTBLDEM)

All baseline characteristics were balanced between treatment groups in the secondary prevention population, except the following two, which were not considered to be clinically relevant.

- For the primary prevention population, patient demographic characteristics were balanced between treatment groups
- For the secondary prevention population, baseline systolic blood pressure was significantly higher in raloxifene assigned patients and a significantly higher proportion of placebo assigned patients were current smokers

Risk Assessment Characteristics at Baseline (11.2.2.)

Table xxx 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 ≥1.66%.

Table xxx. 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.

Program: RMP.H3SGGIO.SASPGM(BCCTBL1)
Data: RMP.SAS.H3SK.L.MCGGIOSA.FINAL.MAIN Output: RMP.H3SO.GGIO.FINAL(BCTBLRSK)

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.

Program: RMP.H3SSGGIO.SASPGM(BCCTBL1)
Data: RMP.SAS.H3SM.L.MCGGIOSA.FINAL.MAIN Output: RMP.H3SO.GGIO.FINAL(BCTBLRSK)

Cardiovascular Risk Assessment Characteristics at Baseline

Table xxx presents CV risk assessment characteristics at baseline. CV risk assessment characteristics were balanced between treatment groups at baseline except for a significantly greater CV risk score in patients assigned to raloxifene compared with patients assigned to placebo. This difference was due to a significantly greater proportion of patients in the raloxifene group reporting a history of CABG. The magnitude of these differences is small and was not deemed clinically relevant.
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.

Table xxx. 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.

Program: RMP.H3SSGGIO.SASPGM(CVCTBL1)
Data: RMP.SAS.H3SM.L.MCGGIOSA.FINAL.MAIN Output: RMP.H3SO.GGIO.FINAL(CVTBLRSK)

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.

Program: RMP.H3SSGGIO.SASPGM(CVCTBL1)
Data: RMP.SAS.H3SM.L.MCGGIOSA.FINAL.MAIN Output: RMP.H3SO.GGIO.FINAL(CVTBLRSK)

In the primary prevention population, the CV risk assessment characteristics were balanced between treatment groups.
In the secondary prevention population, the CV risk score was significantly greater in the raloxifene group compared with the placebo group due to the significant difference in the reported history of coronary artery bypass graft (CABG). Significantly more patients in the placebo group were current smokers compared with the raloxifene group. The magnitudes of these differences are small and were not deemed clinically relevant. All other CV risk assessment characteristics were balanced between treatment groups in the secondary prevention population.

VTE Risk Assessment Characteristics at Baseline

Table GGIO.11.4 presents VTE risk assessment characteristics at baseline. The VTE risk assessment characteristics were balanced between treatment groups at baseline.

Table GGIO.11.4. VTE Risk Assessment Characteristics at Baseline (All Randomized Patients)

Characteristic	Placebo (N=5057)	Ralox (N=5044)	Total (N=10101)	p-Value*
Prior varicose veins				
No. patients	5045	5034	10079	.339
Yes, n (%)	1543(30.58)	1584(31.47)	3127(31.02)	
No, n (%)	3502(69.42)	3450(68.53)	6952(68.98)	
Prior surgery on veins in legs				
No. patients	5055	5042	10097	.305
Yes, n (%)	592(11.71)	624(12.38)	1216(12.04)	
No, n (%)	4463(88.29)	4418(87.62)	8881(87.96)	
Prior pregnancy				
No. patients	5056	5043	10099	.375
Yes, n (%)	4669(92.35)	4633(91.87)	9302(92.11)	
No, n (%)	387(7.65)	410(8.13)	797(7.89)	
Prior visible swelling in legs for > 24 hours				
No. patients	5012	5010	10022	.439
Yes, n (%)	218(4.35)	234(4.67)	452(4.51)	
No, n (%)	4794(95.65)	4776(95.33)	9570(95.49)	
Family history of deep vein clot				
No. patients	4725	4710	9435	.654
Yes, n (%)	320(6.77)	330(7.01)	650(6.89)	
No, n (%)	4405(93.23)	4380(92.99)	8785(93.11)	

Abbreviations: VTE=venous thromboembolic event.
*p-Value is obtained from a Pearson's Chi-square test.

Program: RMP.H3SSGGIO.SASPGM(CVCTBL2)
Data: RMP.SAS.H3SM.L.MCGGIOSA.FINAL.MAIN Output: RMP.H3SO.GGIO.FINAL(CVTBLVTE)

In the primary prevention population, the baseline VTE risk assessment characteristics were balanced between treatment groups. In the secondary prevention population, a significantly greater proportion of patients in the raloxifene group reported a history of varicose veins but this finding was not considered clinically relevant. All other baseline VTE risk assessment characteristics for this population were balanced between treatment groups.

