Comparative Effectiveness Review

Number 17

Comparative Effectiveness of Medications To Reduce Risk of Primary Breast Cancer in Women

Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • **www.ahrq.gov**

This report is based on research conducted by the Oregon Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-2007-10057-1). The findings and conclusions in this document are those of the authors, who are responsible for its contents, and do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help clinicians, policymakers, and others make informed decisions about the provision of health care services. This report is intended as a reference and not as a substitute for clinical judgment.

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

Comparative Effectiveness Review

Number 17

Comparative Effectiveness of Medications To Reduce Risk of Primary Breast Cancer in Women

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. 290-2007-10057-1

Prepared by:

Oregon Evidence-based Practice Center

Investigators
Heidi D. Nelson, M.D., M.P.H.
Rochelle Fu, Ph.D.
Linda Humphrey, M.D., M.P.H.
M. E. Beth Smith, D.O.
Jessica C. Griffin, M.S.
Peggy Nygren, M.A.

AHRQ Publication No. 09-EHC028-EF
September 2009

None of the investigators has any affiliations or financial involvement that conflicts with the material presented in this report.

Suggested citation: Nelson HD, Fu R, Humphrey L, Smith ME, Griffin JC, Nygren P. Comparative Effectiveness of Medications To Reduce Risk of Primary Breast Cancer in Women. Comparative Effectiveness Review No. 17. (Prepared by Oregon Evidence-based Practice Center under Contract No. 290-2007-10057-1.) Rockville, MD: Agency for Healthcare Research and Quality. September 2009. Available at: www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Preface

The Agency for Healthcare Research and Quality (AHRQ) conducts the Effective Health Care Program as part of its mission to organize knowledge and make it available to inform decisions about health care. As part of the Medicare Prescription Drug, Improvement, and Modernization Act of 2003, Congress directed AHRQ to conduct and support research on the comparative outcomes, clinical effectiveness, and appropriateness of pharmaceuticals, devices, and health care services to meet the needs of Medicare, Medicaid, and the State Children's Health Insurance Program (SCHIP).

AHRQ has an established network of Evidence-based Practice Centers (EPCs) that produce Evidence Reports/Technology Assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care. The EPCs now lend their expertise to the Effective Health Care Program by conducting Comparative Effectiveness Reviews of medications, devices, and other relevant interventions, including strategies for how these items and services can best be organized, managed, and delivered.

Systematic reviews are the building blocks underlying evidence-based practice; they focus attention on the strength and limits of evidence from research studies about the effectiveness and safety of a clinical intervention. In the context of developing recommendations for practice, systematic reviews are useful because they define the strengths and limits of the evidence, clarifying whether assertions about the value of the intervention are based on strong evidence from clinical studies. For more information about systematic reviews, see http://effectivehealthcare.ahrq.gov/reference/purpose.cfm.

AHRQ expects that Comparative Effectiveness Reviews will be helpful to health plans, providers, purchasers, government programs, and the health care system as a whole. In addition, AHRQ is committed to presenting information in different formats so that consumers who make decisions about their own and their family's health can benefit from the evidence.

Transparency and stakeholder input are essential to the Effective Health Care Program. Please visit the Web site (www.effectivehealthcare.ahrq.gov) to see draft research questions and reports or to join an e-mail list to learn about new program products and opportunities for input. Comparative Effectiveness Reviews will be updated regularly.

Carolyn M. Clancy, M.D.
Director
Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.
Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Acknowledgments

The authors thank Andrew Hamilton, M.L.S, M.S., and Rose Campbell, M.L.I.S., M.S., for literature searches, and Jennifer Nguyen for administrative assistance at the Oregon Evidence-based Practice Center at the Oregon Health & Science University. We also acknowledge the contributions of AHRQ Officers Shilpa Amin, M.D., M.B.Sc., and Kenneth Lin, M.D., and members of the Technical Expert Panel and Peer Reviewers listed below.

Technical Expert Panel

Katrina Armstrong, M.D., University of Pennsylvania, Philadelphia, Pennsylvania
Elizabeth Barrett-Connor, M.D., University of California at San Diego, San Diego, California
William Gradishar, M.D., Northwestern University, Evanston, Illinois
Linda Kinsinger, M.D., Veterans Health Administration, Durham, North Carolina
Diana Petitti, M.D., University of Arizona, Phoenix, Arizona
George Sawaya, M.D., University of California San Francisco, San Francisco, California
Victor Vogel, M.D., Magee Women's Hospital, Pittsburgh, Pennsylvania

Peer Reviewers

Nananda Col, M.D., M.P.P., M.P.H., FACP
Brian E. Henderson, M.D.
Linda S. Kinsinger, M.D.
Joy Simha

AHRQ Contacts

Beth A. Collins-Sharp, Ph.D., R.N.
Director
Evidence-based Practice Center Program
Agency for Healthcare Research and Quality
Rockville, MD

Shilpa H. Amin, M.D., M.B.Sc, FAAFP
Task Order Officer
Evidence-based Practice Center Program
Agency for Healthcare Research and Quality
Rockville, MD

Kenneth W. Lin, M.D.
Medical Officer
U.S. Preventive Services Task Force Program
Center for Primary Care, Prevention, and Clinical Partnerships
Agency for Healthcare Research and Quality
Rockville, MD

Contents

Tables:

Figures:

Appendixes

Executive Summary

Background

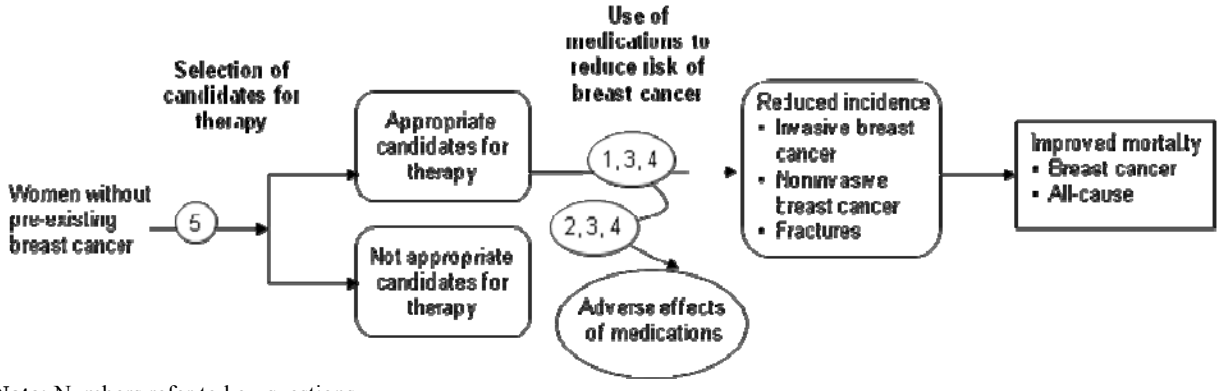
Breast cancer is the most frequently diagnosed noncutaneous cancer and the second leading cause of cancer death after lung cancer among women in the United States. In 2008, an estimated 182,460 cases of invasive breast cancer and 67,770 cases of *in situ* breast cancer were diagnosed, and 40,480 women died of breast cancer in the United States.

Recent clinical trials have demonstrated the efficacy of three medications—tamoxifen citrate, raloxifene, and tibolone—to reduce the risk of invasive breast cancer in women without pre-existing cancer. This therapy is sometimes referred to as "chemoprevention" in the literature, although this is not a fully accurate representation of the intervention. Tamoxifen and raloxifene are approved by the U.S. Food and Drug Administration for this indication and tibolone is not. Raloxifene is approved for use by postmenopausal women only. Current clinical recommendations, including those from the U.S. Preventive Services Task Force issued in 2002, support tamoxifen use for primary breast cancer prevention in women considered at high risk for breast cancer by the Gail model or other criteria and low risk for adverse events. However, use of risk-reducing medications for breast cancer is believed to be low in the United States.

The purpose of this review is to evaluate the comparative effectiveness of tamoxifen citrate, raloxifene, and tibolone to reduce the risk of primary breast cancer; assess the nature and magnitude of harms; and examine how benefits and harms vary by age, breast cancer risk status, and other factors. The review was originally entitled "Comparative Effectiveness of Chemotherapy Agents in the Prevention of Primary Breast Cancer in Women." Peer review comments suggested that the terms "chemotherapy" and "prevention" were misnomers. The term "medications to reduce risk" is a better representation of the intervention and therefore, all references to "chemoprevention" are edited, including the key questions and report title.

The review also examines issues related to clinical effectiveness, such as patient choice, concordance, adherence, and persistence of use, and evaluates methods to appropriately select patients for risk-reducing medications for clinical applications. The target population includes women without pre-existing breast cancer, noninvasive breast cancer, or precursor conditions who are not known carriers of breast cancer susceptibility mutations (BRCA1, BRCA2, or others). The analytic framework and key questions guiding this review are described below.

Figure A. Analytic framework



Note: Numbers refer to key questions.

Key Question 1. In adult women without pre-existing breast cancer, what is the comparative effectiveness of selective estrogen receptor modulators (SERMs) tamoxifen citrate and raloxifene, and the selective tissue estrogenic activity regulator (STEAR) tibolone, when used to reduce risk for primary breast cancer on improving short-term and long-term outcomes including invasive breast cancer, noninvasive breast cancer, including ductal carcinoma *in situ* (DCIS), breast cancer mortality, all-cause mortality, and osteoporotic fractures?

Key Question 2. What is the evidence for harms of tamoxifen citrate, raloxifene, and tibolone when used to reduce risk for primary breast cancer?

Key Question 3. How do outcomes for tamoxifen citrate, raloxifene, and tibolone when used for primary prevention of breast cancer vary by heterogeneity in subpopulations?

Key Question 4. What is the evidence that harms or secondary potential benefits listed above affect treatment choice, concordance, adherence, and persistence to treatment with tamoxifen citrate, raloxifene, and tibolone when used for primary prevention of breast cancer?

Key Question 5. What methods, such as clinical risk-assessment models, have been used to identify women who could benefit from medications to reduce risk of breast cancer?

Conclusions

Key Question 1. Comparative effectiveness of tamoxifen citrate, raloxifene, and tibolone for the primary prevention of breast cancer, mortality, and fractures:

- Eight large randomized controlled trials provide data on breast cancer risk reduction in women without pre-existing breast cancer. These include one good-quality head-to-head trial of tamoxifen and raloxifene and seven fair- and good-quality placebo-controlled trials (four tamoxifen, two raloxifene, and one tibolone). Results of placebo-controlled trials cannot be directly compared between types of medications because of important differences between study subjects.
- Tamoxifen (risk ratio [RR] 0.70; 0.59, 0.82; four trials), raloxifene (RR 0.44; 0.27, 0.71; two trials), and tibolone (RR 0.32; 0.13, 0.80; one trial) reduce the incidence of invasive breast cancer in midlife and older women by approximately 30 percent to 68 percent. Tamoxifen and raloxifene had similar effects in the STAR (Study of Raloxifene and Tamoxifen) head-to-head trial.
- Reduction of invasive breast cancer continued at least 3 to 5 years after discontinuation of tamoxifen in the two trials providing post-treatment followup data.
- Tamoxifen (RR 0.58; 0.42, 0.79; four trials) and raloxifene (RR 0.33; 0.18, 0.61; two trials) reduced estrogen receptor positive invasive breast cancer, but not estrogen receptor negative invasive breast cancer, in placebo-controlled trials. They had similar effects in the STAR head-to-head trial.
- Tamoxifen and raloxifene did not significantly reduce noninvasive breast cancer, including DCIS, in meta-analysis of four placebo-controlled trials, although noninvasive breast cancer was significantly reduced in the NSABP P-1 (National Surgical Adjuvant

Breast and Bowel Project) tamoxifen trial (RR 0.63; 0.45, 0.89). The STAR head-to-head trial indicated no statistically significant differences between raloxifene and tamoxifen (RR 1.40; 0.98, 2.00).

- All-cause mortality is similar for women using raloxifene and those using tamoxifen, and also is similar for tamoxifen, raloxifene, or tibolone compared with placebo, although followup times in most trials were short. Tamoxifen does not reduce breast cancer mortality compared to placebo.
- Tamoxifen and raloxifene had similar effects on fractures at multiple sites in the STAR head-to-head trial. In placebo-controlled trials, raloxifene (RR 0.61; 0.54, 0.69; two trials) and tibolone (RR 0.55; 0.41. 0.74; one trial) reduced vertebral fractures; tamoxifen (RR 0.66; 0.45, 0.98; one trial) and tibolone (RR 0.74; 0.58, 0.93; one trial) reduced nonvertebral fractures; and tibolone reduced wrist (RR 0.54; 0.35, 0.82; one trial) but not hip fractures.

Table A. Summary of primary prevention trials–benefits: number of events reduced with medications and strength of evidence

Major health outcome	Head-to-head trial[a]	Placebo-controlled trials[b]		
	Raloxifene vs. tamoxifen	Tamoxifen vs. placebo	Raloxifene vs. placebo	Tibolone vs. placebo
Invasive breast cancer	No difference	7 (4, 12) +++	9 (4, 14) +++	10 (3, 17) ++
Estrogen receptor positive	No difference	8 (3, 13) +++	8 (4, 12) +++	Insufficient
Estrogen receptor negative	No difference	No difference ++	No difference ++	Insufficient
Noninvasive cancer	No difference	No difference +	No difference ++	Insufficient
All-cause death[c]	No difference	No difference +++	No difference +++	Insufficient
Vertebral fracture	No difference	No difference +	7 (5, 9) +++	44 (25, 61) ++
Nonvertebral fracture	Insufficient	3 (0.2, 5) ++	No difference +++	34 (8, 56) ++

[a]Study of Raloxifene and Tamoxifen (STAR).
[b]Number of events reduced compared to placebo per ,1000 women-years assuming 5 years of use (95-percent confidence interval shown in parentheses).
[c]Based on short-term followup times from trials.

Strength of Evidence Symbols

+++	High: Consistent results from numerous (>5) or large definitive trials show a positive protective effect.
++	Moderate: Some evidence (3-5 studies) suggests a protective effect, but results could be altered by future research.
+	Low: Few (≤2) trials exist, existing trials have inconsistent results and/or limitations, results are likely to be altered by future research.
No difference	Results are not statistically significantly different.
Insufficient	Data are inadequate to calculate outcomes or are not reported.

Key Question 2. Harms of tamoxifen citrate, raloxifene, and tibolone when used for primary prevention of breast cancer:

- In addition to the 8 large randomized controlled trials described in Key Question 1, harms data were provided by 12 placebo-controlled trials and 1 observational study of raloxifene, and 7 placebo-controlled trials and 1 observational study of tibolone.

- Raloxifene caused fewer thromboembolic events (RR 0.70; 0.54, 0.91) than tamoxifen in the STAR head-to-head trial. Tamoxifen (RR 1.93; 1.41, 2.64; four trials) and raloxifene (RR 1.60; 1.15, 2.23; two trials) cause more thromboembolic events than placebo. Risk returned to normal after discontinuation of tamoxifen in the two trials providing post-treatment data. Tibolone does not increase risk for thromboembolic events, although data are limited.
- Tamoxifen, raloxifene, and tibolone do not increase risk for coronary heart disease events, although data for tibolone are limited.
- Tibolone causes more strokes than placebo (RR 2.19; 1.14, 4.23); tamoxifen and raloxifene do not increase risk for stroke.
- In the STAR head-to-head trial, raloxifene caused fewer cases of endometrial hyperplasia (RR 0.16; 0.09, 0.29) and was associated with fewer hysterectomies (RR 0.44; 0.35, 0.56) than tamoxifen, but differences for endometrial cancer were not statistically significant (RR 0.62; 0.35, 1.08).
- Tamoxifen causes more cases of endometrial cancer than placebo (RR 2.13; 1.36, 3.32; three trials); raloxifene does not increase risk for endometrial cancer or uterine bleeding, and tibolone does not increase risk for endometrial cancer in clinical trials but was associated with more cases of endometrial cancer in a large cohort study (RR 1.79; 1.43, 2.25).
- Raloxifene caused fewer cataracts (RR 0.79; 0.68, 0.92) and cataract surgeries (RR 0.82; 0.68, 0.99) than tamoxifen in the STAR head-to-head trial. Tamoxifen was associated with more cataract surgeries than placebo in the NSABP P-1 trial (RR 1.57; 1.16, 2.14). Raloxifene does not increase risk for cataracts or cataract surgery.
- In head-to-head comparisons, women using raloxifene reported more musculoskeletal problems, dyspareunia, and weight gain, while those using tamoxifen had more gynecological problems, vasomotor symptoms, leg cramps, and bladder control symptoms.
- Most common side effects for tamoxifen are hot flashes and other vasomotor symptoms, vaginal discharge, and other vaginal symptoms such as itching or dryness; for raloxifene, vasomotor symptoms and leg cramps; and for tibolone, vaginal bleeding and reduced number and severity of hot flashes.

Table B. Summary of primary prevention trials–harms: number of events increased with medications and strength of evidence

Major health outcome	Head-to-head trial[a]	Placebo-controlled trials[b]		
	Raloxifene vs. tamoxifen	Tamoxifen vs. placebo	Raloxifene vs. placebo	Tibolone vs. placebo
Thromboembolic events	6 (2, 10)[c] More with tamoxifen	4 (2, 9) +++	7 (2, 15) +++	No difference +
Coronary heart disease	No difference	No difference +++	No difference +++	No difference +
Stroke	No difference	No difference ++	No difference ++	11 (1, 36) ++
Endometrial cancer	No difference	4 (1, 10) +++	No difference ++	Insufficient
Cataracts	13 (5, 21) More with tamoxifen	No difference +	No difference +++	Insufficient

[a]Study of Raloxifene and Tamoxifen (STAR).
[b]Number of events increased compared to placebo per 1,000 women-years assuming 5 years of use (95-percent confidence interval).
[c]Number of events increased per 1,000 women-years assuming 5 years of use (95-percent confidence interval).

Strength of Evidence Symbols

+++	High: Consistent results from numerous (>5) or large definitive trials show a harmful effect.
++	Moderate: Some evidence (3-5 studies) suggests a harmful effect, but results could be altered by future research.
+	Low: Few (≤2) trials exist, existing trials have inconsistent results and/or limitations, results are likely to be altered by future research.
No difference	Results are not statistically significantly different.
Insufficient	Data are inadequate to calculate outcomes or are not reported.

Key Question 3. Variability of outcomes in subpopulations:

- Tamoxifen and raloxifene had similar effects on breast cancer outcomes regardless of age and family history of breast cancer in the head-to-head STAR trial.
- Tamoxifen reduces breast cancer outcomes in subgroups evaluated in prevention trials based on age, menopausal status, estrogen use, family history of breast cancer, and history of lobular carcinoma *in situ* or atypical hyperplasia. In the NSABP P-1 trial, cancer rates were highest and risk reduction greatest among women in the highest modified Gail model risk category and among women with prior atypical hyperplasia.
- Raloxifene reduces breast cancer outcomes in subgroups evaluated in prevention trials based on age, age at menarche, parity, age at first live birth, and body mass index. Estimates from subgroups based on prior estrogen use, family history of breast cancer, and prior hysterectomy or oophorectomy are limited by smaller numbers of subjects.
- Thromboembolic events and endometrial cancer were more common in older (>50) than younger women in the NSABP P-1 trial.
- Tibolone causes more strokes in older (>70 years) than younger women.

Key Question 4. Treatment choice, concordance, adherence, and persistence to treatment with tamoxifen citrate, raloxifene, and tibolone when used for primary prevention of breast cancer:

- Comparisons of adherence and persistence rates across medications in prevention trials are limited because few trials report treatment duration, completion rates, or other measures of adherence and persistence, and trials were designed for different treatment purposes.
- Discontinuation rates for tamoxifen or raloxifene are generally higher than placebo. In the few trials reporting discontinuation rates, the difference between treatment and placebo groups was ≤2 percent for adverse events and ≤4 percent for nonprotocol-specified events.
- Women make decisions to use tamoxifen for risk reduction based on their concern for adverse effects as well as their risk for breast cancer, according to small descriptive studies.
- Women weigh their physicians' recommendations highly when deciding whether to take tamoxifen for risk reduction, according to descriptive studies of concordance.
- Studies of treatment choice and concordance for raloxifene and tibolone for breast cancer risk reduction are lacking.

Key Question 5. Clinical risk assessment models to identify women who could benefit from medications to reduce risk of breast cancer:

- Nine risk stratification models that predict an individual's risk for developing breast cancer have been evaluated for use in clinical settings. Models consider multiple risk factors for breast cancer.
- Risk stratification models demonstrate good calibration, with the expected number of breast cancer cases in a study population closely matching the number of breast cancer cases observed.
- All models have low discriminatory accuracy in predicting the probability of breast cancer in an individual. Most models perform only slightly better than age alone as a risk predictor.
- A Gail score of ≥1.66 percent has been used as a risk threshold in prevention trials and in Food and Drug Administration approval of tamoxifen and raloxifene for breast cancer prevention. However, this threshold has low discriminatory accuracy in predicting breast cancer in an individual.

Applicability

Trials met criteria for good applicability: they were conducted in settings appropriate to clinical practice, enrolled subjects selected with broad eligibility criteria, assessed health outcomes, and had followup periods of several years. Also, although inclusion criteria differed between trials, results for breast cancer outcomes were similar. For these reasons, the trials provided information about effectiveness as well as efficacy of the risk-reducing medications.

Clinicians can consider the results of trials to be most applicable to patients with characteristics similar to those of the study populations. Specifically, tamoxifen results apply to younger premenopausal and postmenopausal women meeting breast cancer risk criteria; tibolone results apply to older postmenopausal women with osteoporosis; and raloxifene results apply to

postmenopausal women meeting breast cancer risk criteria and to older postmenopausal women with osteoporosis or cardiovascular disease and/or risk factors for cardiovascular disease. Women not well represented in the trials are those who are younger (<55 years old), have Gail scores <1.66 percent or considered low risk by other criteria used by some of the trials, are nonwhite, or are from outside North America and Europe. Also, premenopausal women were excluded from the raloxifene and tibolone trials.

Remaining Issues

While the efficacy of tamoxifen, raloxifene, and tibolone has been demonstrated for women in the clinical trials, it is not clear which women in clinical practice would optimally benefit from risk reduction. Future research to determine the optimal candidates for risk-reduction medications would help focus prevention efforts. Applying these findings to clinical selection criteria would improve identification of patients for risk-reducing medications in practice.

The results of current trials indicate that adverse effects differ between medications and may drive decisions for risk-reducing medications as much or more than benefits do. Further research to more clearly identify characteristics of individuals experiencing specific adverse effects would guide physicians and patients to regimens that cause the least harm.

Introduction

Background

The purpose of this review is to evaluate the comparative effectiveness of tamoxifen citrate, raloxifene, and tibolone to reduce risk for primary breast cancer, assess the nature and magnitude of harms, and examine how benefits and harms vary by age, breast cancer risk status, and other factors. In addition, it examines issues related to clinical effectiveness, such as patient choice, concordance, adherence, and persistence of use, and evaluates methods to appropriately select patients for medication therapy to reduce risk of breast cancer.

Breast cancer is the most frequently diagnosed non-cutaneous cancer and the second leading cause of cancer death after lung cancer among women in the United States.[1] In 2008, an estimated 182,460 cases of invasive breast cancer and 67,770 cases of *in situ* breast cancer were diagnosed, and 40,480 women died of breast cancer.[2] The National Cancer Institute estimates that 14.7% of women born today will develop breast cancer in their lifetimes, 12.3% with invasive disease.[2] The probability of a woman developing breast cancer in her forties is 1 in 69, in her fifties 1 in 38, and in her sixties 1 in 27.[3]

Breast cancer is a proliferation of malignant cells that arises in the breast tissue, specifically in the terminal ductal-lobular unit. Breast cancer represents a continuum of disease, ranging from noninvasive to invasive carcinoma.[4] Noninvasive carcinoma is confined to either the mammary duct, as with ductal carcinoma *in situ* (DCIS), or to the lobule, as with lobular carcinoma *in situ* (LCIS). LCIS is not considered a precursor lesion for invasive lobular carcinoma, but believed to be a marker for increased risk of invasive ductal or lobular breast cancer development in either breast.[5] DCIS is thought to be a precursor lesion to invasive ductal carcinoma. Unlike *in situ* lesions, invasive breast cancers have metastatic potential.

Although several risk factors have been associated with breast cancer, most cases occur in women with no specific risk factors other than sex and age. Family history of breast and ovarian cancer are strong risk determinants. Family history is further characterized by the number of affected relatives, closeness of the degree of relationships, and ages of diagnosis. Although uncommon, hereditary mutations in tumor suppressor genes *BRCA1* and *BRCA2* increase individual risks for breast cancer 60-85% and may be identified in 5-10% of all breast cancer cases.[6]

Personal history of *in situ* breast cancer, previous abnormal breast biopsy containing LCIS, or atypical ductal or lobular hyperplasia increase risk for invasive breast cancer. High mammographic breast density is also associated with increased risk of breast cancer.[7,8] Endogenous estrogen exposure is associated with increased risk; thus early menarche, late menopause, older age at birth of first child, nulliparity, and obesity are implicated as risk factors. Use of combination postmenopausal hormone therapy (estrogen and progestin) was associated with an increased risk for breast cancer compared to placebo in the Women's Health Initiative (WHI) randomized controlled trial.[9] Use of alcohol at levels more than 1 to 2 drinks per day is also associated with increased breast cancer.[9]

Recent clinical trials have demonstrated the efficacy of tamoxifen citrate and raloxifene, selective estrogen receptor modulators (SERM), and the selective tissue estrogenic activity regulator (STEAR) tibolone, to reduce the risk of invasive breast cancer in women without pre-existing cancer (Table 1). Tamoxifen is approved by the U.S. Food and Drug Administration (FDA) to reduce the incidence of breast cancer in women at high risk of developing the disease

defined as those with a breast biopsy with lobular carcinoma *in situ* or atypical hyperplasia, one or more first-degree relatives with breast cancer, or a 5-year predicted risk of breast cancer of ≥1.66% calculated by the modified Gail model. Tamoxifen is primarily used for the treatment of early and advanced estrogen receptor positive breast cancer in pre and postmenopausal women and for reduction of contralateral breast cancer. Raloxifene was initially approved by the FDA for osteoporosis prevention (1997) and treatment (1999) and has been primarily used for these indications. In September 2007, the FDA approved raloxifene for reducing the risk of invasive breast cancer in postmenopausal women with osteoporosis and in postmenopausal women at high risk for invasive breast cancer.

Tibolone is not currently approved by the FDA for use in the United States, but is approved to treat menopausal symptoms in 90 countries, and to prevent osteoporosis in 45 countries.[10] Tibolone became available in the U.K. in the early 1990's, and since then nearly 9 million women per year have taken it worldwide.[11] A recent evaluation of tibolone's safety profile concluded that it is comparable to combined menopausal hormone therapy, and prescribing considerations for older women need to be taken into account for increased risk of stroke.[11]

Current clinical recommendations, including those from the U.S. Preventive Services Task Force (USPSTF) issued in 2002, support tamoxifen use to reduce risk for primary breast cancer in women considered at high risk for breast cancer by the Gail model or other criteria and low risk for adverse events. However, use of risk reducing medications for breast cancer is believed to be low in the United States.[12] Primary care clinicians cite potential adverse effects, ranging from thromboembolism to hot flashes, as deterrents to prescribing tamoxifen to women without breast cancer. Now that raloxifene has also demonstrated breast cancer risk reduction benefits, recommendations need to be updated to include the most recent trials.

Scope and Key Questions

This report summarizes the available evidence comparing the effectiveness and safety of tamoxifen, raloxifene, and tibolone to reduce risk for primary breast cancer in women. The target population includes women without pre-existing breast cancer, noninvasive breast cancer, or precursor conditions who are not known carriers of breast cancer susceptibility mutations (*BRCA1, BRCA2,* or others). The report addresses the following questions.

Key Question 1. In adult women without pre-existing breast cancer, what is the comparative effectiveness of selective estrogen receptor modulators (SERMs) tamoxifen citrate and raloxifene, and the selective tissue estrogenic activity regulator (STEAR) tibolone, when used for the primary prevention of breast cancer on improving short-term and long-term outcomes including:

- **Invasive breast cancer**
- **Noninvasive breast cancer including ductal carcinoma in situ (DCIS)**
- **Breast cancer mortality**
- **All-cause mortality**
- **Osteoporotic fractures**

Key Question 2. What is the evidence for harms of tamoxifen citrate, raloxifene, and tibolone when used for primary prevention of breast cancer?
Harms may include but are not limited to:

- **Thromboembolic events (deep vein thrombosis, pulmonary embolism)**
- **Cardiovascular events (coronary heart disease, stroke and transient ischemic attack, arrhythmias)**
- **Metabolic disorders**
- **Musculoskeletal symptoms (myalgia, leg cramps)**
- **Mental health (mood changes, other)**
- **Genitourinary outcomes (vaginal dryness, vaginal discharge, dyspareunia, sexual dysfunction, endometrial hyperplasia, abnormal uterine bleeding, other benign uterine conditions, hysterectomy, endometrial cancer, urinary symptoms, other)**
- **Breast outcomes (biopsies, breast density, other)**
- **Other malignancies (incidence, death)**
- **Ophthalmologic disorders (cataracts, other)**
- **Gastrointestinal/hepatobiliary disorders**
- **Other adverse events that would impact quality of life (vasomotor symptoms, sleep disturbances, headaches, cognitive/memory changes, peripheral edema)**

Key Question 3. How do outcomes for tamoxifen citrate, raloxifene, and tibolone when used for primary prevention of breast cancer vary by heterogeneity in subpopulations? Subpopulations include but are not limited to:

- **Age**
- **Menopausal status (pre-, peri-, postmenopausal)**
- **Hysterectomy status**
- **Use of exogenous estrogen**
- **Level of risk of breast cancer (family history, body mass index, parity [number of pregnancies], age at first live birth, age at menarche, personal history of breast abnormalities, prior breast biopsy, estradiol levels, breast density)**
- **Ethnicity and race**
- **Metabolism status (CYP 2D6 mutation)**
- **Risk for thromboembolic events (obesity, others)**

Key Question 4. What is the evidence that harms or noncancer benefits listed above affect treatment choice, concordance, adherence, and persistence to treatment with tamoxifen citrate, raloxifene, and tibolone when used for primary prevention of breast cancer?