Biochemical Markers of Cardiovascular Risk at Baseline

Table GGIO.11.5 presents the biochemical markers of CV risk at baseline. Biochemical markers of CV risk were balanced between treatment groups at baseline. Section 14.3.8 presents by-visit descriptive statistics for the biochemical markers of CV risk.

Table xxx. Biochemical Markers of Cardiovascular Risk at Baseline (All Randomized Patients)

Biochemical marker	Placebo (N=5057)	Raloxifene (N=5044)	Total (N=10101)	p-Value*
Total cholesterol (mmol/L)				
No. patients	4991	4964	9955	.135
Mean	5.67	5.63	5.65	
LDL cholesterol (mmol/L)				
No. patients	4977	4950	9927	.665
Mean	3.16	3.15	3.15	
HDL cholesterol (mmol/L)				
No. patients	4981	4940	9921	.608
Mean	1.36	1.35	1.36	
Non-HDL cholesterol (mmol/L)				
No. patients	4975	4938	9913	.218
Mean	4.31	4.28	4.29	
Ratio of total cholesterol to HDL cholesterol				
No. patients	4975	4938	9913	.741
Mean	4.44	4.44	4.44	
Triglycerides (mmol/L)				
No. patients	4991	4964	9955	.453
Mean	1.80	1.79	1.79	
Fibrinogen** (g/L)				
No. patients	263	258	521	.799
Mean	3.55	3.55	3.55	

Abbreviations: LDL=low-density lipoprotein; HDL=high-density lipoprotein.
*p-Value is obtained from a ranked ANOVA model: ranked response=therapy.
**Collected in only a subset of randomized patients.

Program: RMP.H3SSGGIO.SASPGM(CVCTBL3)
Data: RMP.SAS.H3SM.L.MCGGIOSA.FINAL.LABS Output: RMP.H3SO.GGIO.FINAL(CVTBLLAB)

Concomitant Medications (11.2.3.)

Table GGIO.11.6 presents the number and proportion of patients who reported use of concomitant medications during the study. Concomitant medications were grouped into categories using Anatomical Therapeutic Chemicals (ATC) codes. Lipid-lowering agents were reported as being taken by 72.64% of all randomized patients and antihypertensives by 95.96% of patients. Significantly more patients in the raloxifene group reported taking fibrates, beta-blockers, antiadrenergic agents, angiotensin receptor blockers, warfarin, thiazolidinediones, and secretagogues compared with patients in the placebo group. Significantly more patients in the placebo group reported taking bisphosphonates, HMG-CoA reductase inhibitors, and tamoxifen. The magnitudes of these differences are small and were not considered clinically relevant.

Table GGIO.11.7 presents the number and proportion of patients who reported use of selected concomitant medications at baseline and the percent change in the proportion of patients reporting use of these medications at any time post-baseline. There were no significant differences between treatment groups for baseline concomitant medications.
At baseline, 89.49% of patients reported taking an antihypertensive therapy and 54.92% reported taking a lipid-lowering agent.
From baseline to post-baseline, warfarin use increased in a significantly greater proportion of raloxifene assigned patients compared to placebo assigned patients. Use of bisphosphonates and SERMs increased in a significantly greater proportion of placebo assigned patients compared to raloxifene assigned patients. SERMs were only allowed to be taken when patients stopped study drug. The magnitudes of these differences are small and were not considered clinically relevant.

Table GGIO.14.10 and Table GGIO.14.11 present the number and proportion of patients in the primary and secondary prevention populations, respectively, who reported use of selected concomitant medications at baseline and the percent change in the proportion of patients reporting use of these medications at any time post-baseline. There were no significant differences between treatment groups for baseline concomitant medications in either of these populations. From baseline to post-baseline, HMG-CoA reductase inhibitor and SERM use increased in a significantly greater proportion of placebo assigned patients compared to raloxifene assigned patients in the primary prevention population, and warfarin and SERM use increased in a significantly greater proportion of placebo assigned patients compared to raloxifene assigned patients in the secondary prevention population. The magnitudes of these differences are small and were not considered clinically relevant.