Key Question 5. What methods, such as clinical risk assessment models, have been used to identify women who could benefit from breast cancer medications to reduce risk of breast cancer?

Methods

Topic Development

The topic for this comparative effectiveness review was nominated in a public process. With input from technical experts, the Scientific Resource Center (SRC) for the AHRQ Effective Health Care Program drafted the initial key questions and, after approval from AHRQ, posted them to a public web site. The public was invited to comment on these questions. After reviewing the public commentary, the SRC drafted final key questions and submitted them to AHRQ for approval.

The key questions went through three subsequent revisions. After discussions with a technical expert panel, the key questions were further refined to identify specific outcomes of interest for key questions 1, 2, and 3. These changes were submitted to AHRQ for approval before literature searches were conducted. The second change to the key questions occurred in September 2008 after publication of a new study of breast cancer risk reduction with the medication tibolone. After discussion with AHRQ, it was determined that the current report would be amended to include tibolone. New key questions including tibolone were then approved by AHRQ. The third change was in response to peer reviewers who suggested that the terms "chemotherapy" and "prevention" were misnomers. The term "medications to reduce risk" is a better representation of the intervention. Therefore, all references to "chemoprevention" were edited, including the key questions and report title.

We created an analytic framework incorporating the key questions to guide our examination of a chain of evidence about the effectiveness and potential adverse effects of medications to reduce risk of primary breast cancer (Figure 1). The analytic framework outlines the target population, interventions, and outcomes defined by the scope of this review. The target population includes women without pre-existing invasive or noninvasive breast cancer or precursor conditions, and who are not known carriers of breast cancer susceptibility mutations (*BRCA1, BRCA2,* or others). Outcomes are defined by the key questions and include a wide range of health outcomes as opposed to intermediate outcomes. Health outcomes are signs, symptoms, conditions, or events that individuals experience, such as myocardial infarction. Intermediate outcomes are health measures that individuals do not personally experience, such as laboratory test results.

Search Strategy

We used the National Library of Medicine's Medical Subject Headings (MeSH®) keyword nomenclature developed for MEDLINE® and adapted for use in other databases. With assistance from a research librarian, we searched OVID MEDLINE® (1950 to January Week 3, 2009), Cochrane Central Register of Controlled Trials (4th Quarter 2008), and Cochrane Database of Systematic Reviews (4th Quarter 2008) for relevant studies, systematic reviews, and meta-analyses. The searches were limited to papers published in English language. The texts of the major search strategies are provided in Appendix A1. We also searched clinical trial registries and conducted secondary referencing by manually reviewing reference lists of papers and reviewing citations indicated for key trials by Web of Science.® After identifying several large trials meeting inclusion criteria for the review, we contacted the investigators to request

additional unpublished data specifically addressing the subpopulations described in key question 3. No additional data have been received.

In addition, we received the following materials from the Scientific Resource Center:

- Searches of clinical trial registries: www.clinicaltrials.gov; www.controlled-trials.com; www.clinicalstudyresults.org; www.Drugs@FDA.gov; and the American Society of Clinical Oncology website: (http://www.asco.org/ASCO/Abstracts+%26+Virtual+Meeting/Abstracts).
- Scientific information packet from Eli Lilly for Evista® (raloxifene). Scientific information packets were requested from manufacturers of tamoxifen, however no information was provided. A scientific information packet was requested from the manufacturers of tibolone by the Scientific Resource Center in November 2008. As of publication of this draft, no information has been received.

The searches identified a total of 4,842 unique citations. Some citations were relevant to multiple key questions. Investigators reviewed 4,230 citations for key questions 1, 2, and 3; 1,644 citations for key question 4; and 1,364 citations for key question 5. All citations were imported into an electronic database (EndNote X1).

Study Selection

Prior to our review of abstracts and articles, we developed inclusion and exclusion criteria for studies based on the patient populations, interventions, outcome measures, and types of evidence specified in each key question (Appendix A2). We applied these criteria to the abstracts and articles identified by our searches. After an initial review of citations and abstracts, we retrieved full-text articles of potentially relevant material and conducted a second review to determine inclusion. A second reviewer confirmed results of the initial reviewer. Articles with questionable eligibility were reviewed and discussed by the investigator team before determining their inclusion. Results published only in abstract form were not included in our review because adequate information was not available to assess the validity of the data. Excluded studies and their main reasons for exclusion are listed in Appendix B.

For key question 1 and any outcomes relating to risk reduction benefits for key question 3, we included only randomized controlled trials (RCT) of tamoxifen, raloxifene, or tibolone for primary prevention of breast cancer enrolling women without breast cancer. We included trials that were designed and powered to demonstrate invasive breast cancer incidence as a primary or secondary outcome. The technical expert panel advised including only RCTs for several reasons. These include lack of observational studies of tamoxifen and raloxifene with breast cancer outcomes in women without breast cancer, and concerns for bias among users in observational studies. For example, women using tibolone to treat menopausal hot flashes are more likely to have a hysterectomy/oophorectomy than nonusers, reducing their breast cancer risk.

For key question 2 and outcomes relating to harms for key question 3, we defined our inclusion criteria more broadly. We included RCTs and observational studies of tamoxifen, raloxifene, or tibolone in women without breast cancer that were designed for multiple types of outcomes. However, studies must have had a nonuser comparison group, or direct comparisons between tamoxifen, raloxifene, or tibolone to be included. We included studies with treatment durations of 3 months or more that enrolled 100 or more participants to assure adequate drug exposure and power to support results.

For key question 4, RCTs, observational studies, and descriptive studies evaluating benefits or harms and treatment adherence, persistence, concordance, or treatment choice with tamoxifen, raloxifene, or tibolone in women without breast cancer were included.

For key question 5, we included studies of risk stratification models that could be used in a primary care setting to identify women at higher than average risk for breast cancer. Only studies reporting discriminatory accuracy of the models were included. We did not include models designed to evaluate family history in order to determine risk for deleterious *BRCA* mutations because women with these mutations are outside the target population for this review. We also excluded studies of single risk factors or laboratory tests.

Data Extraction

For the included RCTs and observational studies, we abstracted the following data: study design; setting; participant characteristics (including age, ethnicity, diagnosis); enrollment criteria; interventions (dose and duration); numbers enrolled and lost to follow-up; methods of outcome ascertainment; and results for each outcome. For descriptive studies of treatment choice, we abstracted: study design; intervention; setting; population characteristics; eligibility and exclusion criteria; response rates; procedure for data collection; and results for each outcome. For studies of risk stratification models, we abstracted: study design; population characteristics; eligibility and exclusion criteria; breast cancer incidence rates; risk factors included in the models; and performance measures of the models. A second reviewer confirmed the accuracy of abstracted information.

Quality Assessment

We used predefined criteria developed by the U.S. Preventive Services Task Force to assess the quality of studies of benefits and harms of medications (Appendix C-1).[13] To determine quality of risk assessment instruments, we adapted the U.S. Preventive Services Task Force criteria for diagnostic accuracy studies (Appendix C-1).[13] We did not evaluate descriptive studies for quality because specific criteria are not available for these study designs. Two investigators independently rated the quality of each eligible study (good, fair, poor) and final ratings were determined by consensus.

Applicability

We assessed applicability of studies by following the population, intervention, comparator, outcomes, timing of outcomes measurement, and setting (PICOTS) format (Appendix C-1).[14] When possible, we highlighted *effectiveness* studies conducted in settings relevant to clinical practice, with subjects selected with less stringent eligibility criteria, assess health outcomes, and have longer follow-up periods than most *efficacy* studies. The results of effectiveness studies are more applicable to the spectrum of patients that will use a medication, have a test, or undergo a procedure than results from highly selected populations in efficacy studies. Two investigators independently rated the quality of each eligible study (good, fair, poor) and final ratings were determined by consensus.

Rating the Body of Evidence

We assessed the overall strength of the body of evidence through group consensus using the EPC GRADE (Grading of Recommendations Assessment, Development and Evaluation)

approach.[15] This approach uses a two step process for each key outcome. First, we assessed risk of bias, consistency of effect, directness, and precision for each outcome. We also determined the magnitude of effect for key outcomes using results of meta-analyses of trials as described below. Additional optional domains in the EPC GRADE table were not included in the table because they are not relevant to this review. Definitions and criteria for scoring these domains are described in Appendixes C-2, C-3, and C-4. Second, we determined overall grades based on qualitative combinations of the ratings for each domain. The EPC GRADE classifications for the overall strength of the body of evidence are: high, moderate, low, and insufficient (Appendix C-3). A grade of high indicates high confidence that the evidence reflects the true effect and further research is very unlikely to change our confidence in the estimate of effect. A grade of moderate indicates moderate confidence that the evidence reflects the true effect and that further research may change our confidence in the estimate of effect and may change the estimate. A grade of low indicates low confidence that the evidence reflects the true effect and that further research is likely to change our confidence in the estimate of effect and is likely to change the estimate. A grade of insufficient is given when the evidence either is unavailable or does not permit estimation of an effect.

Data Synthesis

Statistical Analysis

We combined results of eligible placebo-controlled trials in several meta-analyses to obtain more precise estimates of major health outcomes for the target population (key questions 1 and 2), and explore whether the combined estimates differ among subpopulations (key question 3). To determine the appropriateness of meta-analysis, we considered clinical and methodological diversity and assessed statistical heterogeneity.

We abstracted or calculated estimates of risk ratios (rate ratio, hazard ratio, or relative risk) and their standard errors from each trial and used them as the effect measures. For each outcome, we adopted the following steps to obtain the risk ratio and to account for the varying follow-up periods of the trials:

1) If a study reported a rate ratio based on a Poisson model, where women-years of follow-up was incorporated in the estimates, or a hazard ratio from a Cox regression model, we used the reported estimate.
2) If not, but the study reported the number of events and women-years of follow-up, or women-years of follow-up could be calculated from reported data, we calculated the rate ratio based on a Poisson distribution using the number of events and women-years of follow-up.
3) If both 1) and 2) were not possible, we used the reported or calculated relative risk, which does not take into account the women-years of follow-up. However, the estimate of relative risk would be expected to be very close to the estimate of rate ratio since the mean or median follow-up time was similar between the treatment and control arms in the trials.

We assessed the presence of statistical heterogeneity among the studies using standard χ^2 tests, and the magnitude of heterogeneity using the I^2 statistic.[16] We used a random effects model to account for variation among studies.[17] When there is no variation among studies, the random effects model yields the same results as a fixed effects model. For all meta-analyses, we

combined results separately for tamoxifen and raloxifene and provided 95% confidence intervals. We used STATA® 9.1 software for all these analyses (StataCorp, College Station, TX, 2006).

To explore whether combined estimates differ among subpopulations for key question 3, we performed subgroup analysis by age (≤50 yrs vs. > 50 yrs), family history of breast cancer (yes vs. no), use of menopausal hormone therapy (yes vs. no), menopausal status (pre vs. post), and body mass index (BMI) (≤25 vs. >25), when at least two studies reported results. We also performed subgroup analysis for tamoxifen trials stratified by active vs. post treatment periods when studies reported these data.

We also conducted an indirect comparison to compare the major benefits from trials of raloxifene with the one trial of tibolone using meta-regression. Since the raloxifene and tibolone trials recruited much older populations than the tamoxifen trials, we did not conduct indirect comparisons between the tamoxifen trials and raloxifene/tibolone trials.

Event Rates

To facilitate the evaluation of benefits and harms across trials, we abstracted or calculated event rates per 1000 women years for both treatment and placebo groups using steps similar to those obtaining risk ratios. When the event rates were not reported or calculable, we indicated them as such. To obtain the combined event rates, we conducted a meta-analysis of the placebo event rates by using a random effects Poisson model and raw data of number of events and women years of follow-up. We used PROC NLMIXED, SAS 9, 1.3. software for this analysis (SAS Institute Inc., Cary, NC, 2008).

Number Needed To Treat/Harm

To help interpret the clinical impact of the medications, we calculated the number of women needed to treat (NNT) to cause an outcome if each woman were to take the medication for 5 years. These numbers and the corresponding 95% confidence intervals were estimated using the combined risk ratios from the meta-analyses and the combined event rates from the placebo groups of included trials. To obtain the combined event rates, we conducted a meta-analysis of the placebo event rates as described above. We calculated the 95% confidence intervals for NNT by using a simulation method. We assumed that both logs of risk ratios and event rates have normal distributions, and we drew 10,000 random samples from them. The number needed to treat/harm and the number of events prevented/caused were then calculated from each sample, and the 95% confidence intervals were obtained by computing the 2.5% and 97.5% quantiles of the full sample.

Peer Review and Public Commentary

A draft of the report was sent to peer reviewers, anonymous reviewers identified by the United States Preventive Services Task Force, AHRQ representatives and the Scientific Resource Center. The draft report was also posted on the AHRQ Effective Health Care for a public comment period. Changes to the report were made based on comments received from peer and public reviewers. A summary of responses to comments will be publically available on the Effective Health Care website.

Results

From electronic database searches and the scientific information packet, we identified 4,842 abstracts (Figure 2). For key question 1 and the benefits portion of key question 3, we reviewed 72 full-text papers and included 13 in our results. For key question 2 and the harms portion of key question 3, we reviewed 280 full-text papers and included 70. For key question 4, we reviewed 120 full-text papers and included 24. For key question 5, we reviewed 112 full-text papers and included 16. Excluded studies are cataloged in Appendix B.

Description of Primary Prevention Trials

Eight large randomized controlled trials of tamoxifen, raloxifene, and tibolone that enrolled women without breast cancer and reported breast cancer outcomes provide the main results for this comparative effectiveness review. Additional studies are described in subsequent sections. The primary prevention trials include one head-to-head trial of tamoxifen and raloxifene, the Study of Tamoxifen and Raloxifene (STAR);[12,18] four placebo-controlled trials of tamoxifen, including the International Breast Cancer Intervention Study (IBIS-I),[19,20] National Surgical Adjuvant Breast and Bowel Project (NSABP P-1),[21-24] Royal Marsden Hospital Trial,[25,26] and the Italian Tamoxifen Prevention Study;[27-30] two placebo-controlled trials of raloxifene, the Multiple Outcomes of Raloxifene Evaluation (MORE) with long-term follow-up in the Continuing Outcomes Relevant to Evista (CORE) study,[31-45] and the Raloxifene Use for the Heart (RUTH) trial;[46,47] and one placebo-controlled trial of tibolone, the Long-Term Intervention on Fractures with Tibolone (LIFT).[10] Details of individual trials are provided in Tables 2 and 3.

All of the primary prevention trials met criteria for fair or good quality for major outcomes (Appendix C-5). We considered the most important methodological limitation of the trials to be the inclusion of women using estrogen in the Italian (14% of women), Royal Marsden (15% to 27%), and IBIS (40%) tamoxifen trials. Estrogen use could modify or confound breast cancer risk. Estrogen could influence other outcomes, such as thromboembolic events, especially in trials where estrogen use varied between treatment and placebo groups.

Trials met criteria for good applicability, except for the Italian trial that exclusively enrolled women who had undergone prior hysterectomy[28] (Appendix C-5). These women represent a subgroup of the target population. Those with oophorectomies may be at lower than average risk for breast cancer. Although the other trials used differing inclusion criteria, they selected women who would be considered candidates for risk reduction medications in the target population. For each trial, interventions, comparators, outcomes, and timing of outcome measures were appropriate. All trials were multi-center and relevant to primary care.

The primary prevention trials are large, ranging from the Royal Marsden trial[25] enrolling 2,471 women to the STAR trial enrolling 19,747.[12] Subjects were recruited from clinics and communities located in many countries, with North America, Europe, and the United Kingdom most represented. The majority of subjects are white and none of the trials provide outcomes specific to racial or ethnic groups. Subjects range in age from 30s to 80s at baseline.

The tamoxifen trials, including STAR, were designed to determine breast cancer incidence as the primary outcome.[12,19,20,23-30] As such, inclusion criteria considered breast cancer risk in all of these trials except the Italian Tamoxifen Prevention Study.[28] Two trials, STAR and NSABP P-1, utilized the modified Gail model[48,49] to select subjects. In STAR, women were eligible for the trial if they were postmenopausal and had a Gail model 5-year predicted breast

cancer risk of ≥1.66%.[12] The NSABP P-1 trial used this same threshold as well as additional criteria, such as age ≥60 or a history of lobular carcinoma *in situ*.[24] Most women age ≥60 years have a Gail model risk ≥1.66% without additional risk factors because age is an important predictor in the model. The IBIS and Royal Marsden trials defined eligibility criteria based on numbers of relatives affected with breast cancer as well as personal history of prior benign breast biopsies.[19,25] Inclusion criteria are further described in Table 2.

Breast cancer incidence was one of two primary outcomes in RUTH, and was a secondary outcome in MORE and LIFT. The MORE and LIFT trials enrolled women with osteoporosis in order to determine the efficacy of raloxifene or tibolone in preventing fractures.[10,38] Eligibility criteria for both trials included bone mineral density (BMD) T-score ≤-2.5 at the femoral neck or lumbar spine, or low BMD with pre-existing vertebral fractures at baseline. The RUTH trial was designed to determine the efficacy of raloxifene in preventing coronary events and enrolled women with coronary heart disease or multiple risk factors for heart disease.[46] Subjects were required to have a cardiovascular risk score of 4 or more according to a point system that assigned values for specific conditions (Table 2).

Differences in trial designs lead to the enrollment of dissimilar groups of women into trials. The mean age at entry of subjects ranged from 47[25] to 51 years[50] in the tamoxifen trials, and from 67[34] to 68 years[10,46] in the raloxifene and tibolone trials. Risks for most outcomes measured in these trials increase with age, including risks for adverse events such as thromboembolic events and strokes. The 15 to 20-year age difference between subjects in different trials would be expected to influence results and limit comparisons between medications. Differences in other subject characteristics that have known associations with breast cancer could also influence outcomes, such as prior oophorectomy (reduces risk), estrogen and progestin use (increases risk), family history of breast cancer (increases risk), and osteoporosis (may reduce risk). Although the head-to-head design of the STAR trial allows direct comparisons between tamoxifen and raloxifene, there are no head-to-head comparison trials for tibolone.

Trials also varied by treatment and follow-up times. These variations could influence results because individuals with short exposures may not attain the optimal benefits or experience the adverse effects that individuals with longer exposures would. Also, short follow-up times may not allow conditions with slower progression, such as breast cancer, to be detected during the course of the trial. Median treatment times were not provided for every placebo-controlled trial of tamoxifen, but available information indicates treatment times of approximately 4 years.[24,29] Three of the four tamoxifen trials provided explicit median follow-up times ranging from 7 years in NSABP P-1[23] to 13 years in Royal Marsden.[26] The Royal Marsden[26] and IBIS[20] trials provided some results by active vs. post treatment periods, while other trials did not. Results of the MORE trial were reported after 3 and 4 years of treatment.[31-39,41,44] The CORE study is a continuation of MORE that follows a subset of MORE subjects in order to further examine raloxifene's effect on breast cancer incidence. Although subjects continued their randomized assignment to raloxifene or placebo, all had a gap in their use. Median time between participation in MORE and CORE was 10.6 months (2.6-62 months).[51] Results of CORE are reported for 4-year and combined 8-year outcomes (MORE + CORE).[42,43,45] The RUTH, LIFT, and STAR trials have only recently been published and do not provide long-term outcomes. Median exposures to medications are 2.8 years in LIFT,[10] 3.1 to 3.2 years in STAR,[52] and 5.1 years in RUTH.[46]

Although most trials reported similar main outcomes, the ascertainment of outcomes varied by trial (Table 3). The diagnostic criteria for several outcomes were not well described in the trials and it is likely that differences in results between trials for some of these outcomes may be due, at least in part, to how the outcomes were determined and measured. All of the primary prevention trials reported incidence of invasive breast cancer, and most reported results for estrogen receptor positive,[20,23,26,29,46,51] negative,[20,23,26,29,46,51] and noninvasive breast cancer separately.[20,23,26,29,46,51] All-cause mortality was provided in all of the primary prevention trials, and breast cancer specific mortality in five.[20,23,26,29,53] Fracture outcomes were more comprehensively evaluated in the MORE and LIFT trials.[10,35,38,45] Both trials evaluated fractures at multiple anatomic sites, such as the hip and wrist specifically, and detected rigorously defined radiographic vertebral fractures. The NSABP P-1, RUTH, and STAR trials included clinical vertebral fractures;[12,23,46] these are identified by physical findings or symptoms. Other trials included only larger categories of fractures such as all types or osteoporotic types (hip, vertebral, wrist).[20,26]

All primary prevention trials reported thromboembolic events, and some provided specific results for deep vein thrombosis,[24,27,39,46] pulmonary embolus,[24,27,39,46] and superficial phlebitis.[20,27] Coronary heart events were described in all trials and included myocardial infarction, angina, acute ischemic syndrome, and other events. However, specific outcomes included in this broad category varied and were often not well specified. The RUTH trial, designed to measure coronary outcomes primarily, provided the most comprehensive measures.[46] Stroke was measured in all trials and transient ischemic attack in five.[10,12,20,24,29] Endometrial cancer, hysterectomy, endometrial hyperplasia, uterine fluid, and vaginal bleeding were determined in various ways in most trials. Six trials reported cataracts.[12,20,24,26,39,46] Descriptions of other outcomes, such as vasomotor symptoms, edema, pain, and quality of life measures, for example, vary by trial. Additional details of ascertainment of adverse outcomes are described for key question 2.

Key Question 1. In adult women without pre-existing breast cancer, what is the comparative effectiveness of selective estrogen receptor modulators (SERMs) tamoxifen citrate and raloxifene, and the selective tissue estrogenic activity regulator (STEAR) tibolone, when used for the primary prevention of breast cancer on improving short-term and long-term outcomes.

Key Points

- Eight large randomized controlled trials provide data on breast cancer risk reduction in women without pre-existing breast cancer. These include one good-quality head-to-head trial of tamoxifen and raloxifene and seven fair and good quality placebo-controlled trials (four tamoxifen, two raloxifene, and one tibolone). Results of placebo-controlled trials cannot be directly compared between types of medications because of important differences between study subjects.
- Tamoxifen (RR 0.70; 0.59, 0.82; 4 trials), raloxifene (RR 0.44; 0.27, 0.71; 2 trials), and tibolone (RR 0.32; 0.13, 0.80; 1 trial) reduce the incidence of invasive breast cancer in

midlife and older women by approximately 30% to 68%; tamoxifen and raloxifene had similar effects in the STAR head-to-head trial.

- Reduction of invasive breast cancer continued at least 3 to 5 years after discontinuation of tamoxifen in the two trials providing post treatment follow-up data.
- Tamoxifen (RR 0.58; 0.42, 0.79; 4 trials) and raloxifene (RR 0.33; 0.18, 0.61; 2 trials) reduce estrogen receptor positive invasive breast cancer, but not estrogen receptor negative invasive breast cancer, in placebo-controlled trials, and had similar effects in the STAR head-to-head trial.
- Tamoxifen and raloxifene do not significantly reduce noninvasive breast cancer, including ductal carcinoma *in situ* (DCIS) in meta-analysis of four placebo-controlled trials, although noninvasive breast cancer was significantly reduced in the NSABP P-1 tamoxifen trial (RR 0.63; 0.45, 0.89). The STAR head-to-head trial indicated no statistically significant differences between raloxifene compared to tamoxifen (RR 1.40; 0.98, 2.00).
- All-cause mortality is similar for women using raloxifene compared to tamoxifen; or tamoxifen, raloxifene, or tibolone compared to placebo, although follow-up times in most trials were short. Tamoxifen does not reduce breast cancer mortality compared to placebo.
- Tamoxifen and raloxifene had similar effects on fractures at multiple sites in the STAR head-to-head trial. In placebo-controlled trials, raloxifene (RR 0.61; 0.54, 0.69; 2 trials) and tibolone (RR 0.55; 0.41. 0.74; 1 trial) reduce vertebral fractures, tamoxifen (RR 0.66; 0.45, 0.98; 1 trial) and tibolone (RR 0.74; 0.58, 0.93; 1 trial) reduce nonvertebral fractures, and tibolone reduces wrist (RR 0.54; 0.35, 0.82; 1 trial) but not hip fractures.

Detailed Analysis

The eight randomized controlled trials reported in 11 publications described above and in Tables 2 and 3 provide data for key question 1. Results are summarized in Table 4.

Invasive breast cancer

Tamoxifen vs. raloxifene. Raloxifene and tamoxifen had similar effects on invasive breast cancer in the STAR head-to-head trial (RR for raloxifene vs. tamoxifen 1.02; 0.82, 1.28),[12] and there were also no differences for estrogen receptor positive and negative subtypes.

Tamoxifen vs. placebo. Tamoxifen reduced invasive breast cancer in all four prevention trials using long-term follow-up data.[20,23,26,29] Reductions ranged from 20% to 43% with the biggest effect from the largest trial, the NSABP P-1 trial (RR 0.57; 0.46, 0.70).[23] Combining results in meta-analysis indicates a summary RR of 0.72 (0.61, 0.86; 4 trials) for all breast cancer (Figure 3) and 0.70 (0.59, 0.82; 4 trials) for invasive breast cancer specifically (Figure 4). Tamoxifen reduced risks for estrogen receptor positive (RR 0.58; 0.42, 0.79; 4 trials), but not estrogen receptor negative breast cancer (RR 1.19; 0.92, 1.55; 4 trials) (Figure 5).[20,23,26,29]

The IBIS[20] and Royal Marsden[26] trials provided results for invasive and estrogen receptor positive breast cancer for both active treatment (mean duration 5 and 8 years, respectively) and post treatment periods (mean duration 3 and 5.2 years, respectively). These results indicate continued risk reduction after discontinuation of tamoxifen, providing point estimates of even

larger reductions in breast cancer during the post treatment period (Figure 6). However, differences between periods were not statistically significant by subgroup comparison analysis.

Raloxifene vs. placebo. Raloxifene reduced invasive breast cancer by 44% and 66% in the MORE[51] and RUTH[46] trials. Combining results in meta-analysis indicated a summary RR of 0.53 (0.34, 0.84; 2 trials) for all breast cancer (Figure 3) and 0.44 (0.27, 0.71; 2 trials) for invasive breast cancer specifically (Figure 4). Raloxifene reduced risk for estrogen receptor positive (RR 0.33; 0.18, 0.61; 2 trials), but not estrogen receptor negative breast cancer (RR 1.25; 0.67, 2.31; 2 trials) (Figure 5).

Tibolone vs. placebo. Tibolone reduced invasive cancer by 68% in the LIFT trial (RR 0.32; 0.13, 0.80; 1 trial).[10] The LIFT trial did not report specific results for estrogen receptor types or noninvasive breast cancer.

Indirect comparisons. Where we lacked data from direct head-to-head trials, we used meta-regression to compare differences in risk ratios derived from placebo-controlled trials. As described above, invasive cancer outcomes for raloxifene vs. tamoxifen were not significantly different when directly compared in the STAR trial.[12] Indirect comparison of raloxifene vs. tibolone also indicated no significant differences (raloxifene vs. tibolone, ratio of risk ratios [RRR] 1.37; 0.49, 3.84). Tibolone and tamoxifen were not compared indirectly because of important differences in patient populations.

Noninvasive breast cancer including ductal carcinoma *in situ* (DCIS)

Tamoxifen vs. raloxifene. STAR reported nonsignificantly increased risks for noninvasive cancer (RR 1.40; 0.98, 2.00) and DCIS (RR 1.46; 0.90, 2.41) among women using raloxifene vs. tamoxifen.[12]

Tamoxifen vs. placebo. All four tamoxifen trials reported noninvasive cancer outcomes, although specific diagnoses varied between trials. Risks were reduced in the NSABP P-1[23] and IBIS[20] trials, and increased in the Royal Marsden[26] and Italian[29] trials, although results were significant only in the NSABP P-1 trial (RR 0.63; 0.45, 0.89). When combined in meta-analysis, the risk of noninvasive breast cancer was not significantly reduced (RR 0.85; 0.54, 1.35; 4 trials) (Figure 7).

Raloxifene vs. placebo. Both the MORE[51] and RUTH[46] trials indicated increased risks for noninvasive breast cancer, although results were not statistically significant. Combining estimates in meta-analysis indicated a nonsignificant elevation in risk (RR 1.47; 0.75, 2.91; 2 trials) (Figure 7). For DCIS specifically, MORE reported 9 cases for raloxifene and 5 for placebo.

Tibolone vs. placebo. One case of DCIS was noted in the tibolone group and one in the placebo group.[10]

Breast cancer mortality

Tamoxifen vs. raloxifene. Not reported.

Tamoxifen vs. placebo. All four tamoxifen trials reported breast cancer specific death rates using long-term follow-up data.[20,23,26,29] None of these results were significantly different for tamoxifen vs. placebo (RR 1.07; 0.66, 1.74; 4 trials) (Figure 8).

Raloxifene vs. placebo. Very few breast cancer deaths occurred in the MORE/CORE trial and no relative risks were reported.[51]

Tibolone vs. placebo. Not reported.

All-cause mortality

Tamoxifen vs. raloxifene. Total death rates among women in the STAR trial were similar for women treated with tamoxifen or raloxifene (RR 0.94; 0.71, 1.26).[12]

Tamoxifen vs. placebo. All four tamoxifen trials reported all-cause death rates using long-term follow-up data, and none were significantly different for tamoxifen vs. placebo (RR 1.07; 0.90, 1.27; 4 trials) (Figure 8).[20,23,26,29]

Raloxifene vs. placebo. The RUTH and MORE trials reported all-cause death rates that were nonsignificantly reduced compared to placebo (RR 0.91; 0.81, 1.02; 2 trials) (Figure 8).[46,51]

Tibolone vs. placebo. The LIFT trial reported 26 deaths among women using tibolone and 28 among those using placebo (p=0.89).[10]

Osteoporotic fractures

Tamoxifen vs. raloxifene. Results of the STAR trial indicated no differences between tamoxifen and raloxifene for clinical vertebral, hip, wrist, or total fractures, although all rates were slightly less for raloxifene.[12]

Tamoxifen vs. placebo. The NSABP P-1,[23] IBIS,[20] and Royal Marsden[26] trials reported fractures as secondary outcomes. The tamoxifen trials enrolled subjects 15 to 20 years younger and with much lower fracture rates than subjects in trials of raloxifene.

In the NSABP P-1 trial, tamoxifen reduced risk of combined clinical vertebral, wrist, and hip fractures with tamoxifen compared to placebo (RR 0.68; 0.51, 0.92).[23] Point estimates of risk ratios were also reduced for these fractures in the IBIS[20] and Royal Marsden trials,[26] however, results were not statistically significant. Meta-analysis of trials indicates nonsignificant reductions in total (RR 0.84; 0.67, 1.05; 2 trials) and osteoporotic site fractures (i.e., hip, spine, wrist) (0.81; 0.55, 1.18; 2 trials) (Figure 9). Clinical vertebral fractures specifically were not significantly reduced in the NSABP P-1 trial (RR 0.75; 0.48, 1.15) (Figure 10), although hip and wrist fractures combined were (RR 0.66; 0.45, 0.98) (Figure 11).[23]

Raloxifene vs. placebo. The MORE trial recruited women with low BMD (T-score ≤-2.5) and/or prior vertebral fractures.[35,45] At baseline, 37% of women had prior vertebral fractures. In MORE, raloxifene reduced vertebral fractures (RR 0.60; 0.53, 0.69),[35] but not nonvertebral or hip fractures compared to placebo.[45] Results were similar for women with and without prior vertebral fractures and for women using two different doses of raloxifene (60 or 120 mg/day). The RUTH trial measured fractures as secondary outcomes.[46,54] RUTH reported reduced clinical vertebral fractures (RR 0.65; 0.47, 0.89), but not nonvertebral fractures (RR 0.96; 0.84, 1.10) among raloxifene users compared to placebo, consistent with results of MORE.[46] Combining the results of MORE and RUTH in a meta-analysis indicates a vertebral fracture RR 0.61 (0.54, 0.69) (Figure 10) and a nonvertebral fracture RR 0.97 (0.87, 1.09) (Figure 11).

Tibolone vs. placebo. The LIFT trial[10] recruited women with low BMD (T-score ≤-2.5) and/or prior vertebral fractures, similar to the MORE trial. At baseline, 22% of women had prior nonvertebral fractures and 26% had prior vertebral fractures. Tibolone reduced vertebral (RR 0.55; 0.41, 0.74), nonvertebral (RR 0.74; 0.58, 0.93), and wrist (RR 0.54; 0.35, 0.82), but not hip fractures (RR 0.72; 0.32, 1.63). Tibolone appeared to reduce more fractures for women with prior vertebral fractures (vertebral RR 0.39; 0.24, 0.63; nonvertebral 0.53; 0.35, 0.81) than for women without prior vertebral fractures (vertebral RR 0.69; 0.48, 1.00; nonvertebral RR 0.86; 0.65, 1.14).

Key Question 2. What is the evidence for harms of tamoxifen citrate, raloxifene, and tibolone when used for primary prevention of breast cancer?

Key Points

- In addition to the eight large randomized controlled trials described in key question 1, harms data were provided by 12 placebo-controlled trials and one observational study of raloxifene, and seven placebo-controlled trials and one observational study of tibolone.
- Raloxifene caused fewer thromboembolic events (RR 0.70; 0.54, 0.91) than tamoxifen in the STAR head-to-head trial. Tamoxifen (RR 1.93; 1.41, 2.64; 4 trials) and raloxifene (RR 1.60; 1.15, 2.23; 2 trials) cause more thromboembolic events than placebo. Risk returned to normal after discontinuation of tamoxifen in the 2 trials providing post treatment data. Tibolone does not increase risk for thromboembolic events, although data are limited.
- Tamoxifen, raloxifene, and tibolone do not increase risk for coronary heart disease events, although data for tibolone are limited.
- Tibolone causes more strokes than placebo (RR 2.19; 1.14, 4.23); tamoxifen and raloxifene do not increase risk for stroke.
- In the STAR head-to-head trial, raloxifene caused fewer cases of endometrial hyperplasia (RR 0.16; 0.09, 0.29) and was associated with fewer hysterectomies (RR 0.44; 0.35, 0.56) than tamoxifen, but differences for endometrial cancer were not statistically significant (RR 0.62; 0.35, 1.08).
- Tamoxifen causes more cases of endometrial cancer than placebo (RR 2.13; 1.36, 3.32; 3 trials); raloxifene does not increase risk for endometrial cancer or uterine bleeding, and

tibolone does not increase risk for endometrial cancer in clinical trials, but was associated with more cases of endometrial cancer in a large cohort study (RR 1.79; 1.43, 2.25).

- Raloxifene caused fewer cataracts (RR 0.79; 0.68, 0.92) and cataract surgeries (RR 0.82; 0.68, 0.99) than tamoxifen in the STAR head-to-head trial; tamoxifen was associated with more cataract surgeries than placebo in the NSABP P-1 trial (RR 1.57; 1.16, 2.14); raloxifene does not increase risk for cataracts or cataract surgery.
- In head-to-head comparisons, women using raloxifene reported more musculoskeletal problems, dyspareunia, and weight gain, while those using tamoxifen had more gynecological problems, vasomotor symptoms, leg cramps, and bladder control symptoms.
- Most common side effects for tamoxifen are hot flashes and other vasomotor symptoms, vaginal discharge, and other vaginal symptoms such as itching or dryness; for raloxifene, vasomotor symptoms and leg cramps; and for tibolone, vaginal bleeding and reduced number and severity of hot flashes.

Detailed Analysis

A total of 29 studies met inclusion criteria for key question 2. Details are provided in Tables 2, 3, 5 and 6 and Appendixes D-1, D-2, and D-3.

Description of tamoxifen studies

For tamoxifen, information on adverse effects was confined to the four large placebo controlled primary prevention trials,[19-27,29,30,50,55-69] and the STAR head-to-head trial.[12,18,70,71] We identified no other randomized controlled trials or observational studies that evaluated adverse effects in women without breast cancer. We considered all adverse outcomes at all reported follow-up times to capture potential short and long-term adverse effects. However, because the NSABP P-1 trial was unblinded after reporting initial results in 1998, we focused on data from the earlier 1998 publication,[24] and then compared these results with data from the subsequent 2005 publication.[23]

Trials reported adverse effects in different ways depending on the outcome. Most evaluated adverse effects at clinic visits using either self or staff administered questionnaires and checklists. The NSABP P-1 trial documented them by using a global index modeled after the Women's Health Initiative.[23,24,55,57,59,64] Patients were administered a baseline Health Related Quality of Life examination that was repeated at 36 months. Follow-up visits occurred at 3 and 6 months, and then every 6 months thereafter.[55] Endometrial cancer and thromboembolic events were considered secondary end points in this trial. Gynecologic symptoms of hot flashes, vaginal discharge, vaginal dryness, and abnormal vaginal bleeding were monitored, and clinical sites reported additional uterine and ovarian disorders and gynecologic procedures.[57] Medical records for subjects with suspected cardiovascular disease events were collected by the clinical sites and adjudicated by investigators blinded to treatment assignment.[64] Although trial results were initially reported in 1998 and the study was unblinded at that time, most subjects were followed 7 years.[23] During follow-up, nearly 1/3 of women in the placebo group elected to either enter the STAR trial or begin a SERM for breast cancer prevention.[23] Long-term results of the NSABP P-1 trial are limited by fewer years of follow-up in the placebo group, substantial contamination, and unblinded ascertainment of outcomes.

In the IBIS trial, adverse effects were assessed differently during the active and follow-up phases of the study in Europe and the U.K.; in Australia and New Zealand, the same procedures

were used during the entire study.[19,20] During active treatment and post treatment follow-up phases, a checklist of predefined adverse effects with a free text field was used. Predefined adverse outcomes included myocardial infarction, cardiovascular disease events, thromboembolic events, osteoporotic fractures, any non-breast cancer, nausea, vomiting, hot flushes, headaches, vaginal discharge, vaginal dryness, and vaginal bleeding. During the active treatment phase, these questions were asked directly to subjects. During the follow-up phase, a less detailed version of the checklist was mailed to subjects. For postal replies, adverse outcomes were confirmed by medical record review. Approximately 85% of women returned at least one questionnaire during follow-up.

In the Royal Marsden trial, follow-up visits occurred every 6 months during the course of the trial.[25,26] Acute toxicity and other conditions were assessed at each visit and mammograms were performed annually. Further details of the follow-up procedures for adverse effects were not reported.

Subjects underwent a physical examination every 6 months, and blood testing and mammography every 12 months in the Italian trial.[27,29,50,56] After completion of treatment, or in the case of dropouts, women were followed on an annual basis. Information about major endpoints, such as death, serious adverse events, or cancer, was collected continuously and submitted to the data center. Secondary endpoints included cardiovascular disease, psychological measures, and cognitive function. Surveillance for onset of acute or chronic liver injury based on blood levels of transaminases was also included. Only adverse events that occurred during study treatment were reported.

Subjects in the STAR trial were followed every 6 months for 5 years and annually thereafter.[12] Gynecologic examinations, complete blood counts, and routine serum chemistry tests were obtained annually. Information about the occurrence of all protocol-defined endpoints (endometrial cancer, cardiovascular disease, stroke, pulmonary embolism, deep vein thrombosis, transient ischemic attack, osteoporotic fracture, cataracts, death, quality of life, other cancers) was ascertained at each follow-up visit and verified by reviewing relevant records. Self reported symptoms were collected at each contact. In-depth quality of life assessments were also obtained.[18]

Description of raloxifene studies

For raloxifene, we obtained adverse effect data from the two large placebo-controlled prevention trials, MORE and RUTH,[31-35,37-41,46,47,72] the STAR head-to-head trial,[12,18,70,71] 12 smaller trials evaluating either bone density, biochemical profiles, or fractures (Appendixes D-1 and D-2),[73-85] and one observational study.[86] No other observational studies met inclusion criteria. In general, the smaller trials of raloxifene and the observational study contribute little to the evaluation of harms because they involve so few women relative to the large primary prevention trials.

Details of the ascertainment of adverse outcomes were described in the MORE and RUTH trials. Subjects were followed every 6 months in the MORE trial and were queried about potential adverse effects at every visit.[32-34,39] Fasting plasma glucose levels were evaluated annually. Endometrial changes were monitored with transvaginal ultrasound at 17 clinic centers; some centers only performed transvaginal ultrasound on a subset of women. All cases of endometrial cancer were confirmed by a panel blinded to treatment assignment. Medical records and reports were reviewed for subjects reporting possible thromboembolic events by three physician adjudicators blinded to treatment assignment. In RUTH, subjects were followed every

6 months by either a visit or telephone call, and adverse events were ascertained at each evaluation through unsolicited reporting by subjects.[46] Electrocardiograms were performed at baseline, years 2 and 4, and the final visit. Serum lipids were measured at baseline, years 1 and 5, and the final visit. Committees of experts blinded to treatment assignment adjudicated coronary events, breast cancer, stroke, thromboembolism, and death outcomes.

The 12 smaller trials ranged in size from 129 to 1,145 postmenopausal women. Women had osteoporosis in 5 trials.[74,79-81,83] The dose of raloxifene ranged from 30 to 150 mg per day, although all trials evaluated a 60 mg per day dose. The duration of the studies ranged from 6 months to 5 years. Several of the smaller trials adequately collected and reported data for selected adverse outcomes, but reported others inadequately or not at all (Appendix D-1), and none evaluated more than 1 to 3 adverse outcomes. Of the 12 smaller raloxifene trials,[73,74,76-85] only 6 reported thromboembolic events[77-79,81,82,84] and none reported cardiovascular events. Four trials evaluated uterine outcomes,[73,74,79,80] one urinary outcomes,[76] and one cognitive function.[83] The most commonly reported adverse events were hot flashes and vasomotor symptoms reported in eight trials.[74,77,78,80-84] The one included observational study evaluated the effect of raloxifene on vaginal bleeding and endometrial thickness.[86] No other observational studies met inclusion criteria. In general, the smaller trials of raloxifene and the observational study contribute little to the evaluation of harms because they involve so few women relative to the large primary prevention trials.

Description of tibolone studies

The LIFT trial,[10,87] seven additional randomized placebo-controlled trials (Tables 5 and 6 and Appendixes D1 and D-2),[88-96] and one large cohort study, the Million Women Study (Appendixes C-5, D-1 and D-2),[97,98] met inclusion criteria. Trials ranged in size from 106 to 4,538 subjects, daily tibolone treatment doses ranged from 0.3 to 5 mg, and duration of treatment from 3 months to 3 years. In the large Million Women Study, the dose and duration of tibolone use varied, and the average lengths of follow-up were 2.6 years for incidence of outcomes, and 4.1 years for mortality.[98] Primary outcomes in these studies included fracture,[10] cardiovascular disease,[10] breast cancer,[10,98] endometrial cancer,[97] menopausal symptoms,[91,93,94] breast density,[95] depression,[96] bone density,[88,92] carotid intima-media thickness,[88] and lipids,[94,96] although all reported additional secondary outcomes and adverse effects.

Other trials of tibolone were excluded because they enrolled less than 100 subjects, lacked a placebo or nonuse comparison group, or included subjects with a history of breast cancer (Appendix B). For example, the Tibolone Histology of the Endometrium and Breast Endpoints Study (THEBES)[99] did not contain a placebo group, and the Livial Intervention Following Breast Cancer; Efficacy, Recurrence and Tolerability Endpoints (LIBERATE) trial[100] enrolled women with a history of breast cancer. Other observational studies were reviewed and excluded[101-104] due to the lack of non-use comparison groups, small numbers of tibolone users within a larger pool of menopausal hormone therapy users, and/or lack of reported adverse effects.

Overall, the LIFT trial was well powered for several adverse event outcomes, providing data on cancer, stroke, gastrointestinal, and gynecological outcomes for older postmenopausal women with osteoporosis.[10,87] Although most of the remaining tibolone trials reported some data on various adverse events, most were underpowered to determine statistically significant differences for major outcomes such as death, stroke, and cancer. Other less serious adverse effects were reported with varying degrees of detail.

The large 3-year Osteoporosis Prevention and Arterial effects of tiboLone (OPAL) trial compared tibolone to other types of menopausal hormone therapy or placebo in Europe and the U.S.[89,90] A total of 866 predominantly Caucasian, healthy postmenopausal women ages 45 to 79 years were randomized to tibolone (2.5 mg/daily), conjugated equine estrogen (CEE) with medroxyprogesterone acetate (MPA) (0.625 mg/2.5mg respectively), or placebo for 36 months. Primary outcomes included bone mineral density (BMD) and carotid artery intima-medial thickness; adverse effects on the endometrium and vaginal bleeding were secondary outcomes. Approximately 30% of subjects were lost to follow-up compromising results.

A trial to determine bone density effects of tibolone enrolled 770 healthy postmenopausal women over age 45 years from over 47 sites in the U.S.[92] Subjects were randomized to either placebo or one of four daily doses of tibolone (from 0.3 to 2.5 mg) for 24 months. Adverse effects were well documented and included deep vein thrombosis, pulmonary embolus, vaginal symptoms, hot flashes, and others.[92] Loss to follow-up was 34 % in treatment and 29% in comparison groups.

A trial evaluating tibolone's effect on menopausal vasomotor symptoms enrolled 775 Scandinavian women experiencing severe hot flashes and sweating to either daily placebo or one of four doses of tibolone ranging from 0.625 to 5 mg for 3 years.[91] The placebo group had a higher drop-out rate compared to the tibolone group (20% vs. 11%, respectively) largely due to the lack of a therapeutic effect on vasomotor symptoms.

Four smaller trials conducted in various countries randomized between 106 to 396 healthy postmenopausal women to either 2.5 mg tibolone daily or placebo;[95,96] two trials included a 1.25 mg tibolone daily dose.[93,94] The U.S.[93] and Romanian[94] studies measured vasomotor and sexual function outcomes, the Turkish trial lipids and depression,[96] and the Swedish trial breast density.[95] Multiple adverse effects data were well documented in two trials,[93,94] while the other two provided limited data.[95,96] These trials had several methodological limitations, including no description of an intention-to-treat analysis,[94-96] differences between comparison groups for baseline patient characteristics,[93] and inadequate information on randomization procedures.[94] Applicability of the results was also limited because of the enrollment of small, selected populations including women seeking treatment for vasomotor symptoms.

The Million Women Study, a large, population-based prospective cohort study, compared breast and endometrial cancer outcomes of women using various hormone therapy regimens for symptomatic relief of menopausal symptoms with nonusers.[97,98] This study enrolled women age 50 to 64 years who were invited for routine breast cancer screening in the U.K. (N=1,084,110; mean age 56 years). Approximately 6% of the active hormone therapy users in this study were using tibolone. Data included self-reported information on sociodemographic and other personal factors and menopausal status, and cancer incidence and death rates from the National Health Service Central Registers.[98] This study is limited by the biases introduced by its observational design and subjects' self-selection of various regimens for symptomatic relief of menopausal symptoms. Some research indicates possible preferential prescribing of tibolone to women at higher risk for breast or endometrial cancer,[105] confounding associations with these outcomes.

Thromboembolic events

Tamoxifen vs. raloxifene. In the STAR trial, raloxifene caused fewer thromboembolic events compared to tamoxifen, including composite measures of thromboembolic events (RR 0.70;

0.54, 0.91), pulmonary embolism (RR 0.64; 0.41, 1.00), and deep vein thrombosis (RR 0.74; 0.53, 1.03).[12]

Tamoxifen vs. placebo. The four tamoxifen prevention trials identified thromboembolic complications as an adverse effect of active treatment, although the evaluation of this outcome varied by trial.[20,24,26,27] None of the trials indicated if thromboembolic events were adjudicated. All trials measured pulmonary embolus and deep venous thrombosis outcomes, the IBIS trial also measured superficial thrombophlebitis and retinal vein thrombosis,[20] and the Italian trial measured visceral, retinal, and superficial thrombophlebitis.[27] All of these trials excluded women with either a history of prior thromboembolic events or one within 10 years prior to study enrollment.

Active treatment with tamoxifen increased composite measures of thromboembolic events in all four prevention trials resulting in a summary risk ratio of 1.93 (1.41, 2.64; 4 trials) (Figure 12).[20,24,26,27] The IBIS[20] and Royal Marsden[26] trials provided results for both active and post treatment periods indicating no increased risk after discontinuation of active treatment (RR 1.02; 0.53, 2.97; 2 trials) (Figure 12).

Only the NSABP P-1[24] and Italian trials[27] evaluated outcomes by type of thromboembolic event. In the NSABP P-1 trial, tamoxifen increased risks for pulmonary embolism (RR 3.01; 1.15, 9.27); but risk was not statistically significantly increased for deep vein thrombosis (RR 1.60; 0. 91, 2.86).[24] In the Italian trial, risks were not elevated.[27] Summary risk ratios are 2.69 (1.12, 6.47; 2 trials) for pulmonary embolism and 1.45 (0.89, 2.37; 2 trials) for deep vein thrombosis (Figure 13).

Tamoxifen caused superficial thrombophlebitis in the Italian (RR 1.96; 1.10, 3.51)[27] and IBIS trials (RR 2.84; 1.07, 8.78),[20] with a summary risk ratio of 2.14 (1.29, 3.56; 2 trials) (Figure 13). The Italian trial also reported one retinal vein thrombosis in each arm of the trial and one visceral thrombosis in the placebo group.[27]

Raloxifene vs. placebo. Raloxifene increased thromboembolic events in both the MORE (RR 2.10; 1.20, 3.80)[39] and RUTH (RR 1.44; 1.06, 1.95)[46] trials, with similar event rates for women in control groups for both trials (3.50 and 3.67 per 1000 women years, respectively). Further analysis of the MORE trial by year of treatment indicated the highest risks during the first two years of therapy (RR ≥6 in years 1 and 2 vs. 0.9 in year 4).[39] Combining results of both trials in a meta-analysis results in a summary estimate of 1.60 (1.15, 2.23; 2 trials) (Figure 12). Both trials also reported nonstatistically significantly elevated risks for pulmonary embolus (combined RR 2.19; 0.97, 4.97; 2 trials) and deep vein thrombosis specifically (combined RR 1.91; 0.87, 4.23; 2 trials) (Figure 13). Although six other smaller trials reported information on thromboembolic events,[77-79,81,82,84] only two events occurred among women randomized to raloxifene and one among women randomized to placebo in these trials and they were not included in the meta-analyses.

Tibolone vs. placebo. Tibolone did not increase the risk of thromboembolic events,[10] deep vein thrombosis,[91,92] or pulmonary embolism[91,92] in the few trials reporting these outcomes. Rates of thromboembolism in the LIFT trial were 0.8 per 1000 women years in the tibolone group vs. 1.3 in the placebo group.[10]

Cardiovascular events

Tamoxifen vs. raloxifene. The STAR trial reported no differences between raloxifene and tamoxifen for a composite measure of ischemic coronary heart disease events (RR 1.10; 0.85, 1.43).[12] Specific events, such as myocardial infarction, severe angina, and acute ischemic syndrome, were also not significantly different between medications.[12] Stroke and transient ischemic attacks were also similar for raloxifene and tamoxifen in STAR (RR 0.96; 0.92, 1.32 and 1.21; 0.79, 1.88, respectively).[12]

Tamoxifen vs. placebo. Although the four prevention trials evaluated cardiovascular events,[20,24,26,27] definitions of outcomes, and the quality and detail of reporting varied across trials. Only the Italian trial indicated that they excluded women with a history of cardiovascular disease other than stable angina.[27]

The NSABP P-1 trial provided the most detailed information on cardiovascular outcomes, although it did not explicitly describe how these events were defined or adjudicated.[24] In this trial, rates of a composite measure of coronary heart disease, myocardial infarction, acute coronary syndrome, and severe angina were similar for tamoxifen and placebo.[24] The IBIS trial reported no increase in a composite measure of "all cardiac problems," including myocardial infarction, angina and other cardiac problems, as well as myocardial infarction specifically for both active treatment and post treatment periods.[20] Definitions for these outcomes were not provided. The Italian trial indicated no increase in myocardial infarction but identified an elevated rate of atrial fibrillation (RR 1.73; 1.02, 2.98) among women randomized to tamoxifen,[29] however, this is the only trial reporting atrial fibrillation. The Royal Marsden trial reported no differences in "cardiovascular problems."[26]

Since tamoxifen showed no differential effects on multiple specific coronary heart disease outcomes, we combined results of composite measures of coronary heart disease in meta-analysis, resulting in a summary risk ratio of 1.00 (0.79, 1.27; 4 trials) (Figure 14).[20,24,26,29] The risk ratio for myocardial infarction specifically is 1.01 (0.63, 1.64; 2 trials) (Figure 15).[20,24,29]

All four prevention trials evaluated stroke outcomes, and stroke was a predefined outcome in the IBIS trial. None of the trials indicated how stroke was defined or whether it was adjudicated. Tamoxifen did not increase stroke in either the active or post treatment periods of the Royal Marsden[26] and IBIS[20] trials. The Italian[29] and NSABP P-1[24] trials reported elevated risk ratios for stroke during active treatment that did not reach statistical significance (Italian RR 3.11; 0.63, 15.4; NSABP P-1 RR 1.59; 0.93, 2.77). The summary risk ratio for stroke is 1.36 (0.89, 2.08; 4 trials) (Figure 16). After discontinuation of treatment in the IBIS[20] and Royal Marsden[26] trials, tamoxifen had no effect on stroke (RR 0.83; 0.20, 3.42; 2 trials) (Figure 16).

Tamoxifen did not increase risk for transient ischemic attack in the trials evaluating this outcome (RR 0.77; 0.46, 1.30; 3 trials) (Figure 17).[20,24,29]

Raloxifene vs. placebo. Cardiovascular outcomes were extensively evaluated in the MORE and RUTH trials.[32,46] In the MORE trial, raloxifene did not increase risk for a composite measure of coronary heart disease, including myocardial infarction, coronary death, silent myocardial infarction, sudden death, unstable angina, coronary ischemia, and acute coronary syndrome (RR 0.92; 0.66, 1.27).[32] Results using a more narrow definition of coronary heart disease events, including coronary death, myocardial infarction, and unstable angina, were similar. Follow-up in

the CORE trial also showed no relationship between the use of raloxifene for 8 years and major cardiovascular events (HR 1.16; 0.86, 1.56) or coronary events (RR 1.22; 0.82, 1.83).[72]

The RUTH trial was designed to identify whether raloxifene prevented coronary heart disease among women at high risk for heart disease or with existing heart disease. In RUTH, raloxifene showed no benefit in reducing composite coronary heart disease outcomes including coronary heart disease death, non-fatal myocardial infarction, and acute coronary syndrome (RR 0.95; CI 0.84, 1.07).[46] Combining coronary heart disease composite measures from MORE and RUTH provides a summary risk ratio of 0.95 (0.84, 1.06; 2 trials) (Figure 14).

Raloxifene did not increase risk of stroke in the MORE[32] or RUTH[46] trials (RR 0.96; 0.67, 1.38; 2 trials) (Figure 16). In CORE, raloxifene did not increase risk of stroke after eight years of treatment and follow up.[72] None of the trials evaluated transient ischemic attacks.

Tibolone vs. placebo. Tibolone did not increase risk for coronary heart disease in the LIFT trial[10] or in another smaller trial.[93] Reports of sinus bradycardia were higher with tibolone in the LIFT trial.[10]

The LIFT trial ended early because of increased ischemic and hemorrhagic strokes in tibolone users (RR 2.19; 1.14, 4.23).[10] In LIFT, transient ischemic attacks were reported as rare in both tibolone group and placebo groups (0.3 % vs. 0.2 %, respectively).[10]

Genitourinary outcomes

Tamoxifen vs. raloxifene. Raloxifene users had lower rates of endometrial cancer than tamoxifen users in STAR (1.25 vs. 2.0 per 1000 women years, respectively),[12] but differences were not statistically significant (RR 0.62; 0.35, 1.08).[12] Raloxifene users had fewer hysterectomies than tamoxifen users (RR 0.44; 0.35, 0.56),[12] with rates of 6.04 vs. 13.37 per 1000 women years, respectively; and fewer cases of endometrial hyperplasia (RR 0.16; 0.09, 0.29).[12] The STAR trial found no differences in other genitourinary cancers.[12]

Tamoxifen vs. placebo. Three prevention trials reported data on endometrial cancer;[20,24,26] the Italian trial included only women with prior hysterectomies.[50] Trials evaluated endometrial changes in different ways. The Royal Marsden trial evaluated endometrial thickness with ultrasound, although the protocol was not reported.[62] The IBIS trial included endometrial cancer as a predefined outcome. The NSABP P-1 trial monitored gynecologic conditions and procedures during the course of the trial.[57] In the NSABP P-1 trial, women randomized after July 1994 underwent endometrial sampling prior to randomization, suggesting that women with abnormal sampling were excluded from the trial creating a cohort at lower risk for endometrial cancer.[24]

All three trials reported increased risks for endometrial cancer with tamoxifen, although only results from the active treatment period of the NSABP P-1 trial reached statistical significance (RR 2.53; 1.35, 4.97).[24] Combining these results from the three trials provides a summary risk ratio of 2.13 (1.36, 3.32; 3 trials) (Figure 18). As noted above, the NSABP P-1 trial was unblinded in 1998, however, women continued to be followed for both breast cancer and other outcomes. Nearly one-third of women in the placebo arm of this trial went on to either participate in the STAR trial or electively begin tamoxifen. With these limitations in mind, the risk of endometrial cancer reported after 7 years of follow-up in this trial was even higher (RR

3.28; 1.87, 6.03).[23] When this estimate is included in the meta-analysis, the summary risk ratio is 2.43 (1.50, 4.00; 3 trials).

In the NSABP P-1 trial, tamoxifen increased rates of endometrial hyperplasia without atypia (RR 2.06; 1.64, 2.60)[57] and other benign gynecologic conditions for both pre and postmenopausal women. For premenopausal women, these included endometrial polyps (RR 1.9; 1.55, 2.41), leiomyomas (RR 1.3; 1.14, 1.55), endometriosis (RR 1.9; 1.35, 2.70), and ovarian cysts (RR 1.5; 1.2, 1.78), as well as gynecologic surgical procedures including hysterectomy (RR 1.6; 1.88, 11.29).[57] For postmenopausal women, these included endometrial polyps (RR 2.4; 1.65, 3.24), leiomyomas (RR 1.4; 1.04, 1.80), endometriosis (RR 1.9; 1.29, 5.58), and gynecologic procedures (RR 2.2; 1.6, 3.13).[57] Tamoxifen had similar effects in the IBIS trial increasing rates of gynecologic procedures including hysterectomy, abnormal bleeding, endometrial polyps, and ovarian cysts.[19] Tamoxifen was associated with higher rates of hysterectomy in the Royal Marsden trial than placebo (177 vs. 96 per 1000 women years, respectively; $p<0.001$).[26] None of the tamoxifen trials reported rates of ovarian cancer.