Table xxx. Post-baseline Use of Concomitant Medications (All Randomized Patients)

Concomitant medication category	Placebo (N=5057) n (%)	Ralox (N=5044) n (%)	Total (N=10101) n (%)	p-Value*
Bone-active agents	790(15.62)	709(14.06)	1499(14.84)	.0269
Bisphosphonates	464(9.18)	404(8.01)	868(8.59)	.0366
Lipid-lowering agents	3700(73.17)	3637(72.11)	7337(72.64)	.2320
HMG-CoA reductase inhibitors	3435(67.93)	3331(66.04)	6766(66.98)	.0438
Niacin	89(1.76)	91(1.80)	180(1.78)	.8667
Fibrates	636(12.58)	729(14.45)	1365(13.51)	.0058
Bile acid sequestrants	33(0.65)	32(0.63)	65(0.64)	.9092
Cholesterol absorption inhibitors	57(1.13)	71(1.41)	128(1.27)	.2077
Anti-hypertensives	4847(95.85)	4846(96.07)	9693(95.96)	.5620
Beta-blockers	3227(63.81)	3313(65.68)	6540(64.75)	.0493
Antiadrenergic agents	465(9.20)	524(10.39)	989(9.79)	.0436
Alpha-adrenergic antagonists	387(7.65)	436(8.64)	823(8.15)	.0687
ACE inhibitors	3319(65.63)	3315(65.72)	6634(65.68)	.9242
Calcium channel blockers	2687(53.13)	2669(52.91)	5356(53.02)	.8248
Diuretics	3260(64.47)	3316(65.74)	6576(65.10)	.1784
Angiotension receptor blockers	1089(21.53)	1188(23.55)	2277(22.54)	.0152
Cardiac therapy	3052(60.35)	2990(59.28)	6042(59.82)	.2711
Inotropes	473(9.35)	492(9.75)	965(9.55)	.4932
Antiarrhythmics	546(10.80)	537(10.65)	1083(10.72)	.8068
Nitrates	2347(46.41)	2330(46.19)	4677(46.30)	.8266
Anticoagulants	3905(77.22)	3976(78.83)	7881(78.02)	.0512
Warfarin	542(10.72)	603(11.95)	1145(11.34)	.0499
Antiplatelets	3709(73.34)	3759(74.52)	7468(73.93)	.1767
Aspirin	3545(70.10)	3606(71.49)	7151(70.79)	.1245
Non-aspirin antiplatelet agents	733(14.49)	704(13.96)	1437(14.23)	.4393
Clopidogrel	459(9.08)	449(8.90)	908(8.99)	.7587
Dipyridamole	152(3.01)	154(3.05)	306(3.03)	.8895
Ticlopidine	200(3.95)	181(3.59)	381(3.77)	.3337
Pentoxifylline	316(6.25)	316(6.26)	632(6.26)	.9733
Non-steroidal anti-inflammatory drugs	2314(45.76)	2320(46.00)	4634(45.88)	.8112
COX-2 inhibitors	569(11.25)	551(10.92)	1120(11.09)	.5998
Hypoglycemic agents	2452(48.49)	2415(47.88)	4867(48.18)	.5405
Oral hypoglycemic agents	2080(41.13)	2100(41.63)	4180(41.38)	.6081
Alpha-glucosidase inhibitors	217(4.29)	241(4.78)	458(4.53)	.2396
Biguanides	1357(26.83)	1422(28.19)	2779(27.51)	.1265
Sulfonylureas and related	1574(31.13)	1555(30.83)	3129(30.98)	.7473
Thiazolidinediones	176(3.48)	226(4.48)	402(3.98)	.0101
Secretagogues	81(1.60)	115(2.28)	196(1.94)	.0135
Insulin	1134(22.42)	1097(21.75)	2231(22.09)	.4130

Abbreviations: HMG-CoA=hydroxymethylglutaryl-coenzyme A; ET=estrogen therapy; ACE=angiotensin converting enzyme; EPT=estrogen plus progestin therapy; COX-2=cyclooxygenase-2; SERM=selective estrogen receptor modulator.

*p-Value is obtained from a Pearson's Chi-square test if total>=10, Fisher's exact test if 5<=total<10, and N/A otherwise.