Tamoxifen increased vaginal symptoms, including dryness, discharge, and other types, in all of the prevention trials.[20,24,26,29] Over twice as many women using tamoxifen vs. placebo reported vaginal discharge ($p<0.001$) or vaginal symptoms ($p=0.008$) in the Royal Marsden trial.[26] In the NSABP P-1 trial, 13% of women taking placebo and 29% taking tamoxifen reported vaginal discharge that was at least moderately bothersome.[24] Tamoxifen increased risks for vaginal dryness (RR 1.14; 0.97, 1.34) and discharge (RR 3.44; 2.9, 4.09) in the Italian trial.[29]

Tamoxifen increased symptoms of cystitis and incontinence in the Italian trial (RR 1.52; 1.23, 1.89),[29] but not similar symptoms during and after active treatment in the Royal Marsden trial.[26]

Raloxifene vs. placebo. The raloxifene trials differed in their methods of ascertaining endometrial cancer outcomes. In the MORE trial, 17 clinical centers performed annual transvaginal ultrasonography in all subjects with a uterus, carefully monitoring uterine pathology.[39] In the RUTH trial, endometrial cancer was determined on the basis of unsolicited reporting by the participant.[46] In neither trial were the risks of endometrial cancer elevated (combined RR 1.14; 0.65, 1.98; 2 trials) (Figure 18).[39,46]

Raloxifene did not cause uterine bleeding in several trials[33,46,73,74,77-80,82,84] and the one observational study[86] reporting this outcome. Raloxifene increased rates of endometrial cavity fluid, as determined by periodic transvaginal ultrasound in the MORE trial ($p<0.009$).[33] Raloxifene did not increase rates of ovarian cancer in RUTH, the only trial reporting this outcome.[46] Raloxifene increased urinary symptoms in the CORE trial (2.1% raloxifene vs. 1.2% placebo; $p=0.041$).[51]

Tibolone vs.. placebo. Three studies provide conflicting data on tibolone and endometrial cancer. The OPAL trial[90] reported only one case of endometrial cancer in each of the placebo and treatment groups, while women with an intact uterus in the LIFT trial[87] had a trend toward increased risk with tibolone (0 vs. 4 cases, respectively, $p=0.06$). The mean age of women in the LIFT trial was 10 years older than the age of women in the OPAL trial (68 vs. 58.7, respectively). In contrast, tibolone users with a mean age of 58 years and no prior cancer or hysterectomy in the U.K. Million Women's cohort study showed an increased risk for endometrial cancer (RR 1.79; 1.43, 2.25).[97] In the Million Women's Study, endometrial cancer risk was increased for woman age ≥60 and with >3 years use of tibolone compared with younger

women and shorter durations of use.[97] Tibolone did not increase risk for cervical cancer[10] or uterine cancer[89] in the two trials reporting these outcomes.

Tibolone did not increase risk for endometrial hyperplasia and moderate or severe dysplasia;[10] however, tibolone was associated with increased rates of procedures for endometrial thickness, hyperplastic polyps,[87] and endometrial biopsy.[10] Tibolone did not increase endometrial thickness in two other trials, the large OPAL trial in the U.S. and Europe and another small Romanian study.[94]

Tibolone increased vaginal bleeding and spotting in the LIFT and OPAL trials.[87,90] A large Scandanavian trial in younger women reported a dose effect for bleeding and spotting with highest rates with 5 mg/day.[91] Tibolone did not increase vaginal bleeding rates at 6 month follow-up in a trial that reported 12% to 15% bleeding rates.[94] Other trials report bleeding or spotting as tolerable with no differences between tibolone and placebo.[92-94]

Tibolone increased pelvic pain, vaginal infection, and vaginal discharge in LIFT.[10] Tibolone did not increase rates of uterine spasm,[93] enlarged abdomen,[93] genital pruritus,[93] or abdominal pain.[91] Tibolone improved vaginal maturation measures,[93] vaginal dryness, and sexual function.[94]

Non-cancer breast outcomes

Tamoxifen vs. raloxifene. No results.

Tamoxifen vs.. placebo. Tamoxifen is associated with reductions in breast density in both the IBIS and NSABP P-1 trials. In a subsample of 69 women in the IBIS trial, at 18 months, women on tamoxifen had a 7.9% greater decrease in breast density than women on placebo; at 54 months, the difference was 13.7% ($p<0.001$).[20] In the NSABP P-1 trial, between 1 to 3.4 years, 38.5% of tamoxifen users had decreased breast density compared with 6.7% of placebo ($p<0.069$),[55] and between 3.5 and 5 years, the difference was 48% compared with 22% ($p<0.114$).[55] Tamoxifen did not cause breast symptoms in the IBIS and Royal Marsden trials.[20,26]

Raloxifene vs. placebo. Raloxifene did not decrease breast density in a small trial of postmenopausal women with osteoporosis (1.3% reduction for placebo, 1.5 % for raloxifene 60 mg/day, 1.7% raloxifene 120 mg/day).[75]

Tibolone vs. placebo. Tibolone did not reduce breast density[94,95] or cause breast pain.[92,93,95] Breast pain ranged from approximately 5%[93] to 10%[94] in tibolone users. Tibolone users without prior hysterectomies in the LIFT trial had more breast discomfort.[10]

Ophthalmologic disorders

Tamoxifen vs. raloxifene. In the STAR trial, women on raloxifene had fewer cataracts (RR 0.79; 0.68, 0.92) and cataract surgery (RR 0.82; 0.68, 0.99) than women on tamoxifen.[12]

Tamoxifen vs. placebo. All four prevention trials evaluated ocular outcomes,[20,24,26,29] although the Italian trial reported data on the composite category of "ophthalmologic diseases."[29] None of the trials described how women were evaluated for ophthalmologic outcomes. The

NSABP P-1,[24] Royal Marsden,[26] and IBIS[20] trials reported increased cataracts with tamoxifen, although results for the IBIS trial did not reach statistical significance. Combining results in meta-analysis indicates a summary risk ratio of 1.13 (0.70, 1.83; 3 trials) (Figure 19). A sensitivity analysis including 7-year follow-up data from the NSABP P-1 trial[23] (see limitations discussed above) rather than short-term follow-up, indicates a summary risk ratio of 1.27 (1.00, 1.62).[20,23,26] Cataract surgery was also evaluated in the NSABP-1 trial and risk estimates were elevated in the initial (RR 1.57; 1.16, 2.14)[24] and follow-up (RR 1.21; 1.10, 1.34)[23] reports.

Raloxifene vs. placebo. Raloxifene did not cause more cataracts than placebo in the MORE and RUTH trials.[39,46]

Tibolone vs. placebo. Tibolone did not increase rates of retinal detachment in one trial.[91]

Gastrointestinal and hepatobiliary disorders

Tamoxifen vs. raloxifene. No results.

Tamoxifen vs. placebo. Tamoxifen did not cause gastrointestinal symptoms in the Italian and Royal Marsden trials.[26,29]

Raloxifene vs. placebo. In RUTH, raloxifene caused more cholelithiasis and dyspepsia (230 compared with 186; p=0.03), although rates of cholecystectomy were similar.[46]

Tibolone vs. placebo. Tibolone did not cause cholecystitis,[91] but increased liver function tests;[10] gastroenteritis was more common with placebo.[10] In LIFT, tibolone reduced risk for colon cancer (RR 0.31; 0.10, 0.96).[10]

Other outcomes impacting quality of life

Tamoxifen vs. raloxifene. In STAR, mean scores on quality of life instruments (health survey, depression scale, sexual questionnaire) did not differ between women using tamoxifen vs. raloxifene, except sexual function was slightly better for tamoxifen (odds ratio, 1.22%; 1.01, 1.46).[18] Women using raloxifene reported more musculoskeletal problems, dyspareunia, and weight gain, while those using tamoxifen had more gynecological problems, vasomotor symptoms, leg cramps, and bladder control symptoms.[18]

Tamoxifen vs. placebo. Tamoxifen increased vasomotor symptoms in the four prevention trials,[20,24,26,29] although vasomotor and gynecologic symptoms were combined in the IBIS trial.[20] In the Royal Marsden trial, 32% of women taking placebo reported hot flashes vs. 48% of women taking tamoxifen (p<0.001).[26] The NSABP P-1 trial had similar findings; hot flashes in 29% of placebo and 46% of tamoxifen groups.[24] In the Italian trial, the risk ratio for hot flashes with tamoxifen was increased at 1.78 (1.57, 2.0).[29]

Two studies from the NSABP P-1 trial evaluated outcomes of depression and quality of life and identified no increased depression with tamoxifen.[21,22,59] Women randomized to tamoxifen reported 4% more sexual side effects than women randomized to placebo, although women on tamoxifen were slightly more sexually active (p=0.031).[59] Tamoxifen caused weight

gain in the Royal Marsden trial,[26] but not in the Italian trial.[29] Tamoxifen did not increase headaches in the IBIS or Royal Marsden trials.[20,26]

Raloxifene vs. placebo. Raloxifene increased vasomotor symptoms in both the MORE and RUTH trials.[33,46] In MORE, 7% of women using placebo, 11% using raloxifene 60 mg, and 12% using raloxifene 120 mg reported vasomotor symptoms ($p<0.05$).[33] In the RUTH trial, comprised of older women, the rates of vasomotor symptoms were lower in general than in MORE, but higher for women taking raloxifene compared with placebo (4.8% placebo vs. 8.0% raloxifene; $p<0.001$).[46] Raloxifene also caused hot flashes and other vasomotor symptoms in three[77,78,80] of eight smaller trials that evaluated vasomotor effects.[74,77,78,80-84]

Raloxifene caused leg cramps in three[33,46,80] of six trials.[33,46,77,78,80,82] Raloxifene caused peripheral edema in the MORE (6.1% placebo, 7.1% raloxifene 60 mg, 7.9% raloxifene 120 mg; $p=0.026$)[33] and RUTH trials (12.1% placebo, 14.4% raloxifene; ($p<0.001$).[46]

Influenza syndrome symptoms occurred at a higher rate among women taking raloxifene in MORE (14% placebo, 16.2% raloxifene 60 mg, 16.7% raloxifene 120 mg),[33] but not in two other studies.[46,84] Raloxifene caused joint pain in two trials,[46,79] but not in a third.[84] Raloxifene had no effect on mood, depression, and anxiety symptoms in three trials.[46,83,84]

Tibolone vs. placebo. Unlike tamoxifen and raloxifene, tibolone reduces vasomotor symptoms, such as the number and severity of hot flashes.[91,93,94] One study showed reduction in hot flashes for the 2.5 mg/day tibolone dose, but not in the 0.3-1.25 mg/day doses.[92] Tibolone did not increase weight in two trials.[92,93] Measures on the Beck Depression Inventory were improved with tibolone after one year of treatment in one trial.[96]

Tibolone did not cause several other symptoms that impact quality of life in trials measuring these outcomes, such as musculoskeletal disorders,[89] headache,[91-93] back or abdominal pain,[92] upper respiratory[93] or respiratory tract infection,[92] allergy,[92],sinusitis,[92] accidental injury,[92] anxiety and nervousness,[92] nausea,[93,94] fluid retention,[94] and concussion.[91] Tibolone did not cause moniliasis in the 0.3–1.25 mg/day doses, however, was greater in the 2.5 mg/day dose compared to placebo.[92]

Key Question 3. How do outcomes for tamoxifen citrate, raloxifene, and tibolone when used for primary prevention of breast cancer vary by heterogeneity in subpopulations?

Key Points

- Tamoxifen and raloxifene had similar effects on breast cancer outcomes regardless of age and family history of breast cancer in the head-to-head STAR trial.
- Tamoxifen reduces breast cancer outcomes in subgroups evaluated in prevention trials based on age, menopausal status, estrogen use, family history of breast cancer, and history of lobular carcinoma *in situ* or atypical hyperplasia. In the NSABP P-1 trial, cancer rates were highest and risk reduction greatest among women in the highest modified Gail model risk category and among women with prior atypical hyperplasia.
- Raloxifene reduces breast cancer outcomes in subgroups evaluated in prevention trials based on age, age at menarche, parity, age at first live birth, and body mass index.

Estimates from subgroups based on prior estrogen use, family history of breast cancer, and prior hysterectomy or oophorectomy are limited by smaller numbers of subjects.

- Thromboembolic events and endometrial cancer were more common in older (>50) than younger women in the NSABP P-1 trial.
- Tibolone causes more strokes in older (>70 years) than younger women.

Detailed Analysis

Some prevention trials provide data for important subgroups, although outcomes are predominantly confined to breast cancer (all breast cancer, invasive, and estrogen receptor positive). Data are available for subgroups based on age,[12,20,23,29,42,47] menopausal status,[20,26] hysterectomy status,[47] estrogen use,[20,23,29,42,47] family history of breast cancer,[12,23,29,42,47] body mass index,[42,47,106] history of breast abnormalities,[12,23] predicted breast cancer risk,[12,23,47] estradiol levels,[42] and reproductive factors.[47] No trials reported outcomes by race or ethnic groups.

Age

The STAR,[12] IBIS,[20] Italian,[29] NSABP P-1,[23] RUTH,[47] and MORE.[42] trials evaluated breast cancer outcomes by age categories, although categories varied by trial. In STAR, invasive cancer outcomes did not differ significantly for women using raloxifene vs. tamoxifen in the three age categories evaluated (≤49, 50 to 59; ≥60 years), and results were similar across categories.[12] In the three tamoxifen vs. placebo trials, summary risk estimates for invasive or all cancer outcomes were significantly reduced and similar for women ≤50 and >50 years (Figure 20).[20,23,29] The raloxifene trials stratified results for invasive cancer using different age categories (MORE <65 years; RUTH <60 years) and we did not combine them in a meta-analysis. MORE reported a reduced risk ratio for women ≥65 vs. <65 years,[42] and RUTH an increased risk ratio point estimate for women ≥60 vs. <60 years,[47] although confidence intervals overlap (Figure 20).

The NSABP P-1 trial suggested higher risks for deep vein thrombosis, pulmonary embolism, and stroke for women >50 vs. ≤50 years; rates and risk ratios are higher, but results are not statistically significant.[24] Age >60 years was also an important risk factor for venous thrombosis in the Italian trial.[27] The NSABP P-1 trial also found that endometrial cancer was more common among women >50 vs. ≤50 years (RR 4.01; 1.70, 10.90 vs. 1.21; 0.41, 3.60; respectively).[24] In LIFT, rates of stroke were highest among tibolone users age >70 vs. 60 to 70 years (6.6 vs. 2.8 per 1000 women years).[10]

Menopausal status

The IBIS[20] and Royal Marsden[26] trials evaluated breast cancer outcomes by menopausal status (pre vs. post). Point estimates indicate similar risk reduction with tamoxifen vs. placebo for both pre and postmenopausal women, although results were not statistically significant for postmenopausal women in both trials (Figure 21). We detected no significant differences between pre and postmenopausal women by subgroup comparison analysis.

Hysterectomy status

In RUTH, raloxifene did not significantly reduce risk for invasive cancer for women with prior hysterectomies or oophorectomies, while risk reduction was significant in women without

these prior surgeries.[47] However, these differences could reflect the smaller numbers of women in the surgical subgroups.

Use of exogenous estrogen

The IBIS,[20] Italian,[29] NSABP P-1,[23] RUTH,[47] and MORE[42] trials evaluated breast cancer outcomes by use of menopausal hormone therapy (estrogen with or without progestin). In the tamoxifen trials, women were allowed to use hormones during the trial, and use rates varied from <10% in NSABP P-1[24] to 40% in IBIS.[19] Women in the raloxifene trials were not allowed to use hormones during the trial and hormone use status represented prior use. For both tamoxifen and raloxifene trials, point estimates improved and results became statistically significant for hormone nonusers compared to users, although summary estimates were not significantly different (Figure 22). These findings may reflect the smaller numbers of hormone users in the trials.

Risk of breast cancer

Family history. The STAR,[12] Italian,[29] NSABP P-1,[23] RUTH,[47] and MORE[42] trials evaluated breast cancer outcomes by family history of breast cancer, most commonly referring to the number of first-degree relatives with breast cancer. In STAR, invasive cancer did not differ significantly for women using raloxifene vs. tamoxifen in the three family history categories evaluated (0, 1; >2), and results were similar across categories.[12] Tamoxifen reduced invasive and all breast cancer for women without a family history in the two tamoxifen vs. placebo trials, but had dissimilar results for women with a family history (Figure 23). In the NSABP P-1 trial, risk was similar for women in both family history groups; in the Italian trial, risks were reduced for women with no family history and increased for women with family history, although results were not statistically significant (Figure 23). The raloxifene trials indicate similar significantly reduced risk estimates for women without family history and dissimilar results for women with family history (Figure 23). These results may reflect the smaller numbers of women with positive family history for breast cancer in these trials rather than true medication effects. We did not combine results for women with family history for tamoxifen or raloxifene trials in a meta-analysis.

Body mass index. A nested case-control analysis of data from the NSABP P-1 trial indicates that elevated body mass index is associated with higher risk of thrombembolic events among women in both the placebo and control groups (RR 3.69; 2.09, 6.65).[106] Additional analysis of the prothrombin gene mutation and Factor V Leiden deficiency indicated no interaction with tamoxifen and risk of thromboembolic events. This analysis also indicated that the risk of thromboembolic events was elevated only during the first 3 years of use of tamoxifen. The RUTH and MORE trials evaluated invasive breast cancer by body mass index (BMI ≤25 vs. >25).[42,47] While MORE indicated similar significantly reduced risk estimates for women with low and high BMI, RUTH reported lower risk estimates for women with high BMI (Figure 24), although estimates were not significantly different between women with low or high BMI.

History of breast abnormalities. In STAR, tamoxifen and raloxifene had similar effects on invasive breast cancer regardless of history of LCIS or atypical hyperplasia.[12] In NSABP P-1, tamoxifen reduced invasive cancer compared to placebo regardless of history of LCIS or atypical

hyperplasia, although reduction was greatest among women with prior atypical hyperplasia (RR 0.25; 0.10, 0.52).[23]

Predicted breast cancer risk. In STAR, tamoxifen and raloxifene had similar effects on invasive breast cancer for women in all risk categories determined by the modified Gail model (5-year predicted risk ≤3.00; 3.01 to 5.00; ≥5.01).[12] In NSABP P-1, tamoxifen reduced risk for invasive cancer compared to placebo for women in all modified Gail model risk categories (5-year predicted risk ≤2.00, 2.01 to 3.00; 3.01 to 5.00, ≥5.01).[23] Cancer rates were highest and risk reduction greatest among women in the highest risk group in this trial. In RUTH, raloxifene reduced risk for invasive cancer compared to placebo for women in all modified Gail model risk categories (5-year predicted risk ≤2.00, 2.01 to 3.00; 3.01 to 5.00), although results were statistically significant only for the large number of women in the lowest risk group.[47]

Estradiol levels. Raloxifene had less effect on invasive cancer outcomes among women with estradiol levels <5 pmol/L (RR 0.52; 0.26, 1.06) than women with higher levels (5 to 10 pmol/L, RR 0.33 ; 0.13, 0.84; >10 pmol/L, RR 0.25; 0.14, 0.47) in MORE/CORE.[42]

Reproductive factors. Raloxifene reduced risk for invasive cancer regardless of age at menarche (<11, ≥11 years), parity (0, 1 to 2, ≤3), or age at first live birth (<20, ≥20 years) in the RUTH trial.[47]

Key Question 4. What is the evidence that harms or secondary potential benefits listed above affect treatment choice, concordance, adherence, and persistence to treatment with tamoxifen citrate, raloxifene, and tibolone when used for primary prevention of breast cancer?

Key Points

- Comparisons of adherence and persistence rates across medications in prevention trials are limited because few trials report treatment duration, completion rates, or other measures of adherence and persistence, and trials were designed for different treatment purposes.
- Discontinuation rates for tamoxifen or raloxifene are generally higher than placebo. In the few trials reporting discontinuation rates, the differences between treatment and placebo groups were ≤2% for adverse events and ≤4% for nonprotocol specified events.
- Women make decisions to use tamoxifen for risk reduction based on their concern for adverse effects as well as their risk for breast cancer according to small descriptive studies.
- Women weigh their physicians' recommendations highly when deciding whether to take tamoxifen for risk reduction according to descriptive studies of concordance.
- Studies of treatment choice and concordance for raloxifene and tibolone for breast cancer risk reduction are lacking.

Detailed Analysis

A total of 24 studies met inclusion criteria for key question 4.[10,12,20,24,26,29,34,46,73,76,79-81,84,90,107-115] Quality ratings for the 16 randomized controlled trials are detailed in prior key questions (Appendix C-5).[10,12,20,24,26,29,34,46,73,76,79-81,84,90,109] The remaining eight studies were not evaluated for quality because they use descriptive methods that are not included in quality rating criteria.[107,108,110-115]

Comparisons of rates of adherence and persistence are limited because few trials reported mean duration of treatment, percentage of subjects completing the planned treatment duration, or other measures of adherence and persistence. Also, the trials were designed for different treatment purposes. The raloxifene trials were intended to prevent fractures in women with preexisting osteoporosis, and were designed for long-term treatment. Tamoxifen trials were designed to test a time-limited prevention intervention in women without pre-existing conditions. This difference makes inferences about comparative adherence difficult. The STAR trial might be able to provide information regarding adherence or compliance of tamoxifen and raloxifene in a comparable population, however the published reports of the trial do not include adherence or persistence data.

Rates of adherence and persistence

Adherence is the extent to which a patient acts in accordance with the prescribed interval and dose of a dosing regimen.[116] Persistence is the duration of time from initiation to discontinuation of therapy.[116]

Adherence was reported by one tamoxifen trial,[26] four raloxifene trials,[34,46,76,84] and one tibolone trial,[10] and was lacking for several trials including STAR (Table 7).[12] Of trials reporting adherence, results indicate at least 70% adherence with the planned treatment dose, however, these data do not allow direct comparisons between trials. In the Royal Marsden trial, adherence was 8% lower with tamoxifen vs. placebo (p=0.002).[26] In RUTH, there were no differences between raloxifene and placebo; 70% vs. 71% took at least 70% of the study medication, respectively.[46] Adherence was not reported separately in MORE; 92% of the entire study population took at least 80% of the assigned study medication.[34] In LIFT, 91% received at least 80% of the assigned study medication.[10]

Persistence was measured as duration of treatment in the STAR trial,[12] one tamoxifen trial,[29] three raloxifene trials,[46,76,80] and one tibolone trial;[10] and as completion of the planned course of treatment by two tamoxifen trials,[20,29] six raloxifene trials,[46,76,79-81,84] and two tibolone trials.[90,109]

In the STAR trial, treatment was ongoing at the time of publication and final persistence rates have not been published, although the mean duration of treatment was similar for raloxifene and tamoxifen (3.2 vs. 3.1 years, respectively).[12] In the Italian trial, designed for 60 months of treatment, women using tamoxifen had lower completion rates than placebo (59.8% vs. 61.8%, respectively).[29] The IBIS trial had similar results, although both groups had higher completion rates than the Italian trial (63.9% vs. 72%, respectively).[20] In RUTH, women using raloxifene had slightly higher completion rates than placebo (80% vs. 79%; p=0.02), although the median duration of treatment was 5.05 years for both groups.[46] Additional trials of raloxifene reported 60% to 91% of subjects completing the planned duration of treatment.[76,79-81,84] In LIFT, prematurely discontinued due to preset stopping rules, the median duration of treatment with tibolone was 34 months.[10] Completion rates in OPAL were 69% for tibolone and 70% for placebo,[90] and 89% overall in another tibolone trial.[109]

Harms or benefits affecting adherence and persistence

Evidence that harms or secondary potential benefits affect adherence and persistence was sporadically reported in tamoxifen and tibolone trials as protocol specified and non-protocol specified events. Protocol specified events are outcomes explicitly stated in the protocol requiring that a participant discontinue the study medication.

Tamoxifen vs. placebo. Two trials reported treatment discontinuation due to non-protocol specified events.[24,29] In the Italian trial, 7.6% of tamoxifen vs. 6.9% of placebo groups withdrew from treatment due to protocol specified events, and 26.7% vs. 25.3% due to non-protocol specified events.[29] In the NSABP P-1 trial, 23.7% of tamoxifen vs. 19.7% of placebo groups discontinued due to non-protocol specified events.[24]

Raloxifene vs. placebo. Eight raloxifene trials provided information on discontinuation rates due to adverse events.[34,46,73,76,79-81,84] In RUTH, 22% of raloxifene and 20% of placebo groups discontinued study medications due to adverse events (p=0.01); specific adverse events were not described.[46] In the MORE trial, significantly more women receiving raloxifene than placebo withdrew from treatment due to hot flashes.[34] In another trial to evaluate the effect of raloxifene on hot flashes in postmenopausal women, vasomotor symptoms caused discontinuation in two women using raloxifene and four using placebo, and 14 other patients discontinued due to other adverse events that were not described.[84] In the OPAL trial, discontinuation rates for hot flashes (5%) and leg cramps (1%) were higher for raloxifene than placebo (1% vs. 0%).[80] In a trial to assess the uterine effects of raloxifene in healthy postmenopausal women, discontinuation due to gynecologic adverse events were not significantly different between groups (3 placebo, 1 raloxifene 60 mg/day, 2 raloxifene 120 mg/day).[73] Three other trials reported discontinuation rates due to adverse events that were not further described.[76,79,81]

Tibolone vs. placebo. The LIFT trial reported higher rates of discontinuation due adverse events for tibolone, but did not provide data.[10] A trial designed to evaluate the effects of 1.25 and 2.5 mg/day doses on early postmenopausal bone loss reported discontinuation rates due to adverse events as 7% for tibolone vs. 17.4% for placebo.[109]

Surveys of treatment choice and concordance

Concordance occurs when a health care provider and patient reach a shared agreement about therapeutic goals. In concordance, the patient is informed of the condition and options for treatment and is involved in the treatment decision.[117] Seven studies described treatment choice for breast cancer risk reducing medications,[108,110-115] and three of these also investigated the relationship between physician recommendations and patient choice (Table 8).[110,114,115] This collection of small descriptive studies suggests that women are making decisions based on their concern for side effects as well as their risk for breast cancer.[110-114] Also, women weigh their physicians' recommendation when deciding whether to use risk reducing medications.[110,114,115] One additional survey of physicians evaluated risk reducing medication prescribing practices.[107] All studies considered tamoxifen use.

In an interview-based cross-sectional study, 17.6% of women were inclined to use tamoxifen following an educational session about its indications and adverse effects.[112] More than half of the subjects listed breast cancer (68.8%), pulmonary embolism (67.2%), endometrial

carcinoma (62.7% of women without a hysterectomy), and deep vein thrombosis (58.4%) as "very important" in making their decisions.

In a study testing a new decision guide for identifying women with high risk for breast cancer and informing them about risk reduction with tamoxifen, women who were interested in taking tamoxifen were allowed to choose between accepting a prescription for tamoxifen or enrolling in STAR.[111] Results indicated that 11.8% of women selected tamoxifen, 76.5% declined, and 11.8% were undecided. Major side effects (60.7%) and small benefit from tamoxifen (32.1%) were the most common reasons for declining. However, 90% of women stated that they would take a medication with the same benefit as tamoxifen if it had no side effects. Approximately half of women also stated that if a medication were developed with the same side effects but could eliminate the chance of getting breast cancer, they would take the medication.

In a pre/post survey study, women completed a questionnaire after receiving information about tamoxifen.[113] Of the 43 subjects, 4.7% selected tamoxifen, 34.8% declined, and 60.5% were undecided. Upon later follow-up, none of the 60.5% who were undecided changed to selecting tamoxifen. Of the patients who did not select tamoxifen, 75.6% reported a concern for side effects, including endometrial cancer and thromboembolic events, as a reason for not using tamoxifen. Other reasons were the feeling that not enough information was available (12.2%) and not wanting to discontinue hormone replacement therapy (4.9%).

A telephone survey of 1,287 women with Blue Cross/Blue Shield insurance was designed to determine if women would be "interested in a medication to prevent breast cancer."[108] The 23% of responders interested in risk reducing medications believed themselves to be at greater risk for breast cancer and were more worried about breast cancer than women who were not interested ($p<0.05$).

Three studies evaluated the relationship between physician recommendations and treatment choice.[110,114,115]

A study of concordance with physician recommendations included women age 35 to 80 years who were evaluated for benign breast findings in a breast clinic.[114] They were provided with Gail model estimates of risk and the option of using tamoxifen for risk reduction, and were asked to discuss tamoxifen use with their family physicians. Of the 89 women, 48 discussed the decision with their family physician. Physicians recommended using tamoxifen for 3 women, not using tamoxifen for 37, and made no recommendation and left the decision up to the patient for 8. Only one woman in the study decided to use tamoxifen. While this study did not include raloxifene as a potential breast cancer risk reduction option, another 5 patients reported that their physicians had prescribed raloxifene for osteoporosis with the secondary benefit of breast cancer prevention. Patients identified one or more of the following factors as influencing their decision: concern for adverse effects (46%), breast cancer risk not high enough to warrant therapy (33%), family physician's decision (31%), personal decision (25%), lack of sufficient information (10%).[114]

A study of patient/physician concordance assessed women's decisions to use tamoxifen or raloxifene at 2 and 4 months after risk counseling.[110] At two months follow-up, 29% of women chose to take tamoxifen, another 27% opted for enrolling in the STAR trial, 24% declined treatment, and 20% were undecided. At 4 months follow-up, 12% changed from choosing or undecided to decline, however, it was unclear whether anyone who changed from choosing tamoxifen to declining had started taking risk reduction medications in the intervening 2 months. Not all women made a decision by the 4-month follow-up, with 13.9% remaining

undecided. All women in this trial were advised by a physician of their eligibility for risk reduction with tamoxifen or raloxifene, however, not all women reported receiving a recommendation from their physician to choose treatment or not. For women who received a recommendation from their physician, most recommendations were related to treatment choice ($p<0.0001$). Concern for side effects of tamoxifen was a significant factor in women's treatment decision ($p<0.006$).[110]

A descriptive study was designed specifically to evaluate the effect of physician recommendations to women eligible for the NSABP P-1 trial.[115] Women were surveyed after attending an informational session about the trial, and 175 of 360 attendees reported having discussed their participation with their primary care physicians and receiving a recommendation for participation or non-participation. Among the 175 women who discussed the decision with their physician, the physician recommendation was related to trial participation ($p<0.001$). Women whose physicians recommended enrollment were 13 times more likely to enroll than women whose physician recommended against enrollment.

A mailed survey to 350 physicians indicated that 27% prescribed tamoxifen for risk reduction for their patients within the prior 12 months.[107] Physicians who had prescribed tamoxifen were more likely to have a family member with breast cancer (19.8% vs. 8.7%; $p=0.004$). Prescribers and nonprescribers differed in their responses to several statements including: the benefits of tamoxifen outweigh the risks (62.5% vs. 39.4%; $p<0.001$), physicians in their community are prescribing tamoxifen for breast cancer prevention (33.3% vs. 16.6%; $p<0.001$), it is easy to determine who is eligible to take tamoxifen for breast cancer risk reduction (28.1% vs. 10.9%; $p<0.001$), and many female patients ask for information about taking tamoxifen for breast cancer risk reduction (14.6% vs. 4.8%, $p=0.002$). Physicians did not differ in their beliefs about the following: whether the evidence that tamoxifen significantly reduces breast cancer is controversial; it is too time consuming to discuss taking tamoxifen with women in my practice; the risk of endometrial cancer is too great to prescribe tamoxifen for breast cancer risk reduction; and, the risk of thromboembolic events is too great to prescribe tamoxifen for breast cancer risk reduction.

Key Question 5. What methods, such as clinical risk assessment models, have been used to identify women who could benefit from breast cancer medications to reduce risk of breast cancer?

Key Points

- Nine risk stratification models that predict an individual's risk for developing breast cancer have been evaluated for use in clinical settings. Models consider multiple risk factors for breast cancer.
- Risk stratification models demonstrate good calibration, with the expected number of breast cancer cases in a study population closely matching the number of breast cancer cases observed.
- All models have low discriminatory accuracy in predicting the probability of breast cancer in an individual. Most models perform only slightly better than age alone as a risk predictor.
- A Gail score of ≥1.66% has been used as a risk threshold in prevention trials and in Food and Drug Administration approval of tamoxifen and raloxifene for breast cancer

prevention. However, this threshold has low discriminatory accuracy in predicting breast cancer in an individual.

Detailed Analysis

A total of 16 studies reporting results of evaluations of nine risk stratification models met inclusion criteria (Table 9).[118-132] Of these, 12 met criteria for good quality because they were adequately described, relevant to primary care practice, used appropriate reference standards, and included large sample sizes. (Appendix C-6)[49,119-122,124-128,130,131] Four met criteria for fair quality because they were developed using secondary data sources,[123] assessed only a 1-year risk for breast cancer,[129] were of questionable feasibility for a primary care setting.[118] or included a small population selected from a nonprimary care setting.[132]

Risk stratification models

The Gail model was the first major breast cancer risk stratification model to be used clinically.[49] This model was derived from multivariate logistic regression analysis of identified risk factors for breast cancer.[49] In the original version of the Gail model, breast cancer incidence rates and baseline hazard rates were determined for invasive cancer, DCIS, and LCIS from a cohort of women in the Breast Cancer Detection and Demonstration Project (BCDDP). The model was subsequently modified by using U.S. national data for invasive cancer from the Surveillance, Epidemiology, and End Results (SEER) program.[124] From these data, a model was developed to allow the prediction of individualized absolute risk (probability) of developing breast cancer in women undergoing annual screening mammography.

Subsequent risk stratification models use a similar approach as the Gail model, however, they vary in their use of reference standards. Age-specific breast cancer rates and attributable risk estimates to determine baseline age-specific hazard ratios should ideally be obtained from an applicable population reference standard, such as SEER data in the U.S. Several studies of newer models do not provide information about their reference standards.[119,120,122,125,131]

Models also vary by the variables they include (Table 10). The original Gail model included age, age at menarche, age of first birth, family history of breast cancer in first degree relatives, number of prior breast biopsies, and history of atypical hyperplasia.[49] Subsequent models include one or more of these variables in addition to other factors. These include race,[125,126,129,130] body mass index or height,[118,119,123,125,128,129] estrogen and progestin use,[118,119,125,129] parity,[119,125] history of breast feeding,[125] menopause status or age,[119,123,129] smoking,[125] alcohol use,[118,119,125] physical activity,[118,125] breast density,[128-130] and diet.[118]

Studies of calibration

Calibration is a measure of how well predicted probabilities agree with actual observed risk. The calibration of a model refers to its ability to predict the average risk in a subset of the population. When the predicted risk matches the proportion that actually develops disease, a model is considered to be well calibrated. In a perfect prediction model, the predicted risk in a population (% expected) would equal the observed number of cases (% observed) such that the % expected/% observed (E/O) equals 1.0.

Of the nine models reviewed, calibration was calculated for all except the Chen, Chlebowski and Boyle model.[118,125,128] For most models, the expected numbers of cases of breast cancer closely match the observed numbers (Figure 25).[118,119,121,123-126,128-130] Six studies evaluated the Gail model,[118,121,122,124,125,132] demonstrating E/O ratios ranging from 0.69 (0.54,

0.90)[132] to 1.03 (0.88, 1.21).[124] Two studies reported values <0.90, indicating under prediction of breast cancer cases.[125,132] In one study, under prediction could be attributed to dissimilarities of the study population; women were included who were undergoing assessment at a family history clinic, rather than a primary care setting, were younger than women in other studies, and were not all undergoing routine mammography screening.[132] The Gail model demonstrated good calibration for estrogen receptor positive cancers (E/O 1.06), but inferior calibration when estrogen receptor negative cancers were included (E/O 0.79).[125]

The Gail model was modified to evaluate its utility in a population of Italian women. The Italian Gail Model (IT-GM) differed from the Gail model by one ordinal value for one variable, and the Italian-1 Gail Model (IT1-GM) differed by using categorical rather than ordinal variables.[121] Both versions demonstrated good calibration in two studies.[118,121] In one study, E/O ratio for the IT-GM was 0.96 (0.84, 1.17) and 1.00 (0.88, 1.16) for the IT1-GM.[121] A second study demonstrated good calibration for the IT-GM (E/O 1.04).[118] The Gail model itself also demonstrated good calibration in this population (E/O 0.96; 0.84, 1.17;[121] E/O 1.12).[118]

All of the other models demonstrated good calibration across the studies (E/O 1.00 to 1.09),[126,129-132] except for the Tyrer-Cuzick model assessing risk in a population with biannual mammography screening (E/O 0.81; 0.62, 1.08).[132] Categories based on age demonstrated good calibration in the Gail [121,122,124] and BCSC-Tice models,[130] except for women <50 years in an Italian population (E/O 0.61; 0.49, 0.80)(Figure 25). When age alone was used to calculate risk of developing breast cancer in an Italian population, breast cancer was under predicted (E/O 0.73; 0.64, 0.86).[121] Two models that include race also demonstrated good calibration, the Gail-AA model for use in the U.S. African American population[126] and the BCSC-Tice model.[130]

Studies of discriminatory accuracy

Discriminatory accuracy is a measure of how well the model can separate those who do and do not have the disease of interest. In diagnostic testing, it is the ability to identify individuals with or without the disease of interest. In prognostic modeling, it is the ability to correctly classify individuals at higher risk from those at lower risk, and is measured by the model's concordance statistic or c-stat. The c-stat is determined by the area under the receiver operator curve, a plot of sensitivity (true positive rate) versus 1-specificity (false-positive rate). Perfect discrimination is a c-stat of 1.0 and occurs when all cases attain higher risk scores than all non-cases. A c-stat of 0.5 would result from chance alone. An acceptable level of discrimination is considered as ≥0.70 and <0.80, excellent ≥0.80 and <0.90, and outstanding ≥0.90.[133]

Thirteen studies of nine models indicate that discriminatory accuracy for most models is <0.70 (Figure 26).[118,120-122,124-132] Only one study reported levels >0.70 for both the Gail-2 and the Tyrer-Cuzick models, with c-stats of 0.74 (0.67, 0.80) and 0.76 (0.70, 0.82), respectfully.[132] However, this study was small (<100 cases) and did not include a primary care population, limiting its clinical applicability. The BCSC-Tice model, drawing from large U.S. national populations, provided the next highest discriminatory accuracy, with a c-stat of 0.66 (0.65, 0.66).[130] The model with the lowest level of discrimination was the Gail-AA, with a c-stat of 0.56.[126,127] The discriminatory accuracies of age[129,131] or breast density alone[129] as a predictor of breast cancer risk ranged from 0.55 to 0.57 and 0.55 to 0.56, respectfully.

Studies of risk quintiles

In some of the breast cancer primary prevention trials, women were assessed for their individual risks for developing breast cancer, and only those meeting established risk thresholds were eligible to participate.[12,19,23-25] Three studies evaluated this approach to risk stratification by determining calibration and/or discriminatory accuracy based on risk quintiles,[121,125,130] and one study determined these values based on a low (<1.67%) vs. high (≥1.67%) risk threshold (Table 11).[130] This threshold was used as inclusion criteria in the NSABP P-1 and STAR trials, and is included in the FDA's approval of the use of SERMS for risk reduction. The BCSC-Tice model demonstrated high calibration (E/O 0.99 to 1.03), and consistent, although low, discriminatory accuracy across the quintiles (c-stat 0.61 to 0.64).[130] The Gail and the Italian Gail Model demonstrated high calibration in the higher risk quintiles, but variable results in the lower quintiles (Table 11).[121,125]

Summary and Discussion

EPC GRADE

Results for major clinical outcomes are summarized in an EPC GRADE table of evidence (Table 12). Major clinical outcomes are those explicitly stated in key questions 1 and 2; identified as important outcomes by members of the Technical Expert Panel because they are most relevant to patients, clinicians, and policymakers; and have adequate data from studies meeting eligibility criteria for the comparative effectiveness review. Outcomes included in the GRADE table are invasive breast cancer, estrogen receptor positive breast cancer, estrogen receptor negative breast cancer, noninvasive breast cancer, all-cause death, vertebral fractures, nonvertebral fractures, thromboembolic events, coronary heart disease events, stroke, endometrial cancer, and cataracts.

The EPC GRADE table includes the four required domains–risk of bias, consistency, directness, and precision (terms defined in Appendix C-2).[15] Additional optional domains were not included in the table because they are not relevant to this review (Appendix C-4). The table summarizes the strength of evidence; estimates of effect using risk ratios from trials and meta-analyses detailed in the report; and estimates of magnitude of effect expressed as the number of events reduced or increased per 1000 women years assuming 5 years of use of tamoxifen, raloxifene, or tibolone.

Risk of Bias

Risk of bias incorporates both study design and study conduct.[15] In general, we ranked risk of bias low because results for all major outcomes were derived from randomized controlled trials with good aggregate quality. These included eight large randomized controlled trials that each met criteria for fair or good quality based on their use of appropriate clinical trial methodology and analysis (Appendix C-5). Additional smaller trials provided data on harms. Although these studies are included in the review and GRADE table, they rarely reported the major clinical outcomes addressed by the table. No nonrandomized effectiveness studies of medications to reduce risk for primary breast cancer have been published. No relevant observational studies of tamoxifen, and only one of raloxifene were identified in our searches or by our technical experts. Observational studies of tibolone, such as the Million Women Study, are likely biased for some of the major outcomes in the GRADE table because they focus on women using tibolone to treat menopausal symptoms.[98,103] This design introduces multiple uncontrolled confounders compromising results.

Consistency

Consistency refers to the degree of similarity in the effect sizes of different studies within an evidence base.[15] In most cases, we ranked this domain as consistent for tamoxifen and raloxifene because the effect sizes of randomized controlled trials for the major clinical outcomes were generally in the same direction of effect, they usually had narrow ranges of effect sizes, and results of placebo-controlled trials were generally consistent with results of the STAR head-to-head trial. We also considered measures of heterogeneity from our meta-analyses in evaluating consistency (Figures 3 to 19). We ranked this domain inconsistent for noninvasive breast cancer and cataracts for tamoxifen because the results of the placebo-controlled NSABP

P-1 trial differed from the meta-analysis of tamoxifen trials. The NSABP P-1 trial is particularly relevant because it is based in the United States, is the largest trial, and meets criteria for good quality and applicability. Results for tibolone were based on a single trial and consistency could not be evaluated.

Directness

Directness has two meanings: (1) evidence links the interventions directly to health outcomes, and (2) evidence compares two or more interventions in head-to-head trials.[15] All trials included in this review linked the evidence directly to health outcomes. The EPC GRADE table focuses on the second meaning for directness, whether evidence came from direct (head-to-head) or indirect (placebo-controlled) trials. Direct evidence comparing tamoxifen and raloxifene resolved important discrepancies arising from the placebo-controlled trials, such as magnitudes of effect. Women enrolled in the raloxifene placebo-controlled trials were 15 to 20 years older than women in the tamoxifen placebo-controlled trials. This age difference accounts for the higher incidence rates of most of the clinical outcomes in the raloxifene trials. Older women have higher risks for breast cancer, thromboembolic events, and other outcomes than younger women and would likely demonstrate larger medication effects for benefits as well as harms. The STAR trial allows direct comparisons between similar groups of women providing a better assessment of advantages and disadvantages of one medication over the other. Women in STAR were more similar to women in the tamoxifen than the raloxifene trials because they were closer in age and inclusion criteria were based on breast cancer risk as determined by the Gail model. No head-to-head trials including tibolone are available.

Precision

Precision is the degree of certainty surrounding an estimate of effect for specific outcomes.[15] The methodology for determining precision for EPC GRADE tables emphasizes the need to include both clinical and statistical considerations (Appendix C-2). For this comparative effectiveness review, we considered estimates precise if they provided statistically significant differences between medications, or between medications and placebo, for major clinical outcomes that would support clinical decisions (conceptual confidence). Estimates were also considered precise if they showed no statistically significant differences between comparators, and confidence intervals did not range beyond 0.67 to 1.50 (statistical precision of effect estimation). Estimates indicating no statistically significant differences between comparators with wider confidence intervals were considered imprecise because they could be compatible with different clinical conclusions.

Most comparisons in the EPC GRADE table met criteria for precise estimates (Table 12). Estimates are imprecise for some comparisons with placebo including estrogen receptor negative breast cancer (tamoxifen, raloxifene), noninvasive breast cancer (tamoxifen, raloxifene), vertebral fractures (tamoxifen), thromboembolic events (tibolone), coronary heart disease events (tibolone), stroke (tamoxifen), endometrial cancer (raloxifene), and cataracts (tamoxifen). For head-to-head comparisons of raloxifene and tamoxifen, estimates are imprecise for estrogen receptor negative breast cancer, noninvasive breast cancer, vertebral fractures, stroke, and endometrial cancer.

Strength of Evidence

We qualitatively rated the overall strength of evidence as high, moderate, low, or insufficient for each outcome based on the required domains and other relevant factors (Appendix C-3). Strength of evidence is high for outcomes with low risk of bias, consistency, and adequate precision. Outcomes with results from placebo-controlled trials that were consistent with results from the head-to-head STAR trial provided additional support for the high strength of evidence grade. Outcomes with high strength of evidence include invasive breast cancer (tamoxifen, raloxifene), estrogen receptor positive breast cancer (tamoxifen, raloxifene), all-cause death (short-term) (tamoxifen, raloxifene), vertebral fractures (raloxifene), nonvertebral fractures (raloxifene), thromboembolic events (tamoxifen, raloxifene), coronary heart disease events (tamoxifen, raloxifene), endometrial cancer (tamoxifen), and cataracts (raloxifene) (Table 12).

The strength of evidence for outcomes with imprecise estimates, inconsistency between trials, or based on only one trial was downgraded to moderate. These include invasive breast cancer (tibolone), estrogen receptor negative breast cancer (tamoxifen, raloxifene), noninvasive breast cancer (raloxifene), vertebral fractures (tibolone), nonvertebral fractures (tamoxifen, tibolone), stroke (tamoxifen, raloxifene, tibolone), and endometrial cancer (raloxifene).

Strength of evidence was ranked low if multiple deficiencies existed. Strength of evidence for tamoxifen was low for noninvasive breast cancer and cataracts because placebo-controlled trials were both inconsistent and imprecise; also, results of placebo-controlled trials were inconsistent with STAR for cataracts. Strength of evidence for tamoxifen was ranked low for vertebral fractures because the one placebo-controlled trial reporting this outcome was imprecise and was not designed to detect vertebral fractures as rigorously as trials of the other medications. We graded the strength of evidence for tibolone low for thromboembolic events and coronary heart disease events because estimates were based on only one trial and were imprecise. Strength of evidence for tibolone was insufficient for estrogen receptor positive breast cancer, estrogen receptor negative breast cancer, noninvasive breast cancer, all-cause death, endometrial cancer, and cataracts because these outcomes were either not reported, or the numbers of events were too low and duration of treatment and follow-up too short to provide useful estimates.

Applicability

All primary prevention trials except the Italian trial met criteria for good applicability. The Italian trial exclusively enrolled women who had undergone prior hysterectomy for reasons other than cancer28 as described in Results (Appendix C-5). For each trial, interventions, comparators, outcomes, and timing of outcome measures were appropriate. All trials were multicenter and relevant to primary care. In addition, trials were conducted in settings appropriate to clinical practice, enrolled subjects selected with broad eligibility criteria, assessed health outcomes, and had follow-up periods of several years. For these reasons, the trials provided information about effectiveness as well as efficacy of the medications.

Although inclusion criteria differed between the primary prevention trials, results for breast cancer outcomes were similar. These findings support good aggregate applicability to the target population of women without pre-existing breast cancer. Most older women with osteoporosis enrolled in the MORE and LIFT trials, and those with cardiovascular disease or risk factors enrolled in the RUTH trial, would have met risk factor eligibility criteria for the STAR and NSABP P-1 tamoxifen trials based on age. Women not well represented in the trials are

those who are younger (<55 years old), have Gail scores <1.66% or considered low risk by other criteria used by some of the trials, are nonwhite, or are from outside North America, the UK, and Europe. Also, premenopausal women were excluded from the raloxifene and tibolone trials.

Clinicians can consider the results of trials to be most applicable to patients with similar characteristics as the study populations. Specifically, tamoxifen results apply to younger pre and postmenopausal women meeting breast cancer risk criteria; tibolone results apply to older postmenopausal women with osteoporosis; and raloxifene results apply to postmenopausal women meeting breast cancer risk criteria, and older postmenopausal women with osteoporosis or cardiovascular disease and/or risk factors for cardiovascular disease.

Applicability may be more limited for other outcomes. Fracture reduction is better for women with osteoporosis than for those without it.[134] It would be expected that fracture reduction would be greater in the MORE trial of raloxifene and LIFT trial of tibolone that enrolled women with known osteoporosis. However, osteoporosis is common and often undiagnosed in the target population, as well as among women enrolled in the other primary prevention trials. Fractures were reduced in most trials, including those that did not specifically enroll women with osteoporosis, supporting the applicability of this effect.

The applicability of trials for adverse effect outcomes is more difficult to determine. Trials varied in how these outcomes were measured and reported, it is not known how risk factors for adverse effect outcomes varied among subjects, and results were not reported for specific sub-groups. Most studies were small and included highly selected participants from outside the United States. Several studies of tibolone enrolled women seeking treatment of menopausal vasomotor symptoms.

Summary of Results

Benefits (Key Questions 1 and 3)

All three medications, tamoxifen, raloxifene, and tibolone, reduced the incidence of invasive breast cancer in midlife and older women without pre-existing breast cancer by 30% to 68%. The direct comparison trial, STAR, indicated similar effects for tamoxifen and raloxifene. Indirect comparison analysis indicated that results of a placebo-controlled trial of tibolone were not significantly different than results of placebo-controlled trials of raloxifene. Reduction of invasive breast cancer continued after discontinuation of tamoxifen in trials providing follow-up data. Tamoxifen and raloxifene reduced estrogen receptor positive but not estrogen receptor negative breast cancer, and had similar effects on these subtypes when directly compared. Tamoxifen reduced noninvasive breast cancer, including ductal carcinoma *in situ* (DCIS), in the NSABP P-1 trial, but not in the other tamoxifen trials. Raloxifene did not decrease noninvasive cancer, and the STAR trial suggested that more women using raloxifene had noninvasive breast cancer than those using tamoxifen.

Tamoxifen and raloxifene reduced invasive breast cancer for all population subgroups evaluated. They had similar effects regardless of age and family history of breast cancer in the STAR trial. Tamoxifen reduced breast cancer outcomes in subgroups evaluated in placebo-controlled primary prevention trials based on age, menopausal status, estrogen use, family history of breast cancer, and history of lobular carcinoma *in situ* (LCIS) or atypical hyperplasia. In the NSABP P-1 trial, cancer rates were highest and risk reduction greatest among women in the highest Gail model risk category and among women with prior atypical hyperplasia. Raloxifene reduced breast cancer outcomes in subgroups based on age, age at menarche, parity,

age at first live birth, and body mass index. Estimates from subgroups based on prior estrogen use, family history of breast cancer, and prior hysterectomy or oophorectomy were limited by small numbers of subjects. Population subgroups have not been evaluated for tibolone.

All-cause mortality was similar for women using raloxifene compared to tamoxifen, or tamoxifen, raloxifene, or tibolone compared to placebo. Tamoxifen did not reduce breast cancer mortality compared to placebo. Tamoxifen and raloxifene had similar effects on fractures at multiple sites in the STAR trial. In placebo-controlled trials, raloxifene and tibolone reduced vertebral fractures, tamoxifen and tibolone reduced nonvertebral fractures, and tibolone reduced wrist but not hip fractures.

Harms (Key Question 2 and 3)

Tamoxifen and raloxifene increased risk for thromboembolic events compared to placebo. Raloxifene caused fewer thromboembolic events than tamoxifen in the STAR trial. Tamoxifen caused more thromboembolic events for older (>50 or 60 years) than younger women. Risk returned to normal after discontinuation of tamoxifen in the trials providing post treatment data. Tibolone did not increase risk for thromboembolic events. None of these medications increased risk for coronary heart disease events. Tibolone caused more strokes than placebo resulting in early discontinuation of the LIFT trial. Subgroup analysis indicated that risk for stroke was higher for older (>70 years) than younger women. Tamoxifen and raloxifene did not increase risk for stroke.

Raloxifene caused fewer cases of endometrial hyperplasia and was associated with fewer hysterectomies than tamoxifen in the STAR trial; differences for endometrial cancer were not significantly different. In placebo-controlled trials, tamoxifen caused more cases of endometrial cancer, and risk was higher in older than younger women. Raloxifene did not increase risk for endometrial cancer or uterine bleeding compared to placebo. Tibolone did not increase risk for endometrial cancer in clinical trials, but was associated with more cases of endometrial cancer in a large cohort study.

Raloxifene caused fewer cataracts and cataract surgeries than tamoxifen in the STAR trial and did not increase risk for cataracts or cataract surgery in placebo-controlled trials. Tamoxifen was associated with more cataract surgeries than placebo in one trial.

Medications caused several additional symptoms. In direct comparisons, women using raloxifene reported more musculoskeletal problems, dyspareunia, and weight gain, while those using tamoxifen had more gynecological problems, vasomotor symptoms, leg cramps, and bladder control symptoms. Some of the most common side effects for tamoxifen were hot flashes and other vasomotor symptoms, vaginal discharge, and other vaginal symptoms such as itching or dryness. For raloxifene, common side effects were vasomotor symptoms and leg cramps. Tibolone increased vaginal bleeding, but in contrast to the SERMs, it reduced the number and severity of hot flashes and reduced risk for colon cancer.

Patient Choice, Concordance, Adherence, and Persistence (Key Question 4)

Evidence about patient treatment choice, concordance, adherence, and persistence to treatment was lacking. Comparisons of adherence and persistence rates across medications in primary prevention trials were limited because few trials reported treatment duration, completion rates, or other measures of adherence and persistence. Also, trials were designed for different treatment purposes. From the few trials reporting data about discontinuation, rates for tamoxifen or raloxifene were generally higher than placebo, but differences between treatment and placebo

groups were low (≤2% for adverse events and ≤4% for nonprotocol specified events). No data were available for tibolone.

Regarding treatment choice, small descriptive studies indicate that women make decisions to use tamoxifen to reduce breast cancer risk based on their concern for adverse effects as well as their risk for breast cancer. They weigh their physicians' recommendations highly when deciding whether to take tamoxifen. Similar data for raloxifene and tibolone are lacking. No studies about how women choose among multiple risk reducing medications have been published.

Risk Assessment (Key Question 5)

Research on risk assessment to identify women who could benefit from medications to reduce breast cancer risk focuses on nine risk stratification models evaluated for use in clinical settings. Models consider multiple risk factors for breast cancer, and some are derived from the original Gail model. Risk stratification models demonstrate high calibration, with the expected number of breast cancer cases in a study population closely matching the number of breast cancer cases observed. All models have low discriminatory accuracy in predicting the probability of breast cancer in an individual. Most models perform only slightly better than age alone as a risk predictor. Models that include breast density, postmenopausal hormone use, and a more extensive family history show promise in improving the predictive risk. A Gail score of ≥ 1.66% has been used as a risk threshold in primary prevention trials and in U.S. Food and Drug Administration approval of tamoxifen and raloxifene for reducing risk for primary breast cancer. However, this threshold has low discriminatory accuracy in predicting the probability of breast cancer in an individual. Most women age 60 and older without other risk factors would meet this threshold by age alone.

Clinical Implications and Limitations

Based on our meta-analysis of placebo-controlled primary prevention trials, the calculated number needed to treat (NNT) to prevent one case of invasive breast cancer assuming 5 years of use is similar for all three medications: 142 (95% CI 84, 280) for tamoxifen, 112 (71, 236) for raloxifene, and 105 (58, 302) for tibolone (Tables 13, 14, 15). The STAR trial indicates similar results for tamoxifen and raloxifene (Table 16). Women and clinicians may interpret these findings as beneficial and consider use of these medications as a promising approach to reducing risk for breast cancer. In the United States, the current choices are raloxifene and tamoxifen, both also capable of reducing risk for fractures.

Although raloxifene and tamoxifen demonstrate these potential benefits, they are also capable of increasing risks for serious and potentially life threatening adverse events. Thromboembolic events are the most common serious complication of both medications, more so with tamoxifen than raloxifene (Table 16). Risk was increased by 60% to 90% in the placebo-controlled primary prevention trials that enrolled women with no prior history of thromboembolic events. Clinicians considering these medications will need to be vigilant in assessing prior history and risk factors for thromboembolic events in treatment candidates. Tamoxifen's effects on endometrial cancer, endometrial hyperplasia, and hysterectomy are also significant. These problems could be avoided if its use were limited to women with prior hysterectomies. However, since tamoxifen is the only medication approved for use in premenopausal women with or without hysterectomies, close monitoring of adverse uterine effects would be required for some users. Raloxifene and tamoxifen are also capable of causing

adverse effects that could impact quality of life such as hot flashes, vaginal symptoms, and musculoskeletal symptoms.

Women need to understand their own risks of death as a result of breast cancer and the unwanted effects of risk reducing medications before using them. The decision to use these medications would ideally occur after an accurate assessment of a woman's individual risks for breast cancer and adverse effects. Those at highest risk for breast cancer would be most likely to benefit. However, methods to identify candidates for risk reducing medications have low discriminatory accuracy. Average risk women age 60 and older meet the Gail model eligibility threshold of 1.66% 5-year risk for breast cancer. Women and clinicians have few clinical tools to work with when making decisions about using these medications.

This review is limited by potential biases. These include publication bias and biases resulting from our selection criteria, such as using English-only publications. Trials may not have been truly blinded because side effects of active medications may have lead to differential ascertainment of outcomes. Active surveillance ended with completion of therapy in most trials and important long-term outcomes may have been underreported. For some tamoxifen trials, participants randomized to placebo switched to active medications following closure of the trial, compromising long-term tracking of outcomes. All efficacy trials were powered to detect statistical differences in breast cancer incidence not adverse outcomes. Risks for some adverse outcomes may be underestimated because of lack of statistical power. Underestimation of adverse outcomes may also relate to inadequate ascertainment. For example, rates of cataracts and cataract surgery in the NSABP P-1 trial are substantially higher than rates in the other trials most likely because the trial enlisted a more aggressive detection method.

These issues highlight the limitations of the comparative effectiveness review as well as limitations of research in this area. Although many factors influence the decision to use medications to reduce risk of breast cancer, they are outside the scope of this comparative effectiveness review. However, these need to be considered when applying the results of this review to health policy, insurance coverage, or patient decisions. Research is lacking in many essential areas including optimal doses, duration of use, persistence of effects after treatment, and outcomes in population subgroups. Data are lacking for nonwhite women, premenopausal women, and women with co-morbidities or taking additional medications for other indications.

Future Research

Although several essential questions have been addressed by current studies, many more remain. More research is needed on tibolone's role in reducing risk for breast cancer and its harms. Although tibolone is not currently approved for use in the United States, it is widely used elsewhere and may be approved in the future. To avoid increasing risk for stroke, future trials of tibolone will need to focus on younger women. Future trials could confirm results of the LIFT trial and compare tibolone's efficacy in head-to-head trials with other medications. More research is needed to further evaluate findings from other studies of tibolone and determine their relevance to women using it for breast cancer risk reduction. For example, a recent multi-center trial of 3,148 breast cancer patients with vasomotor symptoms was stopped early because women using tibolone had higher breast cancer recurrence rates compared with placebo (HR 1.40;1.14, 1.70).[100,135] The Tibolone Histology of the Endometrium and Breast Endpoints Study (THEBES) comparing tibolone and continuous combined conjugated equine estrogen plus medroxyprogesterone acetate indicated that tibolone did not cause endometrial hyperplasia or carcinoma in postmenopausal women and had a more favorable vaginal bleeding profile.[99]

Trials of other emerging medications to reduce breast cancer risk, such as aromatase inhibitors and retinoids, will be needed as these are developed. Well designed and powered head-to-head trials could contribute much needed information on outcomes, duration and timing of treatment, and identification of optimal candidates. Controlled trials of lifestyle modification interventions to reduce risk for breast cancer, such as weight loss and exercise, should also be explored. These interventions could be incorporated into comparative trials that also include medications.