Program: RMP.H3SGGIO.SASPGM(MSCTMED1)
Data: RMP.SAS.H3SM.L.MCGGIOSA.FINAL.MAIN Output: RMP.H3SO.GGIO.FINAL(MSTMDPBL)

Concomitant medication category	Placebo (N=5057) n (%)	Ralox (N=5044) n (%)	Total (N=10101) n (%)	p-Value*
Sex hormones	265(5.24)	247(4.90)	512(5.07)	.4315
ET products	213(4.21)	203(4.02)	416(4.12)	.6356
EPT products	42(0.83)	36(0.71)	78(0.77)	.5025
Tibolone	6(0.12)	4(0.08)	10(0.10)	.5295
SERMs	77(1.52)	35(0.69)	112(1.11)	<.0001
Raloxifene	33(0.65)	25(0.50)	58(0.57)	.2966
Tamoxifen	44(0.87)	11(0.22)	55(0.54)	<.0001
Aromatase inhibitors	12(0.24)	8(0.16)	20(0.20)	.3737
Antidepressants	948(18.75)	887(17.59)	1835(18.17)	.1302
Thyroid therapy	724(14.32)	743(14.73)	1467(14.52)	.5553
Thyroid hormones	654(12.93)	700(13.88)	1354(13.40)	.1632

Abbreviations: HMG-CoA=hydroxymethylglutaryl-coenzyme A; ET=estrogen therapy; ACE=angiotensin converting enzyme; EPT=estrogen plus progestin therapy; COX-2=cyclooxygenase-2; SERM=selective estrogen receptor modulator.
*p-Value is obtained from a Pearson's Chi-square test if total>=10, Fisher's exact test if 5<=total<10, and N/A otherwise.

Program: RMP.H3SGGIO.SASPGM(MSCTMED1)
Data: RMP.SAS.H3SM.L.MCGGIOSA.FINAL.MAIN Output: RMP.H3SO.GGIO.FINAL(MSTMDPBL)

Table xxx. Baseline Use of Concomitant Medications and Change from Baseline to Post-baseline (All Randomized Patients)

	Baseline				% Change from baseline to posthaseline		
Concomitant medication category	Placebo (N=5057) n (%)	Ralox (N=5044) n (%)	Total (N=10101) n (%)	p-Value*	Placebo (%)	Ralox (%)	p-Value**
Bisphosphonates	138(2.73)	116(2.30)	254(2.51)	.1684	6.45	5.71	.0125
Lipid-lowering agents	2750(54.38)	2797(55.45)	5547(54.92)	.2790	18.79	16.65	.9931
HMG-CoA reductase inhibitors	2361(46.69)	2382(47.22)	4743(46.96)	.5889	21.24	18.81	.3225
Anti-hypertensives	4522(89.42)	4517(89.55)	9039(89.49)	.8296	6.43	6.52	.6211
Nitrates	1710(33.81)	1686(33.47)	3396(33.64)	.7105	12.60	12.73	.6786
Warfarin	195(3.86)	222(4.40)	417(4.13)	.1684	6.86	7.55	.0167
Aspirin	2865(56.65)	2846(56.42)	5711(56.54)	.8151	13.45	15.07	.3865
Non-aspirin antiplatelet agents	142(2.81)	156(3.09)	298(2.95)	.3977	11.69	10.86	.7435
Non-steroidal anti-inflammatory drugs	599(11.84)	626(12.41)	1225(12.13)	.3838	33.91	33.58	.4983
COX-2 inhibitors	37(0.73)	31(0.61)	68(0.67)	.4719	10.52	10.31	.4908
Oral hypoglycemic agents	1711(33.83)	1716(34.02)	3427(33.93)	.8432	7.30	7.61	.6124
Insulin	726(14.36)	694(13.76)	1420(14.06)	.3878	8.07	7.99	.2372
Sex hormones	71(1.40)	74(1.47)	145(1.44)	.7898	3.84	3.43	.5725
SERMs	0(0.00)	0(0.00)	0(0.00)	N/A	1.52	0.69	<.0001
Aromatase inhibitors	0(0.00)	0(0.00)	0(0.00)	N/A	0.24	0.16	.3737

Abbreviations: HMG-CoA=hydroxymethylglutaryl-coenzyme A; COX-2=cyclooxygenase-2; SERM=selective estrogen receptor modulator.
*p-Value is obtained from a Pearson's Chi-square test if total>=10, Fisher's exact test if 5<=total<10, and N/A otherwise.
**p-Value is obtained from a Cochran-Mantel-Haenszel (CMH) test, stratified by an indicator variable of baseline.

Program: RMP.H3S5GGIO.SASPGM(MSCTMED2) Data: RMP.SAS.H3SM.L.MCGGIOSA.FINAL.MAIN Output: RMP.H3S0.GGIO.FINAL(MSTMDCHG)