While the efficacy of tamoxifen, raloxifene, and tibolone has been demonstrated for women in randomized controlled trials, it is not clear which women in clinical practice would optimally benefit from risk reducing medications. Inclusion criteria for three of the placebo-controlled tamoxifen trials (NSABP P-1, IBIS, Royal Marsden) and STAR included an assessment of risk for breast cancer, and only women reaching a specified threshold were enrolled. However, for the other raloxifene and tibolone trials, no breast cancer risk assessment was performed and women of all risk groups were included. Despite these differences, trials of all the medications demonstrated efficacy in reducing invasive breast cancer. Our further analysis by various population subgroups, such as by age, menopausal status, and others, also indicated no major differences, suggesting that everyone would benefit. Future research to determine the optimal candidates for these medications would help focus risk reducing efforts. Applying these findings to clinical selection criteria would improve identification of candidates in practice settings.

In addition to improving our understanding of which women are optimal candidates, research is needed to further evaluate clinical risk instruments to identify high-risk women who are most likely to benefit from risk reducing interventions. Current research indicates that prediction models that include breast density offer marginal improvement in diagnostic accuracy. Addition of other factors such as diet, alcohol use, physical activity, smoking status, and height offer little improvement in diagnostic accuracy. The use of previously acknowledged risk factors, such as prior postmenopausal hormone therapy, needs to be reconsidered as new research indicating no associations with breast cancer are reported.[136] New models need to build on research findings from older models, and research needs to expand beyond diagnostic accuracy studies. Models need to be evaluated in relevant clinical settings and populations to

determine their effectiveness in identifying high-risk women for clinical decision-making. Effective models should also be validated in various racial and ethnic populations, among non-English speakers, and across multiple age groups. This work should include research regarding optimal methods for communicating risks and benefits to women.[137]

The results of trials indicate that adverse effects differ between medications and may drive decisions for risk reducing medications as much or more than benefits. Further research to more clearly identify characteristics of individuals experiencing specific adverse effects would guide physicians and patients to regimens that cause the least harm. Strategies could be tested that optimize benefits and minimize harms. For example, the effects of adding aspirin in conjunction with tamoxifen or raloxifene could improve the benefit/harm balance for women by reducing risks of thromboembolic adverse events, stroke,[138,139] and possibly breast cancer itself.[140] Further analysis of data from the MORE and RUTH trials could address this question because a large proportion of subjects were using aspirin in these trials. Future trials could evaluate the benefits and harms of using tamoxifen or raloxifene with an anticoagulant such as warfarin, heparin, or low molecular weight heparin.

Primary prevention trials need to be continually evaluated for long-term and unanticipated outcomes. For example, tamoxifen users in the NSABP P-1 trial who developed estrogen receptor negative breast cancer had shorter times to diagnosis and were more likely to be detected by routine mammograms than placebo users who developed estrogen receptor negative breast cancer.[141] Additional research to assess the use of raloxifene since its recent FDA approval for reducing risk for breast cancer will also be useful.

Evaluating the timing of medication use may also lead to effective clinical strategies. Results of current trials suggest that breast cancer risk reduction persists after treatment while some harms diminish. It is important to understand these changes over time. Use of medication for risk reduction at younger ages (45 to 55 years) could provide better long-term benefit and short-term harm for individuals at lower risk of thromboembolism or stroke than use at older ages (>60 years). Further analysis of data from currently available trials could compare risk/benefit profiles for women of various ages and risk groups. Additional analysis could also indicate optimal treatment durations. Shortening treatment duration would reduce harms, but also could compromise efficacy.

Despite prior recommendations to identify women at high-risk for breast cancer and offer medications to reduce their risks,[142] and the availability of two SERMs for this purpose, use is believed to be low in the United States.[107] This contrasts sharply with the use of statin medications to reduce cholesterol levels and cardiovascular disease.[143] Understanding the differences and similarities in these approaches to risk reduction would be useful for clinicians. This requires research regarding the attitudes of physicians toward recommending 5 years of medication therapy to reduce risk as well as attitudes of patients regarding receptivity to this recommendation and adherence over time. Research on the physician and patient decision-making process could identify factors important for selecting use of medications to reduce breast cancer risk beyond empirical risk.

References

1. American Cancer Society. Cancer Facts and Figures 2007 [updated 2007]. Available at: http://www.cancer.org/docroot/stt/stt_0.asp.
2. American Cancer Society. 2008 Statistics [updated 2008]; Available at: http://www.cancer.org/docroot/stt/stt_0.asp.
3. Ries LAG, Melbert D, Krapcho M, et al. SEER Cancer Statistics Review, 1975-2004; 2007 [updated 2007] Available at: http://seer.cancer.gov/csr/1975_2004/.
4. Simpson JF, Wilkinson EJ. Malignant Neoplasia of the Breast: Infiltrating Carcinomas. In: Bland KI, Copeland EM, editors. The Breast: Comprehensive Management of Benign and Malignant Disorders. 3rd Ed. St Louis: Saunders; 2004.
5. Schwartz GF. Biology and Management of Lobular Carcinoma In Situ of the Breast. In: Bland KI, Copeland EM, editors. The Breast: Comprehensive Management of Benign and Malignant Disorders. 3rd Ed. St Louis: Saunders; 2004.
6. Nelson HD, Huffman LH, Fu R, et al. Genetic risk assessment and BRCA mutation testing for breast and ovarian cancer susceptibility: Systematic evidence review for the U.S. Preventive Services Task Force. Ann Intern Med 2005;143(5):362-79.
7. Martin AM, Weber BL. Genetic and hormonal risk factors in breast cancer. J Natl Cancer Inst. 2000; 92: 1126-35.
8. Cummings SR, Tice JA, Bauer S, et al. Prevention of breast cancer in postmenopausal women: approaches to estimating and reducing risk. J Natl Cancer Inst 2009;101(6):384-98.
9. Boyd NF, Guo H, Martin LJ, et al. Mammographic density and the risk and detection of breast cancer. N Engl J Med 2007;356(3):227-36.
10. Cummings SR, Ettinger B, Delmas PD, et al. The effects of tibolone in older postmenopausal women. N Engl J Med 2008;359(7):697-708.
11. Medicines and Healthcare Products Regulatory Agency. UK Public Assessment Report Tibolone (Livial): benefit-risk evaluation. Available at: http://www.mhra.gov.uk/home/groups/plp/documents/websiteresources/con2032229.pdf.
12. Vogel VG, Costantino JP, Wickerham DL, et al. Effects of tamoxifen vs raloxifene on the risk of developing invasive breast cancer and other disease outcomes: the NSABP Study of Tamoxifen and Raloxifene (STAR) P-2 trial. JAMA 2006;295(23):2727-41.
13. Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force: a review of the process. Am J Prev Med 2001;20(3 Suppl):21-35.
14. Methods Reference Guide for Effectiveness and Comparative Effectiveness Reviews, Version 1.0. Rockville, MD; Draft posted Oct. 2007. Available at: http://effectivehealthcare.ahrq.gov/repFiles/2007_10DraftMethodsGuide.pdf.
15. Lohr KN, Helfand M, Owens DK, et al. Grading the strength of a body of evidence. J Clin Epidemiol [in press].
16. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21(11):1539-58.
17. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177-88.
18. Land SR, Wickerham DL, Costantino JP, et al. Patient-reported symptoms and quality of life during treatment with tamoxifen or raloxifene for breast cancer prevention: the NSABP Study of Tamoxifen and Raloxifene (STAR) P-2 trial [erratum appears in JAMA 2007;298(9):973]. JAMA 2006;295(23):2742-51.
19. Cuzick J, Forbes J, Edwards R, et al. First results from the International Breast Cancer Intervention Study (IBIS-I): a randomised prevention trial. Lancet 2002;360(9336):817-24.

20. Cuzick J, Forbes JF, Sestak I, et al. Long-term results of tamoxifen prophylaxis for breast cancer—96-month follow-up of the randomized IBIS-I trial. J Natl Cancer Inst 2007;99(4):272-82.

21. Day R, Ganz PA, Costantino JP. Tamoxifen and depression: more evidence from the National Surgical Adjuvant Breast and Bowel Project's Breast Cancer Prevention (P-1) Randomized Study. J Natl Cancer Inst 2001;93(21):1615-23.

22. Day R. Quality of life and tamoxifen in a breast cancer prevention trial: a summary of findings from the NSABP P-1 study. National Surgical Adjuvant Breast and Bowel Project. Ann N Y Acad Sci 2001;949:143-50.

23. Fisher B, Costantino JP, Wickerham DL, et al. Tamoxifen for the prevention of breast cancer: current status of the National Surgical Adjuvant Breast and Bowel Project P-1 study. J Natl Cancer Inst 2005;97(22):1652-62.

24. Fisher B, Costantino JP, Wickerham DL, et al. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. J Natl Cancer Inst 1998;90(18):1371-88.

25. Powles T, Eeles R, Ashley S, et al. Interim analysis of the incidence of breast cancer in the Royal Marsden Hospital tamoxifen randomised chemoprevention trial. Lancet 1998;352(9122):98-101.

26. Powles TJ, Ashley S, Tidy A, et al. Twenty-year follow-up of the Royal Marsden randomized, double-blinded tamoxifen breast cancer prevention trial. J Natl Cancer Inst 2007;99(4):283-90.

27. Decensi A, Maisonneuve P, Rotmensz N, et al. Effect of tamoxifen on venous thromboembolic events in a breast cancer prevention trial. Circulation 2005;111(5):650-6.

28. Veronesi A, Pizzichetta MA, Ferlante MA, et al. Tamoxifen as adjuvant after surgery for breast cancer and tamoxifen or placebo as chemoprevention in healthy women: different compliance with treatment. Tumori 1998;84(3):372-5.

29. Veronesi U, Maisonneuve P, Rotmensz N, et al. Tamoxifen for the prevention of breast cancer: late results of the Italian Randomized Tamoxifen Prevention Trial among women with hysterectomy [see comment]. J Natl Cancer Inst 2007;99(9):727-37.

30. Veronesi U, Maisonneuve P, Rotmensz N, et al. Italian randomized trial among women with hysterectomy: tamoxifen and hormone-dependent breast cancer in high-risk women. J Natl Cancer Inst 2003;95(2):160-5.

31. Barrett-Connor E, Cauley JA, Kulkarni PM, et al. Risk-benefit profile for raloxifene: 4-year data From the Multiple Outcomes of Raloxifene Evaluation (MORE) randomized trial. J Bone Miner Res 2004;19(8):1270-5.

32. Barrett-Connor E, Grady D, Sashegyi A, et al. Raloxifene and cardiovascular events in osteoporotic postmenopausal women: four-year results from the MORE (Multiple Outcomes of Raloxifene Evaluation) randomized trial. JAMA 2002;287(7):847-57.

33. Cauley JA, Norton L, Lippman ME, et al. Continued breast cancer risk reduction in postmenopausal women treated with raloxifene: 4-year results from the MORE trial. Breast Cancer Res Treat 2001;65(2):125-34.

34. Cummings SR, Eckert S, Krueger KA, et al. The effect of raloxifene on risk of breast cancer in postmenopausal women: results from the MORE randomized trial. Multiple Outcomes of Raloxifene Evaluation. JAMA 1999;281(23):2189-97.

35. Delmas PD, Ensrud KE, Adachi JD, et al. Efficacy of raloxifene on vertebral fracture risk reduction in postmenopausal women with osteoporosis: four-year results from a randomized clinical trial. J Clin Endocrinol Metab 2002;87(8):3609-17.

36. Delmas PD, Genant HK, Crans GG, et al. Severity of prevalent vertebral fractures and the risk of subsequent vertebral and nonvertebral fractures: results from the MORE trial. Bone 2003;33(4):522-32.

37. Duvernoy CS, Kulkarni PM, Dowsett SA, et al. Vascular events in the Multiple Outcomes of Raloxifene Evaluation (MORE) trial: incidence, patient characteristics, and effect of raloxifene. Menopause 2005;12(4):444-52.

38. Ettinger B, Black DM, Mitlak BH, et al. Reduction of vertebral fracture risk in postmenopausal women with osteoporosis treated with raloxifene: results from a 3-year randomized clinical trial. JAMA 1999;282(7):637-45.

39. Grady D, Ettinger B, Moscarelli E, et al. Safety and adverse effects associated with raloxifene: multiple outcomes of raloxifene evaluation. Obstet Gynecol 2004;104(4):837-44.

40. Johnell O, Cauley JA, Kulkarni PM, et al. Raloxifene reduces risk of vertebral fractures and breast cancer in postmenopausal women regardless of prior hormone therapy. J Fam Pract 2004;53(10):789-96.

41. Keech CA, Sashegyi A, Barrett-Connor E. Year-by-year analysis of cardiovascular events in the Multiple Outcomes of Raloxifene Evaluation (MORE) trial. Curr Med Res Opin 2005;21(1):135-40.

42. Lippman ME, Cummings SR, Disch DP, et al. Effect of raloxifene on the incidence of invasive breast cancer in postmenopausal women with osteoporosis categorized by breast cancer risk [see comment]. Clin Cancer Res 2006;12(17):5242-7.

43. Martino S, Disch D, Dowsett SA, et al. Safety assessment of raloxifene over eight years in a clinical trial setting. Curr Med Res Opin 2005;21(9):1441-52.

44. Silverman SL, Delmas PD, Kulkarni PM, et al. Comparison of fracture, cardiovascular event, and breast cancer rates at 3 years in postmenopausal women with osteoporosis. J Am Geriatr Soc 2004;52(9):1543-8.

45. Siris ES, Harris ST, Eastell R, et al. Skeletal effects of raloxifene after 8 years: results from the continuing outcomes relevant to Evista (CORE) study. J Bone Miner Res 2005;20(9):1514-24.

46. Barrett-Connor E, Mosca L, Collins P, et al. Effects of raloxifene on cardiovascular events and breast cancer in postmenopausal women. N Engl J Med 2006;355(2):125-37.

47. Grady D, Cauley JA, Geiger MJ, et al. Reduced incidence of invasive breast cancer with raloxifene among women at increased coronary risk. J Natl Cancer Inst 2008;100(12):854-61.

48. Gail MH. The estimation and use of absolute risk for weighing the risks and benefits of selective estrogen receptor modulators for preventing breast cancer. Ann N Y Acad Sci 2001;949:286-91.

49. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 1989;81(24):1879-86.

50. Veronesi U, Maisonneuve P, Costa A, et al. Prevention of breast cancer with tamoxifen: preliminary findings from the Italian randomised trial among hysterectomised women. Italian Tamoxifen Prevention Study. Lancet 1998;352(9122):93-7.

51. Martino S, Cauley JA, Barrett-Connor E, et al. Continuing outcomes relevant to Evista: breast cancer incidence in postmenopausal osteoporotic women in a randomized trial of raloxifene. J Natl Cancer Inst 2004;96(23):1751-61.

52. Vogel VG. Recent results from clinical trials using SERMs to reduce the risk of breast cancer. Ann N Y Acad Sci 2006;1089:127-42.

53. Martino S, Costantino J, McNabb M, et al. The role of selective estrogen receptor modulators in the prevention of breast cancer: comparison of the clinical trials. Oncologist 2004;9(2):116-25.

54. Ensrud K, Genazzani AR, Geiger MJ, et al. Effect of raloxifene on cardiovascular adverse events in postmenopausal women with osteoporosis. Am J Cardiol 2006;97(4):520-7.

55. Brisson J, Brisson B, Cote G, et al. Tamoxifen and mammographic breast densities. Cancer Epidemiol Biomarkers Prev 2000;9(9):911-5.

56. Bruno S, Maisonneuve P, Castellana P, et al. Incidence and risk factors for non-alcoholic steatohepatitis: prospective study of 5408 women enrolled in Italian tamoxifen chemoprevention trial. BMJ 2005;330(7497):932.

57. Chalas E, Costantino JP, Wickerham DL, et al. Benign gynecologic conditions among participants in the Breast Cancer Prevention Trial. Am J Obstet Gynecol 2005;192(4):1230-7.

58. Cuzick J, Warwick J, Pinney E, et al. Tamoxifen and breast density in women at increased risk of breast cancer. J Natl Cancer Inst 2004;96(8):621-8.

59. Day R, Ganz PA, Costantino JP, et al. Health-related quality of life and tamoxifen in breast cancer prevention: a report from the National Surgical Adjuvant Breast and Bowel Project P-1 Study. J Clin Oncol 1999;17(9):2659-69.

60. Powles TJ, Hardy JR, Ashley SE, et al. Chemoprevention of breast cancer. Breast Cancer Res Treat 1989;14(1):23-31.

61. Powles TJ, Hardy JR, Ashley SE, et al. A pilot trial to evaluate the acute toxicity and feasibility of tamoxifen for prevention of breast cancer. Br J Cancer 1989;60(1):126-31.

62. Powles TJ, Jones AL, Ashley SE, et al. The Royal Marsden Hospital pilot tamoxifen chemoprevention trial. Breast Cancer Res Treat 1994;31(1):73-82.

63. Powles TJ, Tillyer CR, Jones AL, et al. Prevention of breast cancer with tamoxifen—an update on the Royal Marsden Hospital pilot programme. Eur J Cancer 1990;26(6):680-4.

64. Reis SE, Costantino JP, Wickerham DL, et al. Cardiovascular effects of tamoxifen in women with and without heart disease: breast cancer prevention trial. J Natl Cancer Inst 2001;93(1):16-21.

65. Tan-Chiu E, Wang J, Costantino JP, et al. Effects of tamoxifen on benign breast disease in women at high risk for breast cancer. J Natl Cancer Inst 2003;95(4):302-7.

66. Veronesi U, Costa A. Prevention of breast cancer with tamoxifen: The Italian study in hysterectomized women. Breast 1995;4(4):267-72.

67. Veronesi U, Maisonneuve P, Sacchini V, et al. Tamoxifen for breast cancer among hysterectomised women. Lancet 2002;359(9312):1122-4.

68. Vogel VG. Follow-up of the breast cancer prevention trial and the future of breast cancer prevention efforts. Clin Cancer Res 2001;7(12 Suppl):4413s-8s.

69. Vogel VG, Costantino JP, Wickerham DL, et al. National surgical adjuvant breast and bowel project update: prevention trials and endocrine therapy of ductal carcinoma in situ. Clin Cancer Res 2003;9(1 Pt 2):495S-501S.

70. Vogel VG, Costantino JP, Wickerham DL, et al. The study of Tamoxifen and Raloxifene: preliminary enrollment data from a randomized breast cancer risk reduction trial. Breast Cancer Res Treat 2001;69(3):225.

71. Vogel VG, Costantino JP, Wickerham DL, et al. The study of tamoxifen and raloxifene: preliminary enrollment data from a randomized breast cancer risk reduction trial. Clin Breast Cancer 2002;3(2):153-9.

72. Ensrud KE, Stock JL, Barrett-Connor E, et al. Effects of raloxifene on fracture risk in postmenopausal women: the Raloxifene Use for the Heart Trial. J Bone Miner Res 2008;23(1):112-20.

73. Cohen FJ, Watts S, Shah A, et al. Uterine effects of 3-year raloxifene therapy in postmenopausal women younger than age 60. Obstet Gynecol 2000;95(1):104-10.

74. Delmas PD, Bjarnason NH, Mitlak BH, et al. Effects of raloxifene on bone mineral density, serum cholesterol concentrations, and uterine endometrium in postmenopausal women. N Engl J Med 1997;337(23):1641-7.

75. Freedman M, San Martin J, O'Gorman J, et al. Digitized mammography: a clinical trial of postmenopausal women randomly assigned to receive raloxifene, estrogen, or placebo. J Natl Cancer Inst 2001;93(1):51-6.

76. Goldstein SR, Johnson S, Watts NB, et al. Incidence of urinary incontinence in postmenopausal women treated with raloxifene or estrogen. Menopause 2005;12(2):160-4.

77. Johnston Jr CC, Bjarnason NH, Cohen FJ, et al. Long-term effects of raloxifene on bone mineral density, bone turnover, and serum lipid levels in early postmenopausal women: three-year data from 2 double-blind, randomized, placebo-controlled trials. Arch Intern Med 2000;160(22):3444-50.

78. Jolly EE, Bjarnason NH, Neven P, et al. Prevention of osteoporosis and uterine effects in postmenopausal women taking raloxifene for 5 years. Menopause 2003;10(4):337-44.

79. Lufkin EG, Whitaker MD, Nickelsen T, et al. Treatment of established postmenopausal osteoporosis with raloxifene: a randomized trial. J Bone Miner Res 1998;13(11):1747-54.

80. McClung MR, Siris E, Cummings S, et al. Prevention of bone loss in postmenopausal women treated with lasofoxifene compared with raloxifene. Menopause 2006;13(3):377-86.

81. Meunier PJ, Vignot E, Garnero P, et al. Treatment of postmenopausal women with osteoporosis or low bone density with raloxifene. Raloxifene Study Group [erratum appears in Osteoporos Int 1999;10(5):433]. Osteoporos Int 1999;10(4):330-6.

82. Morii H, Ohashi Y, Taketani Y, et al. Effect of raloxifene on bone mineral density and biochemical markers of bone turnover in Japanese postmenopausal women with osteoporosis: results from a randomized placebo-controlled trial. Osteoporos Int 2003;14(10):793-800.

83. Nickelsen T, Lufkin EG, Riggs BL, et al. Raloxifene hydrochloride, a selective estrogen receptor modulator: safety assessment of effects on cognitive function and mood in postmenopausal women. Psychoneuroendocrinology 1999;24(1):115-28.

84. Palacios S, Farias ML, Luebbert H, et al. Raloxifene is not associated with biologically relevant changes in hot flushes in postmenopausal women for whom therapy is appropriate. Am J Obstet Gynecol 2004;191(1):121-31.

85. Walsh BW, Kuller LH, Wild RA, et al. Effects of raloxifene on serum lipids and coagulation factors in healthy postmenopausal women. JAMA 1998;279(18):1445-51.

86. Christodoulakos GE, Botsis DS, Lambrinoudaki IV, et al. A 5-year study on the effect of hormone therapy, tibolone and raloxifene on vaginal bleeding and endometrial thickness. Maturitas 2006;53(4):413-23.

87. Ettinger B, Kenemans P, Johnson SR, et al. Endometrial effects of tibolone in elderly, osteoporotic women. Obstet Gynecol 2008;112(3):653-9.

88. Bots ML, Evans GW, Riley W, et al. The Osteoporosis Prevention and Arterial effects of tiboLone (OPAL) study: design and baseline characteristics. Control Clin Trials 2003;24(6):752-75.

89. Bots ML, Evans G, Riley W, et al. The osteoporosis prevention and arterial effects of tibolone (OPAL) study: study design and baseline characteristics. Control Clin Trials 2001;22(Suppl).

90. Langer RD, Landgren BM, Rymer J, et al. Effects of tibolone and continuous combined conjugated equine estrogen/medroxyprogesterone acetate on the endometrium and vaginal bleeding: results of the OPAL study. Am J Obstet Gynecol 2006;195(5):1320-7.

91. Landgren MB, Bennink HJ, Helmond FA, et al. Dose-response analysis of effects of tibolone on climacteric symptoms. BJOG 2002;109(10):1109-14.

92. Gallagher JC, Baylink DJ, Freeman R, et al. Prevention of bone loss with tibolone in postmenopausal women: results of two randomized, double-blind, placebo-controlled, dose-finding studies. J Clin Endocrinol Metab 2001;86(10):4717-26.

93. Swanson SG, Drosman S, Helmond FA, et al. Tibolone for the treatment of moderate to severe vasomotor symptoms and genital atrophy in postmenopausal women: a multicenter, randomized, double-blind, placebo-controlled study. Menopause 2006;13(6):917-25.

94. Hudita D, Posea C, Ceausu I, et al. Efficacy and safety of oral tibolone 1.25 or 2.5 mg/day vs. placebo in postmenopausal women. Eur Rev Med Pharmacol Sci 2003;7(5):117-25.

95. Lundstrom E, Christow A, Kersemaekers W, et al. Effects of tibolone and continuous combined hormone replacement therapy on mammographic breast density. Am J Obstet Gynecol 2002;186(4):717-22.

96. Onalan G, Onalan R, Selam B, et al. Mood scores in relation to hormone replacement therapies during menopause: a prospective randomized trial. Tohoku J Exp Med 2005;207(3):223-31.

97. Beral V, Bull D, Reeves G, et al. Endometrial cancer and hormone-replacement therapy in the Million Women Study. Lancet 2005;365(9470):1543-51.

98. Beral V, Million Women Study Collaborators. Breast cancer and hormone-replacement therapy in the Million Women Study [erratum appears in Lancet. 2003 Oct 4;362(9390):1160]. Lancet 2003;362(9382):419-27.

99. Archer DF, Hendrix S, Gallagher JC, et al. Endometrial effects of tibolone. J Clin Endocrinol Metab 2007;92(3):911-8.

100. Kenemans P, Bundred NJ, Foidart JM, et al. Safety and efficacy of tibolone in breast-cancer patients with vasomotor symptoms: a double-blind, randomised, non-inferiority trial. Lancet Oncol 2009;10(2):135-46.

101. Beral V, Green J, Reeves G, et al. Fatal stroke in postmenopausal users of tibolone and hormonal therapy. 2006; Available at: http://www.bmj.com/cgi/eletters/332/7542/667#143473.

102. de Vries CS, Bromley SE, Thomas H, et al. Tibolone and endometrial cancer: a cohort and nested case-control study in the UK. Drug Saf 2005;28(3):241-9.

103. Opatrny L, Dell'Aniello S, Assouline S, et al. Hormone replacement therapy use and variations in the risk of breast cancer. BJOG 2008;115(2):169-75.

104. Stahlberg C, Pedersen AT, Lynge E, et al. Increased risk of breast cancer following different regimens of hormone replacement therapy frequently used in Europe. Int J Cancer 2004;109(5):721-7.

105. Velthuis-Te Wierik EJM, Hendricks PT, Martinez C. Preferential prescribing of tibolone and combined estrogen plus progestogen therapy in postmenopausal women. Menopause 2007;14(3 Pt 1):518-27.

106. Abramson N, Costantino JP, Garber JE, et al. Effect of Factor V Leiden and prothrombin G20210-->A mutations on thromboembolic risk in the national surgical adjuvant breast and bowel project breast cancer prevention trial. J Natl Cancer Inst 2006;98(13):904-10.

107. Armstrong K, Quistberg DA, Micco E, et al. Prescription of tamoxifen for breast cancer prevention by primary care physicians. Arch Intern Med. 2006;166(20):2260-5.

108. Bastian LA, Lipkus IM, Kuchibhatla MN, et al. Women's interest in chemoprevention for breast cancer. Arch Intern Med 2001;161(13):1639-44.

109. Berning B, van Kuijk C, Bennink HJ, et al. Absent correlation between vaginal bleeding and oestradiol levels or endometrial morphology during tibolone use in early postmenopausal women. Maturitas 2000;35(1):81-8.

110. Bober SL, Hoke LA, Duda RB, et al. Decision-making about tamoxifen in women at high risk for breast cancer: clinical and psychological factors. J Clin Oncol 2004;22(24):4951-7.

111. McKay A, Martin W, Latosinsky S. How should we inform women at higher risk of breast cancer about tamoxifen? An approach with a decision guide. Breast Cancer Res Treat 2005;94(2):153-9.

112. Melnikow J, Paterniti D, Azari R, et al. Preferences of Women Evaluating Risks of Tamoxifen (POWER) study of preferences for tamoxifen for breast cancer risk reduction. Cancer 2005;103(10):1996-2005.

113. Port ER, Montgomery LL, Heerdt AS, et al. Patient reluctance toward tamoxifen use for breast cancer primary prevention. Ann Surg Oncol 2001;8(7):580-5.

114. Taylor R, Taguchi K. Tamoxifen for breast cancer chemoprevention: low uptake by high-risk women after evaluation of a breast lump. Ann Fam Med 2005;3(3):242-7.

115. Yeomans Kinney A, Vernon SW, Shui W, et al. Validation of a model predicting enrollment status in a chemoprevention trial for breast cancer. Cancer Epidemiol Biomarkers Prev 1998;7(7):591-5.

116. Cramer JA, Roy A, Burrell A, et al. Medication compliance and persistence: terminology and definitions. Value Health 2008;11(1):44-7.

117. Blenkinsopp B, Britten, Feely, et al. Achieving shared goals in medicine taking. A working party report. London: Royal Pharmaceutical Society of Great Britain and Merck Sharp and Dohme; 1997.

118. Boyle P, Mezzetti M, La Vecchia C, et al. Contribution of three components to individual cancer risk predicting breast cancer risk in Italy. Eur J Cancer Prev 2004;13(3):183-91.

119. Colditz GA, Rosner B. Cumulative risk of breast cancer to age 70 years according to risk factor status: data from the Nurses' Health Study. Am J Epidemiol 2000;152(10):950-64.

120. Colditz GA, Rosner BA, Chen WY, et al. Risk factors for breast cancer according to estrogen and progesterone receptor status. J Natl Cancer Inst 2004;96(3):218-28.

121. Decarli A, Calza S, Masala G, et al. Gail model for prediction of absolute risk of invasive breast cancer: independent evaluation in the Florence-European Prospective Investigation Into Cancer and Nutrition cohort. J Natl Cancer Inst 2006;98(23):1686-93.

122. Rockhill B, Spiegelman D, Byrne C, et al. Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. J Natl Cancer Inst 2001;93(5):358-66.

123. Tyrer J, Duffy SW, Cuzick J, et al. A breast cancer prediction model incorporating familial and personal risk factors [erratum appears in Stat Med 2005;24(1):156]. Stat Med 2004;23(7):1111-30.

124. Costantino JP, Gail MH, Pee D, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. J Natl Cancer Inst 1999;91(18):1541-8.

125. Chlebowski RT, Anderson GL, Lane DS, et al. Predicting risk of breast cancer in postmenopausal women by hormone receptor status [see comment]. J Natl Cancer Inst 2007;99(22):1695-705.

126. Gail MH, Costantino JP, Pee D, et al. Projecting individualized absolute invasive breast cancer risk in African American women [erratum appears in J Natl Cancer Inst 2008;100(5):373]. J Natl Cancer Inst 2007;99(23):1782-92.

127. Adams-Campbell LL, Makambi KH, Palmer JR, et al. Diagnostic accuracy of the Gail model in the Black Women's Health Study. Breast J 2007;13(4):332-6.

128. Chen J, Pee D, Ayyagari R, et al. Projecting absolute invasive breast cancer risk in white women with a model that includes mammographic density. J Natl Cancer Inst 2006;98(17):1215-26.

129. Barlow WE, White E, Ballard-Barbash R, et al. Prospective breast cancer risk prediction model for women undergoing screening mammography. J Natl Cancer Inst 2006;98(17):1204-14.

130. Tice JA, Cummings SR, Smith-Bindman R, et al. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. Ann Intern Med 2008;148(5):337-47.

131. Rockhill B, Byrne C, Rosner B, et al. Breast cancer risk prediction with a log-incidence model: evaluation of accuracy. J Clin Epidemiol 2003;56(9):856-61.

132. Amir E, Evans DG, Shenton A, et al. Evaluation of breast cancer risk assessment packages in the family history evaluation and screening programme. J Med Genet 2003;40(11):807-14.

133. Hosmer D, Lemeshow S. Applied Logistic Regression. New York: John Wiley & Sons; 2000.

134. Nelson H, Haney B, Chou R, et al. Screening for Osteoporosis: Systematic Review to Update the 2002 US Preventive Services Task Force Recommendation (Prepared by the Oregon Health & Science University Evidence-based Practice Center under Contract No 290-97-0018) Rockville, MD: Agency for Healthcare Research and Quality; in press.

135. Kenemans P, Kubista E, Foidart JM, et al. Safety of tibolone in the treatment of vasomotor symptoms in breast cancer patients—design and baseline data 'LIBERATE' trial [erratum appears in Breast 2008;17(2):214-5 Note: Egberts J [added]; van Os S [added]; Planellas J [added]]. Breast 2007;16 Suppl 2:S182-9.

136. Chlebowski RT, Kuller LH, Prentice RL, et al. Breast cancer after use of estrogen plus progestin in postmenopausal women. N Engl J Med 2009;360(6):573-87.

137. Kurz-Milcke E, Gigerenzer G, Martignon L. Transparency in risk communication: graphical and analog tools. Ann N Y Acad Sci 2008;1128:18-28.

138. Wolff T, Miller T, Ko S. Aspirin for the primary prevention of cardivascular events: An update of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 2009;150(6):405-10.

139. U.S. Preventive Services Task Force. Aspirin for the prevention of cardiovascular disease: recommendation statement. Ann Intern Med 2009;150(6):396-404.

140. Gierach GL, Lacey Jr JV, Schatzkin A, et al. Nonsteroidal anti-inflammatory drugs and breast cancer risk in the National Institutes of Health-AARP Diet and Health Study. Breast Cancer Res 2008;10(2):R38.

141. Shen Y, Costantino JP, Qin J. Tamoxifen chemoprevention treatment and time to first diagnosis of estrogen receptor-negative breast cancer. J Natl Cancer Inst 2008;100(20):1448-53.

142. U.S. Preventive Services Task Force. Chemoprevention of breast cancer: recommendations and rationale. Ann Intern Med. 2002;137(1):56-8.

143. RXList.com. Top 200 Drugs by Prescriptions Dispensed. Available at: http://www.rxlist.com/script/main/hp.asp. Accessed March 30, 2009.

Abbreviations

Acronym/ Abbreviation	Definition
AHRQ	Agency for Healthcare Research and Quality
BCDDP	Breast Cancer Detection and Demonstration Project
BCPCG	Breast Cancer Prevention Collaborative Group
BCSC	Breast Cancer Surveillance Consortium
BMD	Bone Mineral Density
BMI	Body Mass Index
BRCA1	Breast Cancer Susceptibility Gene 1
BRCA2	Breast Cancer Susceptibility Gene 2
CEE	Conjugated Equine Estrogens
CER	Comparative Effectiveness Reviews
CHD	Coronary Heart Disease
CI	Confidence Interval
CORE	Continuing Outcomes Relevant to Evista Trial
CT	Computed Tomography
DCIS	Ductal carcinoma *in situ*
DVT	Deep Vein Thrombosis
E/O	Expected/Observed Ratio
EPC	Evidence-based Practice Center
ER-	Estrogen Receptor Negative
ER+	Estrogen Receptor Positive
FDA	Food and Drug Administration
GRADE	Grades of Recommendation Assessment, Development and Evaluation
HR	Hazard Ratio
IBIS-1	International Breast Cancer Intervention Study
IT1-GM	Italian-1 Gail Model
IT-GM	Italian Gail Model
LCIS	Lobular Carcinoma *in situ*
LIFT	Long-Term Intervention on Fractures with Tibolone Trial
MeSH	Medical Subject Headings
MI	Myocardial Infarction
MORE	Multiple Outcomes of Raloxifene Evaluation Trial
MPA	Medroxyprogesterone Acetate
MWS	Million Women's Study
NNT	Number needed to treat
NR	Not reported
NSABP P-1	National Surgical Adjuvant Breast and Bowel Project
OPAL	Osteoporosis Prevention and Arterial effects of tiboLone Trial
PE	Pulmonary Embolism
PICOTS	Population, Intervention, Comparator, Outcomes, Timing of outcomes measurement and Setting
RCT	Randomized Controlled Trial
RH	Relative Hazard
RR	Risk Ratio
RUTH	Raloxifene Use for the Heart Trial
SCHIP	State Children's Health Insurance Program
SEER	Surveillance, Epidemiology, and End Results
SERMs	Selective Estrogen Receptor Modulators
SRC	Scientific Resource Center
STAR	Study of Tamoxifen and Raloxifene
STEAR	Selective Tissue Estrogenic Activity Regulator
THEBES	The Tibolone Histology of the Endometrium and Breast Endpoints Study
TIA	Transient Ischemic Attack
UK	United Kingdom
US	United States of America
USPSTF	United States Preventive Services Task Force
VTE	Venous Thrombotic Event
WHI	Women’s Health Initiative

Tables

Table 1. Medications included in Comparative Effectiveness Review

Medication	Type	Trade name(s)	Half-life or other relevant pharmacokinetic feature	Labeled indications	Dosing for primary prevention of breast cancer	Dose adjustments for special populations
Tamoxifen citrate	Selective estrogen receptor modulator (SERM)	Nolvadex Soltamox	Elimination half-life 5 to 7 days	Reducing the incidence of breast cancer among women at high risk for breast cancer. Adjuvant treatment of breast cancer. Treatment of metastatic breast cancer in men and women. Treatment of intraductal breast cancer *in situ* after surgery and radiation to reduce the risk of invasive breast cancer.	20 mg per day for 5 years	None noted
Raloxifene	Selective estrogen receptor modulator (SERM)	Evista	Elimination half life 27.7 to 32.5 hours	Reducing the risk of breast cancer among postmenopausal women at high risk. Reducing the incidence of breast cancer among postmenopausal women with osteoporosis. Treatment of osteoporosis among postmenopausal women. Prevention of post menopausal osteoporosis.	60 mg per day; optimal duration not described	None noted
Tibolone*	Selective tissue estrogenic activity regulator (STEAR)	Livial	Elimination half-life 10 hours	Prevention of postmenopausal osteoporosis. Treatment of vasomotor menopausal symptoms.	2.5 mg per day for vasomotor symptoms; 1.25 mg per day for median 2.8 years in LIFT breast cancer prevention trial	None noted

*Not currently approved by the U.S. Food & Drug Administration.

Abbreviations: LIFT, Long-Term Intervention on Fractures with Tibolone.

Mechanisms of action (http://www.cancer.gov/Templates/drugdictionary):

Tamoxifen competitively inhibits the binding of estradiol to estrogen receptors, thereby preventing the receptor from binding to the estrogen-response element on DNA. The result is a reduction in DNA synthesis and cellular response to estrogen. In addition, tamoxifen up-regulates the production of transforming growth factor B (TGFb), a factor that inhibits tumor cell growth, and down-regulates insulin-like growth factor 1 (IGF-1), a factor that stimulates breast cancer cell growth. Tamoxifen also down-regulates protein kinase C (PKC) expression in a dose-dependant manner, inhibiting signal transduction and producing an antiproliferative effect in tumors such as malignant glioma and other cancers that overexpress PKC.

Raloxifene binds to estrogen receptors (ER) as a mixed estrogen agonist/antagonist; it displays both an ER-alpha-selective partial agonist/antagonist effect and a pure ER-beta-selective antagonist effect. This agent functions as an estrogen agonist in some tissues (bones, lipid metabolism) and as an estrogen antagonist in others (endometrium and breasts), with the potential for producing some of estrogen's beneficial effects without producing its adverse effects.

Tibolone is a synthetic anabolic steroid with estrogenic, androgenic and progestagenic activities. The 3alpha- and 3beta-hydroxy metabolites of tibilone activate estrogenic receptors (ERs) in bone and vaginal tissue leading to a decrease in bone turnover, and decreased vaginal dryness, respectively; derived from the 3beta-hydroxy metabolite, its delta4-isomer activates androgenic receptors (ARs) in the brain and liver and progestogenic receptors (PRs) in endometrial tissue, affecting sexual function, lipid metabolism, and endometrial function, respectively. In breast and endometrial tissue, tibolone metabolites inhibit sulfatase, preventing the conversion of circulating estrone sulfate and estradiol sulfate to estrone and estradiol, respectively; estrogen-mediated effects in the breast and uterus are thus reduced.

Table 2. Randomized controlled trials of primary prevention for breast cancer

Trial	Included Publications	N	Subjects	Primary Outcome	Duration	Study Quality/ Applicability
Tamoxifen (20 mg/day) vs. Raloxifene (60 mg/day)						
Study of Tamoxifen and Raloxifene (STAR)	Vogel, 2006[12]; Land, 2006[18]	9872 tamoxifen 9875 raloxifene	Postmenopausal women with a 5-year predicted breast cancer risk of ≥1.66% based on the modified Gail model.* Age ≥35 years, mean age 58.5 years; 94% white; 52% post hysterectomy; none using estrogen. US based with nearly 200 clinical sites in North America.	Invasive breast cancer	Mean follow-up 3.9 years with mean exposure 3.1 to 3.2 years.	Good/Good
Tamoxifen (20 mg/day) vs. Placebo						
National Surgical Adjuvant Breast and Bowel Project P-1 Study (NSABP-1)	Fisher, 1998[24]; Fisher, 2005[23]; Day, 2001a[21]; Day, 2001b[22]	6576 tamoxifen 6599 placebo	Age ≥60 years or age 35 to 59 years with a 5-year predicted risk of breast cancer ≥1.66% based on the modified Gail model,* or a history of lobular carcinoma *in situ.* 39% of women were <50 years old; 97% white; 38% post hysterectomy; none using estrogen. US based with multiple clinical sites in North America.	Invasive and noninvasive breast cancer	Median follow-up 4.6 years with median exposure 4.0 years for initial results; median follow-up 7.0 years for long-term results.	Good/Good
International Breast Cancer Intervention Study (IBIS-I)	Cuzick, 2002[19]; Cuzick, 2007[20]	3573 tamoxifen3566 placebo	Increased breast cancer risk based on family history and other factors.† Age 35 to 70 years, mean age 50.8 years; 35% post hysterectomy; 40% using estrogen. UK, Australia, NZ, Europe.	Invasive and noninvasive breast cancer	Median follow-up 4.2 years for initial results; 8.0 years for long-term results.	Fair/Good
Royal Marsden Hospital Trial	Powles, 1998[25]; Powles, 2007[26]	1238 tamoxifen 1233 placebo	Family history of breast cancer.‡ Age 30 to 70 years; median age 47 years; 15% of tamoxifen and 27% of placebo group using estrogen at the beginning of trial; UK.	Invasive breast cancer	Median follow-up 5.8 years for initial results; 13.2 years for long-term results.	Fair/Good

Trial	Included Publications	N	Subjects	Primary Outcome	Duration	Study Quality/ Applicability
Italian Tamoxifen Prevention Study	Veronesi, 1998[28]; Veronesi, 2003[30]; Veronesi, 2007[29]; Decensi, 2005[27]	2700 tamoxifen 2708 placebo	Women with hysterectomy for reasons other than cancer. Age 35 to 70 years; median age 51 years; 14% using estrogen; Italy based with 55 clinical centers in Europe and South America.	Breast cancer incidence and mortality	Median follow-up 3.8 years for intial results; 11.2 years follow-up and 4.0 years exposure for long-term results.	Fair/Fair
Raloxifene (60 or 120 mg/day) vs. Placebo						
Multiple Outcomes of Raloxifene Evaluation (MORE) and Continuing Outcomes Relevant to Evista (CORE)	Ettinger, 1999[38]; Cummings, 1999[34]; Cauley, 2001[33]; Barrett-Connor, 2002[32]; Delmas, 2002[35]; Delmas, 2003[36]; Grady, 2004[39]; Barrett-Connor, 2004[31]; Silverman, 2004[44]; Johnell, 2004[40]; Martino, 2005[43]; Duvernoy, 2005[37]; Keech, 2005[41]; Siris, 2005[45]; Lippman, 2006[42]	MORE: 5129 raloxifene (60 or 120 mg/day) 2576 placebo CORE: 2725 raloxifene (60 mg/day) 1286 placebo	Postmenopausal women with osteoporosis.§ Age 31 to 80 years; median age 66.9 years; 96% white; 23% post hysterectomy; none using systemic estrogen. US based with 180 clinical centers in 25 countries. CORE is comprised of a subset of MORE participants to further examine raloxifene's effect on breast cancer incidence.	MORE: Incident radiographic vertebral fractures and verified clinical nonvertebral fractures excluding pathologic, traumatic, and nonosteoporosis-related fractures (i.e., face, skull, finger, toe). CORE: Breast cancer.	Follow-up time varies with publications and outcomes; MORE results reported at 3 and 4 years and CORE at 4 and 8 years (combines the MORE and CORE data).	Good/Good

Trial	Included Publications	N	Subjects	Primary Outcome	Duration	Study Quality/ Applicability
Raloxifene Use for the Heart (RUTH)	Barrett-Connor, 2006[46]; Grady, 2008[47]	5044 raloxifene (60 mg/day) 5057 placebo	Postmenopausal women with coronary heart disease or multiple risk factors for heart disease.‖ Age ≥55 years; median age 67.5 years; 84% white; 23% post hysterectomy; none on estrogen; US based with 177 clinical sites in 26 countries.	Coronary events (death from coronary causes, nonfatal myocardial infarction, acute coronary syndrome) and invasive breast cancer.	Median duration 5.6 years; median exposure 5.1 years.	Good/Good
Tibolone (1.25 mg/day) vs. Placebo						
Long-Term Intervention on Fractures with Tibolone (LIFT)	Cummings, 2008[10]	2267 tibolone 2267 placebo	Women with bone mineral density T-score ≤-2.5 at the hip or spine or T-score ≤-2.0 and radiologic evidence of a vertebral fracture. Age 60 to 85 years; mean 68 years; 22% post hysterectomy; none on estrogen. US based with 80 clinical sites in 22 countries.	Incident radiographic vertebral fractures and verified clinical nonvertebral fractures excluding pathologic, traumatic, and nonosteoporosis-related fractures (i.e., face, skull, finger, toe).	Median exposure 2.8 years	Good/Good

* STAR & NSABP-1: The Gail model includes age, number of first-degree relatives with breast cancer, nulliparity or age at first live birth, number of benign breast biopsies, pathologic diagnosis of atypical hyperplasia, and age at menarche. The original model was further modified to predict expected rates of invasive breast cancer only (not invasive and noninvasive as originally designed) and to allow for race-specific determinations of risk.

† IBIS: 2-fold relative risk for ages 45 to 70, 4-fold relative risk for ages 40 to 44, 10-fold relative risk for ages 35 to 39 based on family history criteria. All criteria permit entry to trial at age 45 years.

1. First-degree relative who developed breast cancer at or before age 50.
2. First-degree relative with bilateral breast cancer (permits entry from age 40; if relative diagnosed before age 40, permits entry at age 35).
3. Two or more first-degree or second-degree relatives with breast cancer (permits entry from age 40 if both developed breast cancer before age 50, permits entry at age 35 if both relatives are first-degree and both developed breast cancer before age 50).
4. Benign breast biopsy and first-degree relative with breast cancer.
5. Lobular carcinoma in situ (permits entry from age 35).
6. Atypical hyperplasia (permits entry from age 40).
7. Nulliparous and a first-degree relative who developed breast cancer.
8. Risk equivalent (strong family history, not fitting specific categories, but judged to be at higher risk than eligibility category by the study chairman).

‡ Family history criteria for Royal Marsden Hospital Trial:

1. One first-degree relative under 50 years old with breast cancer, or
2. One first-degree relative with bilateral breast cancer, or
3. One affected first-degree of any age plus another affected first-degree or second-degree relative
4. Benign breast biopsy and a first-degree relative with breast cancer

§ MORE:
Study Group 1: Femoral neck or lumbar spine bone mineral density (BMD) T-score <-2.5.
Study Group 2: Low BMD and one or more moderate or severe vertebral fractures or 2 or more milder vertebral fractures (20% to 25% reduction in height); or at least 2 moderate fractures (25% to 40% reduction from expected vertebral height), regardless of BMD.
‖ Participants were required to have a cardiovascular risk score of 4 or more according to a point system: established coronary heart disease (4 points), arterial disease of the leg (4 points), at least 70 years old (2 points), diabetes mellitus (3 points), cigarette smoking (1 point), hypertension (1 point), and hyperlipidemia (1 point).

Table 3. Major health outcomes reported in primary prevention trials

Outcomes	Placebo-controlled Trials Reporting Outcomes	Included in Meta-analysis	Reported in STAR Trial
Benefits			
All breast cancer	NSABP-1, Marsden, IBIS, Italian, MORE, RUTH	x	NR
Invasive breast cancer	NSABP-1, Marsden, IBIS, Italian, MORE, RUTH, LIFT	x	x
ER+ breast cancer	NSABP-1, Marsden, IBIS, Italian, MORE, RUTH	x	x
ER- breast cancer	NSABP-1, Marsden, IBIS, Italian, MORE, RUTH	x	x
Noninvasive breast cancer	NSABP-1, Marsden, IBIS, Italian, MORE, RUTH	x	x
DCIS	Marsden, IBIS, MORE, LIFT		x
Breast cancer mortality	NSABP-1, Marsden, IBIS, Italian, MORE	x	NR
All-cause mortality	NSABP-1, Marsden, IBIS, Italian, MORE, RUTH, LIFT	x	x
All fractures	Marsden, IBIS	x	NR
Hip, wrist, spine fractures	NSABP-1, IBIS	x	x
Vertebral fractures	NSABP-1, MORE, RUTH, LIFT	x	x
Nonvertebral fractures	NSABP-1, MORE, RUTH, LIFT	x	NR
Hip fractures	NSABP-1, MORE, LIFT		x
Wrist fractures	NSABP-1, MORE, LIFT		x
Harms			
Thromboembolic events	NSABP-1, Marsden, IBIS, Italian, MORE, RUTH, LIFT	x	x
Deep vein thrombosis	NSABP-1, Italian, MORE, RUTH	x	x
Pulmonary embolus	NSABP-1, Italian, MORE, RUTH	x	x
Superficial phlebitis	Italian, IBIS	x	NR
Coronary heart events	NSABP-1, Marsden, IBIS, Italian, MORE, RUTH, LIFT	x	x
Myocardial infarction	NSABP-1, IBIS, Italian	x	x
Stroke	NSABP-1, Marsden, IBIS, Italian, RUTH, MORE, LIFT	x	x
Transient ischemic attack	NSABP-1, IBIS, Italian, LIFT	x	x
Endometrial cancer	NSABP-1, Marsden, IBIS, MORE, RUTH, LIFT	x	x
Cataracts	NSABP-1, Marsden, IBIS, MORE, RUTH	x	x

Abbreviations: NSABP-1, National Surgical Adjuvant Brest and Bowel Project P-1 Study; IBIS, International Breast Cancer Intervention Study; MORE, Multiple Outcomes of Raloxifene Evaluation; RUTH, Raloxifene Use for the Heart; LIFT, Long-Term Intervention on Fractures with Tibolone; STAR, Study of Tamoxifen and Raloxifene; NR, not reported; ER+, estrogen receptor positive; ER-, estrogen receptor negative; DCIS, ductal carcinoma *in situ.*

Table 4. Results of primary prevention trials—benefits

Outcomes	Tamoxifen vs Raloxifene STAR Trial RR (95% CI)*	Tamoxifen vs Placebo Meta-analysis RR (95% CI)	Trials	Raloxifene vs Placebo Meta-analysis RR (95% CI)	Trials	Tibolone vs Placebo LIFT Trial RH (95% CI)
Breast cancer						
All breast cancer	Not reported	0.72 (0.61, 0.86)	4	0.53 (0.34, 0.84)	2	Not reported
Invasive	1.02 (0.82, 1.28)	0.70 (0.59, 0.82)	4	0.44 (0.27, 0.71)	2	0.32 (0.13, 0.80)
Estrogen-receptor positive	0.93 (0.72, 1.24)	0.58 (0.42, 0.79)	4	0.33 (0.18, 0.61)	2	Not reported
Estrogen-receptor negative	1.15 (0.75, 1.77)	1.19 (0.92, 1.55)	4	1.25 (0.67, 2.31)	2	Not reported
Noninvasive	1.40 (0.98, 2.00)† 1.46 (0.90, 2.41)‡	0.85 (0.54, 1.35)§	4	1.47 (0.75, 2.91)‖	2	Not reported¶
Death						
Breast cancer	Not reported**	1.07 (0.66, 1.74)	4	Not reported††		Not reported
All-cause	0.94 (0.71, 1.26)	1.07 (0.90, 1.27)	4	0.91 (0.81, 1.02)	2	Not reported‡‡
Fractures						
Hip, wrist, vertebral	0.92 (0.69, 1.22)	0.81 (0.55, 1.18)	2	Not reported		Not reported
Vertebral	0.98 (0.65, 1.46)	0.75 (0.48, 1.15)	1§§	0.61 (0.54, 0.69)	2	0.55 (0.41, 0.74)
Nonvertebral	Not reported	0.66 (0.45, 0.98)	1§§	0.97 (0.87, 1.09)	2	0.74 (0.58, 0.93)
Hip	0.88 (0.48, 1.60)	0.68 (0.39, 1.18)	1§§	0.97 (0.62, 1.52)	1‖‖	0.72 (0.32, 1.63)
Wrist	0.85 (0.46, 1.53)	0.69 (0.37, 1.25)	1§§	0.83 (0.66, 1.05)	1‖‖	0.54 (0.35, 0.82)

* Risk ratio for women in the raloxifene group compared with those in the tamoxifen group.
† RR for total noninvasive breast cancer; includes DCIS, LCIS, and mixed.
‡ RR for DCIS only.
§ Combines noninvasive and DCIS in meta-analysis.
‖ Meta-analysis did not include DCIS. Cases of DCIS reported in MORE: 9 raloxifene, 5 placebo.
¶ RH Not reported. Cases of DCIS reported: 1 in each group.
** Cases reported: 4 tamoxifen, 2 raloxifene.
†† Cases reported: 1 raloxifene, 0 placebo.
‡‡ Cases reported: 26 tibolone, 28 placebo (p=0.89).
§§ NSABP-1 (Fisher, 2005).
‖‖ MORE (Delmas, 2004).

Abbreviations: STAR, Study of Tamoxifen and Raloxifene; LIFT, Long-Term Intervention on Fractures with Tibolone; RR, risk ratio; RH, relative hazard; CI, confidence interval; DCIS, ductal carcinoma in situ.

Table 5. Results of primary prevention trials—harms

	Tamoxifen vs Raloxifene	Tamoxifen vs Placebo		Raloxifene vs Placebo		Tibolone vs Placebo
Outcomes	**STAR Trial RR (95% CI)***	**Meta-analysis RR (95% CI)**	**Trials**	**Meta-analysis RR (95% CI)**	**Trials**	**LIFT Trial RH (95% CI)**
Thromboembolic events	0.70 (0.54, 0.91)	1.93 (1.41, 2.64)	4	1.60 (1.15, 2.23)	2	0.57 (0.19, 1.69)
Deep vein thrombosis	0.74 (0.53, 1.03)	1.45 (0.89, 2.37)	2	1.91 (0.87, 4.23)	2	Not reported
Pulmonary embolus	0.64 (0.41, 1.00)	2.69 (1.12, 6.47)	2	2.19 (0.97, 4.97)	2	Not reported
Superficial phlebitis	Not reported	2.14 (1.29, 3.56)	2	Not reported	2	Not reported
Cardiovascular events						
Coronary heart disease events	1.10 (0.85, 1.43)	1.00 (0.79, 1.27)	4	0.95 (0.84, 1.06)	2	1.37 (0.77, 2,45)
Myocardial infarction	0.77 (0.48, 1.20)	1.01 (0.63, 1.64)	3	Not reported	2	Not reported
Stroke	0.96 (0.64, 1.43)	1.36 (0.89, 2.08)	4	0.96 (0.67, 1.38)	2	2.19 (1.14, 4.23)
Transient ischemic attack	1.21 (0.79, 1.88)	0.77 (0.46, 1.30)	3	Not reported	2	Not reported
Endometrial cancer	0.62 (0.35, 1.08)	2.13 (1.36, 3.32)	3	1.14 (0.65, 1.98)	2	Not reported†
Cataracts	0.79 (0.68, 0.92)	1.25 (0.93, 1.67)	3	0.93 (0.84, 1.04)	2	Not reported

* Risk ratio for women in the raloxifene group compared with those in the tamoxifen group.

† RH not reported. Cases reported: 4 tibolone, 0 placebo.

Abbreviations: STAR, Study of Tamoxifen and Raloxifene; LIFT, Long-Term Intervention on Fractures with Tibolone; RR, risk ratio; RH, relative hazard; CI, confidence interval; DCIS, ductal carcinoma in situ.

Table 6. Additional outcomes reported in the primary prevention trials*

	Tamoxifen Trials				Raloxifene Trials		Tibolone Trial
	Royal Marsden Powles, 2007[26]	Italian Veronesi, 2007[29]	IBIS Cuzick, 2007[20]	NSABP Fisher 1998[24]; Day, 1999[59]; Day, 2001[21]	MORE Cauley, 2001[31]	RUTH Barrrett-Connor, 2006[46]	LIFT Cummings 2008[10] Ettinger 2008[87]
Atrial fibrillation						o	
Leg cramps	+						
Pain/joint pain	o			o	+	+	
Anxiety		o				o	
Depression/mood change	o			o			
Sexual symptoms		o					
Vaginal symptoms							+ †
Gynecologic cancers						o ‡	o §
Endometrial fluid					+		+
Breast symptoms	o		o				
GI disorders	o	o					- ‖
Gall bladder disease						+	
Sleep disturbance	o						
Headaches	o		o				
Peripheral edema		o			+	+	
Weight gain	+	o					
Influenza syndrome					+	o	
Hot flashes	+		+	+	+	+	
Malaise/lethargy	o						

* Statistically significant differences between treatment and placebo groups are indicated by: + = outcome increased in treatment groups; - = outcome decreased in treatment groups; O = no differences between treatment and placebo groups for the outcome; blank cells = outcome not reported.

† Vaginal bleeding, discharge, and infection were all statistically significantly increased in LIFT.

‡ Ovarian cancer was not significantly different in the raloxifene and placebo groups.

§ Cervical cancer was not significantly different in the tibolone and placebo groups.

‖ Colon cancer and gastroenteritis were significantly lower in the tibolone group.

Table 7. Compliance outcomes for trials of tamoxifen, raloxifene, and tibolone

	Tamoxifen vs. Raloxifene Trial		Tamoxifen Trials							
	STAR Vogel, 2006[52]		NSABP P-1 Fisher, 1998*[24]		IBIS-I Cuzick, 2007[20]		Royal Marsden Powles, 2007[26]		Italian Veronesi, 2007[29]	
Outcomes	**Raloxifene**	**Tamoxifen**	**Tamoxifen**	**Placebo**	**Tamoxifen**	**Placebo**	**Tamoxifen**	**Placebo**	**Tamoxifen**	**Placebo**
Adherence	NR	NR	NR	NR	NR	NR	8% less than placebo (p=0.002)		NR	NR
Duration of treatment	3.2 years††	3.1 years	NR	NR	NR	NR	NR	NR	47.4 months	48.9 months
Completion of treatment	NR	NR	NR	NR	5 years 2287/3579 (63.9%)	5 years 2574/3575 (72%)	NR	NR	5 years: 1615/2700 (59.8%)	5 years: 1674/2708 (60.8%)
Discontinuation due to protocol specified event (major events)	NR	NR	NR	NR	NR	NR	NR	NR	206/2700 (7.6%)	188/2708 (6.9%)
Discontinuation due to non-protocol specified event	NR	NR	23.7%	19.7%	NR	NR	NR	NR	721/2700 (26.7%)	686/2708 (25.3%)
Discontinuation due to "adverse event"	NR	NR	NR	NR	NR	NR	NR†	NR†	NR	NR

	Raloxifene Trials							
	RUTH Barrett-Connor, 2006[46]		MORE Cummings, 1999[34]		Cohen, 2000[73]		Goldstein, 2005[76]	
Outcomes	**Raloxifene**	**Placebo**	**Raloxifene**	**Placebo**	**Raloxifene**	**Placebo**	**Raloxifene**	**Placebo**
Adherence	70% vs 71% (p=0.62)		92%		NR	NR	86% to 90%‡	
Duration of treatment	Median exposure 5.05 years		NR	NR	NR	NR	Mean duration 2.3 years§	
Completion of treatment	80% vs 79% (p=0.02)‖		NR	NR	NR	NR	60%‡	

Raloxifene Trials

Outcomes	RUTH Barrett-Connor, 2006[46] Raloxifene	Placebo	MORE Cummings, 1999[34] Raloxifene	Placebo	Cohen, 2000[73] Raloxifene	Placebo	Goldstein, 2005[76] Raloxifene	Placebo
Discontinuation due to protocol specified event (major events)	NR	NR	NR	NR	NR	NR	NR	NR
Discontinuation due to non-protocol specified event	NR	NR	NR	NR	NR	NR	NR	NR
Discontinuation due to "adverse event"	22% vs 20% (p=0.01)		33/5129 (0.6%) due to hot flashes	2/2576 (0.1%) due to hot flashes (p<.001)	13.9%		17.6%‡	

Raloxifene Trials

Outcomes	Lufkin, 1998[79] Raloxifene	Placebo	McClung, 2006[80] Raloxifene	Placebo	Meunier, 1999[81] Raloxifene	Placebo	Palacios, 2004[84] Raloxifene	Placebo
Adherence	NR	NR	NR	NR	NR	NR	91.6%	87.4%
Duration of treatment	NR	NR	702 to 706 days#		NR	NR	NR	NR
Completion of treatment	130/143 (91%)¶		67%#		109/129 (84.5%)**		89.2%	87.4%
Discontinuation due to protocol specified event (major events)	1/143		NR	NR	NR	NR	NR	NR
Discontinuation due to non-protocol specified event	2/143		NR	NR	NR	NR	NR	NR
Discontinuation due to "adverse event"	8/143 (5.6%)		22/163 (13.5%)	12/83 (14.5%)	7/87 (8%)	4/40 (10%)	non-significant difference between groups	

	Tibolone Trials					
	LIFT Cummings, 2008[10]		Berning, 2000[109]		OPAL Langer, 2006[90]	
Outcomes	**Tibolone**	**Placebo**	**Tibolone**	**Placebo**	**Tibolone**	**Placebo**
Adherence	91% received at least 80% of study drug		NR	NR	NR	NR
Duration of treatment	Median treatment duration 34 months		NR	NR	NR	NR
Completion of treatment	NR	NR	89%		69%	70%
Discontinuation due to protocol specified event (major events)	NR	NR	NR	NR	NR	NR
Discontinuation due to non-protocol specified event	NR	NR	NR	NR	NR	NR
Discontinuation due to "adverse event"	Significantly higher rate in tibolone group than placebo.		5/71 (7%)	4/23 (17.4%)	NR	NR

* Later reports of the NSABP P-1 trial do not report compliance data, therefore the Fisher 1998 paper is used here.
† An earlier report of the Royal Marsden trial prior to completing enrollment stated that the most frequent side effects leading to discontinuation were hot flushes and gynecologic problems (Powles 1998).
‡ Includes conjugated equine estrogen group.
§ 3- year study period.
‖ RUTH trial reported completed "study" rather than "treatment."
¶ 1- year study period.
# Includes lasofoxifene data.
** 2- year study period.
††At the time of this publication, patients were continuing on therapy.

Abbreviations: NSABP-1, National Surgical Adjuvant Brest and Bowel Project P-1 Study; IBIS, International Breast Cancer Intervention Study; MORE, Multiple Outcomes of Raloxifene Evaluation; RUTH, Raloxifene Use for the Heart; LIFT, Long-Term Intervention on Fractures with Tibolone; STAR, Study of Tamoxifen and Raloxifene; NR, not reported.

Table 8. Descriptive studies of treatment decisions for medications to reduce risk of breast cancer

Study/ Method	Population	Response Rate	N	Accept Treatment	Decline Treatment	Undecided	Included Outcomes
Armstrong, 2006[107] Physican survey by mail	Primary care physicians, including family medicine, obstetrics and gynecology, and general internal medicine.	47.2%	350	96/350 prescribed tamoxifen within prior 12 months	NA	NA	Prescription rates of tamoxifen and reasons for prescribing tamoxifen.
Bastian, 2001[108] Survey by phone	Women age 40 to 55 years enrolled in a Blue Cross/Blue Shield Personal Care Plan; 8% had Gail score of at least 1.66% .	1287/2165 (59%)*	1287	NR	NR	NR	Interest in medications to reduce risk of breast cancer.
Bober, 2004[110] Survey in person with telephone follow-up	Women age ≥35 years with a 5-year risk of developing breast cancer ≥1.7%; mean age 52 years.	129/158 (82%)	129	Tamoxifen prescription: 37/129 (29%) STAR trial: 35/129 (27%)†	31/129 (24%)†	26/129 (20%)†	Decision making about medications at two and four month follow-up times.
McKay, 2005[111] Survey with decision guide by mail	Women at higher risk of breast cancer; mean age 52 years; mean Gail score 3.7% (1.7 to 9.4%).	30/39 (77%)‡	51‡	6/51 (11.8%)	38/51 (76.5%)	6/51 (11.8%)	Evaluation of decision making guide and interest in tamoxifen for breast cancer risk reduction.
Melnikow, 2005[112] Cross sectional, mixed methods interview	Women at high risk for breast cancer; 32% age 39 to 64 years, 44% 65 to 74 years, 25% ≥75 years.	255/341	255	45/255 (17.6%)	206/255 (80.7%)	NR	Attitudes and preferences for use of tamoxifen for breast cancer risk reduction.
Port, 2001[113] Education session with pre/post survey	Women at increased risk for breast cancer; mean age 52.8 years (39 to 74 years).	NR	43	2/43 (4.7%)	15/43 (34.8%)	26/43 (60.5%)	Patient interest in and acceptance of electively taking tamoxifen for breast cancer risk reduction.

Study/ Method	Population	Response Rate	N	Accept Treatment	Decline Treatment	Undecided	Included Outcomes
Taylor, 2005[114] Survey by telephone	High risk women (Gail score >1.6%) age 35 to 80 years.	88/89	89	1/48 women who discussed with physician	47/48 women who discussed with physician	NA	Interest in breast cancer risk reduction with tamoxifen.
Yeomans-Kinney, 1998[115] Survey in person	Women eligible for NSABP P-1 trial; mean age 55 years; mean Gail score 14.8%.	360/479 (75%) completed surveys; 81/360 discussed tamoxifen with their physician; 175/181 reported the physician's recommedation.	360	89/175 (51%) enrolled	86/175 (49%) did not enroll	NA	Effect of a physician's recommendation to enroll in the NSABP P-1 trial.

* After excluding ineligibles, completion Rate was 76% and refusal rate was 20%.

† 2 month follow- up data.

‡ 51 women were identified for participation and 39 agreed to participate. The 21 women who refused were included in the analysis as declining tamoxifen.

Abbreviations: NSABP-1, National Surgical Adjuvant Brest and Bowel Project P-1 Study; STAR, Study of Tamoxifen and Raloxifene; NA, not applicable; NR, not reported.

Table 9. Studies of risk stratification models

Study	Model	Population	N	Study Design	Incidence Rates for Comparison	Inclusion/Exclusion Criteria	Quality
Gail, 1989[49]	Gail (invasive breast cancer and LCIS)	BCDDP- white women with *in situ* or invasive cancer vs control between 1973-1979. Age: 35-79	2582 cases, 3146 controls	Derivation study; case-control; abstracted risk factor information from 80% of eligible cases and 83% of eligible controls; follow- up through 1998.	Determined from 243,221 white females in BCDDP registry.	10- year life expectancy, no history of breast cancer, negative mammogram within 180 days, negative clinical breast exam, no history of DCIS (LCIS ok).	Good
Costantino, 1999[124]	Gail (invasive breast cancer)	BCPT- white women between 1992-1998.	5969 women in placebo arm of BCPT; 204 incident cases	Validition study of Gail 1 and 2 comparing BCDDP, CASH, NHS, BCPT cohorts; follow-up 1 to 70 months (avg. 48.4).	Gail 1 - BCDDP rates for invasive or *in situ* cancer; GAIL 2 - SEER data for invasive cancer.	10- year life expectancy, no history of breast cancer, negative mammogram within 180 days, negative clinical breast exam, no history of DCIS, LCIS.	Good
Rockhill, 2001[122]	Gail 5-yr risk (invasive breast cancer)	NHS - white women age 45-71 in 1992; study duration from 1992 to 1997.	1354 cases of 82,109 cohort	Validation study; prospective cohort; follow- up 60 months.	Not reported	White women with complete risk factor data.	Good

Study	Model	Population	N	Study Design	Incidence Rates for Comparison	Inclusion/Exclusion Criteria	Quality
DeCarli 2006[121]	Italian-Gail Model;* Italian 1-Gail Model† (all breast cancer)	Derivation: Italian multicenter case-control study of diet and breast cancer; Florence - European Prospective Investigation into cancer and nutrition; 1991-1994. Derivation: Age of cases 23-74, mean 55; controls 20-74, mean 56. Validation: Age 35-64.	Derivation: 2569 cases with 2588 controls; Validation: 194 cases in 10,031 cohort	Derivation - case control; Validation - cases in cohort	Florence Cancer Registry	Women admitted with breast cancer diagnosed within 1 year of the study interview with no prior history of cancer.No admissions for gynecological, neoplastic, hormonal diseases or those related to increased risk of breast cancer in controls.	Good

Study	Model	Population	N	Study Design	Incidence Rates for Comparison	Inclusion/Exclusion Criteria	Quality
Boyle, 2004[118]	Italian-Gail Model;*	Derivation: Italian multicenter case-control study of diet and breast cancer,1991-1994. Derivation: Age of cases 23-74, mean 55; controls 20-74, mean 56. Validation: Italian Tamoxifen Prevention Study, 1992-1997. Validation: Age of cases 35-70, median 51.	Derivation: 2569 cases with 2588 controls; Validation: 2700 tamoxifen, 2708 placebo	Derivation- case control; Validation-cases in cohort	Regional Cancer Registry Data	Women admitted with breast cancer diagnosed within 1 year of the study interview with no prior history of cancer.No admissions for gynecological, neoplastic, hormonal diseases or those related to increased risk of breast cancer in controls.	Fair
Chlebowski, 2007[125]	Expanded and simplified models vs Gail 2; (ER+ vs ER- invasive breast cancer)	WHI age: 50-79 years, mean 63 years.	3236 cases, 363 excluded due to missing data =2873 for subgroup analysis, 2412 ER+ cases, 461 ER- cases; 144,680 control.	Derivation and validation; case-control; 5 years follow-up.	Not reported	Unlikely to move or die within 3 years; no history of breast cancer or mastectomy.	Good
Gail, 2007[126]	Gail AA (invasive breast cancer)	CARE: African American women; age 35-64; 1994 to 1998 and 1993 to 1998.	1607 cases; 1647 control; women matched for 5-year age group, location, and race; 14,059 from WHI.	Derivation - CARE Validation - WHI case-control; WHI Follow up 7.57 years.	SEER	First primary incident invasive breast cancer in African American women age 35-64 years; must have complete data available.	Good

Study	Model	Population	N	Study Design	Incidence Rates for Comparison	Inclusion/Exclusion Criteria	Quality
Adams-Campbell, 2007[127]	Gail AA (invasive breast cancer)	BWHS: African American women; age ≥ 35 years from 1995 to 2003.	725 cases; 725 age-matched controls; ≥ 35 years.	Validation; nested case-control; follow- up 8 years.	SEER	Incident invasive breast cancer; must have complete data available.	Good
Chen, 2006[128]	Gail plus breast density (invasive breast cancer)	BCDDP: primarily white women age > 40 years; *in situ* or invasive cancer vs control; data collected 1973 to 1979.	Cases total 2852 (1235 with mammograpy density); age-matched controls 3146 (1656 with mammography density)	Case-control; follow- up through 1998.	SEER	Cases with missing data excluded.	Good
Barlow, 2006[129]	BCSC Barlow model (1-year risk of DCIS or invasive breast cancer)	BCSC: women without breast cancer age 35-84 years; from 1996 to 2001.	11,638 cases from 2,392,998 in cohort	Cases within cohort of women being screened with mammography; 1 year follow- up.	BSCS (compared to SEER)	DCIS or invasive breast cancer in women age 35-84 years who had prior mammogram within the last 5 years; no prior breast cancer, no breast augmentation, no prior mammogram but detected breast cancer within one year of first mammogram; if no data on menopause, excluded from subgroup analysis.	Fair to Good

Study	Model	Population	N	Study Design	Incidence Rates for Comparison	Inclusion/Exclusion Criteria	Quality
Tice, 2008[130]	BCSC Tice (invasive breast cancer)	BCSC: women without breast cancer aged 35-84 years; 71% white	1,095,484 in cohort, 14,766 cases or invasive breast cancer; 629,229 for clinical risk factor analysis; 14,766 cases.	Cases within cohort of women being screened with mammography; median follow- up of 5.3 years.	SEER (BCSC vs SEER, state tumor registries, and path databases)	women age 35 years or older with 1 prior mammogram with BI-RAD measurement in BCSC; excluded women with diagnosis of breast cancer, women diagnosed within 6 mo of index mammogram, and women with breast implants.	Good
Colditz, 2000[119]	Colditz-Rosner, Model 2	NHS: age 35-70 years; 1980 to1994.	1761 cases among 58,520.	Cases within cohort of NHS; derivation; 14 years follow- up.	Not compared	Incident invasive breast cancer; exclusions: pregnancy/offspring history discrepancies, inaccurate age of menarche, unknown age of menopause or death, missing height weight or hormone use data, hysterectomy with 1 or no ovaries removed or missing menopause data.	Good
Rockhill, 2003[131]	Colditz-Rosner, Model 2	NHS: age 45-73 1992 to 1997.	757 cases among 45,210	Cases within cohort of NHS; validation.	Not reported	Invasive breast cancer; no prior cancer, natural menopause or hysterectomy without oophorectomy, complete data.	Good
Colditz, 2004[120]	Colditz-Rosner, Model 2	NHS: age 35-79; 1980 to 2000.	2096 cases (1281 ER+/PR+, 417 ER-/PR-, 318 ER+/PR-, 80 ER-/PR+) among 66,D17145 women	Cases within cohort of NHS; validation.	Not reported	Invasive breast cancer with reported estrogen receptor status.	Good

Study	Model	Population	N	Study Design	Incidence Rates for Comparison	Inclusion/Exclusion Criteria	Quality
Tyrer, 2004[123]	Tyrer-Cuzick (invasive breast cancer)	UK national statistics of breast cancer incidence rates in general population; BRCA risk tables from UK	NR	data from other sources; derivation model	UK rates of breast cancer and positive BRCA.	NR	Fair to Good
Amir, 2003[132]	Tyrer-Cuzick (10- year risk of invasive breast cancer)	Family history clinic at University Hospital of South Manchester, high risk population; total population age 21-73, median 44; screened population age 25-73, median 46; from 987 to 2001.	64 cases among 3150 women; sub-analysis on screening population- 52 cases among 1933 cohort.	Women whose risk estimate could be derived by all the models were compared and only incident cases included.	UK - Northwest cancer registry	Complete risk data for all models being compared (Gail, Claus, Ford, Tyrer-Cuzick); excluded incomplete data.	Fair

* Italian-Gail Model varies from Gail by only 1 ordinal value on one variable

† Italian-1-Gail Model varies from Gail by classifying by categorical rather than ordinal variables

Abbreviations: BCPT, Breast Cancer Prevention Trial; NHS, Nurses' Health Study; WHI, Women's Health Initiative; BCSC, Breast Cancer Surveillance Consortium; BCDDP, Breast Cancer Detection and Demonstration Project; IMCCSDBC, Italian Multicenter Case-control Study of Diet and Breast Cancer; EPIC, European Prospective Investigation into Cancer and Nutrition; CARE, Women's Contraceptive and Reproductive Experiences; FHESP, Family History and Evaluation Screening Program- University Hospital of South Manchester; ER+, Estrogen Receptor positive; ER-, Estrogen Recptor negative; DCIS, Ductal Carcinoma *in situ*; LCIS, Lobular Carcinoma *in situ;* NR, Not reported.

Table 10. Variables included in risk stratification models

Model Study	Age	Age at Menarch	Age at 1st birth	Family History of Breast Cancer in 1st Degree Relative	Number of Breast Biopsies	History of Atypical Hyperplasia	Other Factors not included in Gail Model
Gail (invasive, DCIS, LCIS) Gail, 1989[49]	<50 ≥50	≥14 12-13 ≤12	<20 20-24 25-29 or none ≥30	0 1 ≥2	0 1 ≥2	0 ≥1	Not Applicable
Gail (invasive) Costantino, 1999[124]	<50 ≥50	≥14 12-13 ≤12	<20 20-24 25-29 or none ≥30	0 1 ≥2	0 1 ≥2	0 ≥1	None
Italian- Gail Model* DeCarli, 2006[121]	<50 ≥50	≥14 12-13 ≤12	<20 20-24 25-29 or none ≥30	0 1 ≥2	0 ≥1	0 ≥1	None
Italian- 1- Gail Model† DeCarli, 2006[121]	X‡	X	X	X	X	X	None
Gail- African American (invasive) Gail, 2007[126]	<50 ≥50	≤13 >13		0 1 ≥2	0 1 ≥2		African American race
Boyle Model Boyle, 2004[118]	<50 ≥50	≥14 12-13 ≤12	<20 20-24 25-29 or none ≥30	0 1 ≥2			Age of relative at diagnosis, Diet score, Alcohol use, BMI, HRT, Physical activity

Model Study	Age	Age at Menarch	Age at 1st birth	Family History of Breast Cancer in 1st Degree Relative	Number of Breast Biopsies	History of Atypical Hyperplasia	Other Factors not included in Gail Model
Chlebowski-Expanded (ER+ vs ER-, invasive) Chlebowski, 2007[125]	50-59 60-69 70-79	≥14 12-13 ≤12	<20 20-24 25-29 or none ≥30	0 ≥1	0 1 ≥2	Coded as unknown if prior biopsy	BMI: <25,25 to <30,≥30 kg/m^2 Menopause age Hormone Use: Estrogen only, estrogen + progesterone Duration of estrogen only use: 0, <5, 5 to <10, 10 to <15, ≥15 years Duration of combined estrogen + progesterone use: 0, <5, 5 to <10, 10 to <15, ≥15 years Race Alcohol use: ≤1 or >1 drink/day Parity: 0,1, 2, ≥3 Cumulative time breast-feeding: 0, ≤1, >1 year Smoking status: never, past, current Physical activity: <5, 5-12, 12 METS
Chlebowski-Simplified (ER+, invasive) Chlebowski, 2007[125]	<50(0) ≥50(1)			0 (0) ≥1 (1)	0 (0) 1 (1) ≥2 (2)		None
Chen (invasive) Chen, 2006[128]	<50 or ≥50	≥14 12-13 ≤12	<20 20-24 25-29 or none ≥30	0 1 ≥2	0 1 ≥2		Breast density: 0%, 1-24%, 25-49%, 50-74%, 75-100% BMI: 0 - ≤100, 101-125, 126-150, 151-175, 176-200, >200
BCSC Barlow (DCIS or invasive in premenopausal women) Barlow, 2006[129]	5-yr intervals 35-54			0 1 ≥2 unknown	no yes unknown		Breast Density: BIRADS - 0,1,2,3,4§ Hormone use

Model Study	Age	Age at Menarch	Age at 1st birth	Family History of Breast Cancer in 1st Degree Relative	Number of Breast Biopsies	History of Atypical Hyperplasia	Other Factors not included in Gail Model
BCSC Barlow (DCIS or invasive in postmenopausal women) Barlow, 2006[129]	5-yr intervals 45-84		<30 ≥30 nulliparous unknown	0 1 ≥2 unknown	0 ≥1 unknown	Prior false-positive mammogram	Breast Density: BIRADS - 0,1,2,3,4 BMI: <25, 25-29.99, 30-34.99, ≥35, unknown Menopause: Natural, surgical, unknown Hormone use: No, Yes, Unknown Race/Ethnicity: White, Asian- Pacific Islander, Black, Native, Hispanic
BCSC Tice (invasive) Tice, 2008[130]	5-yr intervals 40-74			yes or no	yes or no		Breast density: BIRADS - 1,2,3,4 Race/ethnicity: White, Asian-Pacific islander, Black, Hispanic. Native excluded due to lack of SEER data.
Colditz-Rosner Colditz, 2000[119]	X	X	X	yes no	Benign breast disease - yes or no		BMI Menopause: natural or bilateral oophorectomy, other; age at menopause Hormone use: Duration of postmenopausal estrogen, estrogen + progesterone, other; current use vs past use. Height Alcohol use Parity: 0 (0), ≥1 (1)
Tyrer-Cuzick Tyrer, 2004[123]	X	≤12 >12	≤30 >30 nulliparity	1, 2, 1 + ≥2 in family, ovarian cancer, other family history combination; age of onset of cancer; bilateral breast cancer, male breast cancer	X	X + LCIS	BMI:<21, 21-23, 23-25, 25-27, >27 Height Age at menopause

*Italian-Gail Model varies from Gail-2 by only 1 ordinal value on one variable
† Italian-1-Gail Model varies from Gail-2 by classifying by categorical rather than ordinal variables
‡X - indicates an included variable but no further data on coding
§BIRADS:0-unknown; 1-entirely fat, 2- scattered fibroglandular densities; 3- heterogeneously dense; 4 - extremely dense
Abbreviations: BCPT, Breast Cancer Prevention Trial; NHS, Nurses' Health Study; WHI, Women's Health Initiative; BCSC, Breast Cancer Surveillance Consortium; BCDDP, Breast Cancer Detection and Demonstration Project; IMCCSDBC, Italian Multicenter Case-control Study of Diet and Breast Cancer; EPIC, European Prospective Investigation into Cancer and Nutrition; CARE, Women's Contraceptive and Reproductive Experiences; FHESP, Family History and Evaluation Screening Program-University Hospital of South Manchester; ER+, Estrogen Receptor positive; ER-, Estrogen Receptor negative; DCIS, Ductal Carcinoma *in situ*; LCIS, Lobular Carcinoma *in situ*.

Table 11. Calibration (expected/observed ratio) and discriminatory accuracy of Gail Model quintiles

	Italian- Gail Model (Decarli, 2006)*[121]	Gail Model (Decarli, 2006)†[121]	Gail Model (Chlebowski, 2007)‡[125]	Gail Model (Tice, 2008)§[130]	Tice Model Tice, 2008[130] c-statistic
Gail Quintile					
1	1.09 (0.71-2.06)	0.91 (0.62-1.58)	0.629	0.99 (0.93-1.05)	0.62
2	0.78 (0.58-1.14)	0.87 (0.64- 1.28)	0.663	0.99 (0.94-1.04)	0.64
3	0.78 (0.60-1.10)	0.73 (0.56-1.02)	0.742	1.01 (0.96-1.06)	0.62
4	0.95 (0.74-1.35)	0.93 (0.71-1.31)	0.817	1.02 (0.98-1.06)	0.62
5	1.19 (0.93-1.60)	1.13 (0.88-1.54)	0.991	1.03 (0.99-1.07)	0.61
Gail Risk Category					
Low‖				1.00 (0.98-1.03)	0.65
High¶				1.03 (0.99-1.07)	0.61

* Quintile values differed across studies. Italian- Gail Model values: 1=0-1.19, 2=1.20-1.53, 3=1.54-1.88, 4=1.89-2.35, 5=2.36-8.73.

† Quintile values for Decarli calibration of the Gail Model: 1=0-1.14, 2=1.15-1.51, 3=1.52-1.87, 4=1.88-2.35, 5=2.36-6.12.

‡ Quintile values for Chlebowski calibration of the Gail Model: 1=1.09, 2=1.09-1.37, 3=1.37-1.68, 4=1.68-2.16, 5= >2.16.

§ Quintile values for the Tice calibration and discriminatory accuracy were undefined.

‖ Low Gail risk is defined as 5-year risk of <1.67%

¶ High Gail risk is defined as 5-year risk of >1.67%

Table 12. GRADE table of evidence for major health outcomes

Number of studies; number of subjects	Domains Pertaining to Strength of Evidence				Strength of Evidence and Magnitude of Effect
	Risk of Bias	Consistency	Directness	Precision	Risk Ratio (95% CI; number of trials) Number of events reduced or increased per 1000 women years assuming 5 years of use (95% CI)
Invasive breast cancer					**High for tamoxifen and raloxifene; moderate for tibolone**
Head to head RCT 1 raloxifene vs tamoxifen; 19,747	Low	Unknown (single study)	Direct	Precise	1.02 (0.82, 1.28; 1 trial)
4 tamoxifen vs placebo RCTs; 28,421	Low	No inconsistency	Indirect	Precise	0.70 (0.59, 0.82; 4 trials) 7 (4, 12) fewer than placebo
2 raloxifene vs placebo RCTs; 17,806	Low	No inconsistency	Indirect	Precise	0.44 (0.27, 0.71; 2 trials) 9 (4, 14) fewer than placebo
1 tibolone vs placebo RCT; 4,506	Low	Unknown (single study)	Indirect	Precise	0.32 (0.13, 0.80; 1 trial) 10 (3, 17) fewer than placebo
Estrogen receptor positive breast cancer					**High for tamoxifen and raloxifene; insufficient for tibolone**
Head to head RCT 1 raloxifene vs tamoxifen; 19,747	Low	Unknown (single study)	Direct	Precise	0.93 (0.72, 1.24; 1 trial)
4 tamoxifen vs placebo RCTs; 28,421	Low	No inconsistency	Indirect	Precise	0.58 (0.42, 0.79; 4 trials) 8 (3, 13) fewer than placebo
2 raloxifene vs placebo RCTs; 17,806	Low	No inconsistency	Indirect	Precise	0.33 (0.18, 0.61; 2 trials) 8 (4, 12) fewer than placebo
Estrogen receptor negative breast cancer					**Moderate for tamoxifen and raloxifene; insufficient for tibolone**
Head to head RCT 1 raloxifene vs tamoxifen; 19,747	Low	Unknown (single study)	Direct	Imprecise*	1.15 (0.75, 1.77; 1 trial)
4 tamoxifen vs placebo RCTs; 28,421	Low	No inconsistency	Indirect	Imprecise*	1.19 (0.92, 1.55; 4 trials)
2 raloxifene vs placebo RCTs; 17,806	Low	No inconsistency	Indirect	Imprecise*	1.25 (0.67, 2.31; 2 trials)
Noninvasive breast cancer					**Moderate for raloxifene; low for tamoxifen; insufficient for tibolone**
Head to head RCT 1 raloxifene vs tamoxifen; 19,747	Low	Unknown (single study)	Direct	Imprecise*	1.40 (0.98, 2.00; 1 trial)

Number of studies; number of subjects	Domains Pertaining to Strength of Evidence				Strength of Evidence and Magnitude of Effect
	Risk of Bias	Consistency	Directness	Precision	Risk Ratio (95% CI; number of trials) Number of events reduced or increased per 1000 women years assuming 5 years of use (95% CI)
4 tamoxifen vs placebo RCTs; 28,421	Low	Inconsistent†	Indirect	Imprecise*	0.85 (0.54, 1.35; 4 trials)
2 raloxifene vs placebo RCTs; 17,806	Low	No inconsistency	Indirect	Imprecise*	1.47 (0.75, 2.91; 2 trials)
All-cause death (short-term)					**High for raloxifene and tamoxifen; insufficient for tibolone**
Head to head RCT 1 raloxifene vs tamoxifen; 19,747	Low	Unknown (single study)	Direct	Precise	0.94 (0.71, 1.26; 1 trial)
4 tamoxifen vs placebo RCTs; 28,421	Low	No inconsistency	Indirect	Precise	1.07 (0.90, 1.27; 4 trials)
2 raloxifene vs placebo RCTs; 14,112	Low	No inconsistency	Indirect	Precise	0.91 (0.81, 1.02; 2 trials)
1 tibolone vs placebo RCT; 4,506	Low	Unknown (single study)	Indirect	Not estimable‡	26 deaths tibolone vs. 28 placebo; p=0.89 in LIFT trial
Vertebral fractures					**High for raloxifene; moderate for tibolone; low for tamoxifen**
Head to head RCT 1 raloxifene vs tamoxifen; 19,747	Low	Unknown (single study)	Direct	Imprecise*	0.98 (0.65, 1.46; 1 trial)
1 tamoxifen vs placebo RCT; 13,388	Low	Unknown (single study)	Indirect	Imprecise*	0.75 (0.48, 1.15; 1 trial)
2 raloxifene vs placebo RCTs; 16,929	Low	No inconsistency	Indirect	Precise	0.61 (0.54, 0.69; 2 trials) 7 (5, 9) fewer than placebo
1 tibolone vs placebo RCT; 4,146	Low	Unknown (single study)	Indirect	Precise	0.55 (0.41, 0.74; 1 trial) 44 (25, 61) fewer than placebo
Nonvertebral fractures					**High for raloxifene; moderate for tamoxifen and tibolone**
1 tamoxifen vs placebo RCT; 13,388	Low	Unknown (single study)	Indirect	Precise	0.66 (0.45, 0.98; 1 trial) 3 (0.2, 5) fewer than placebo
2 raloxifene vs placebo RCTs; 14,112	Low	No inconsistency	Indirect	Precise	0.97 (0.87, 1.09; 2 trials)
1 tibolone vs placebo RCT; 4,506	Low	Unknown (single study)	Indirect	Precise	0.74 (0.58, 0.93; 1 trial) 34 (8, 56) fewer than placebo

Number of studies; number of subjects	Domains Pertaining to Strength of Evidence				Strength of Evidence and Magnitude of Effect
	Risk of Bias	Consistency	Directness	Precision	Risk Ratio (95% CI; number of trials) Number of events reduced or increased per 1000 women years assuming 5 years of use (95% CI)
Thromboembolic events					**High for raloxifene and tamoxifen; low for tibolone**
Head to head RCT 1 raloxifene vs tamoxifen; 19,747	Low	Unknown (single study)	Direct	Precise	0.70 (0.54, 0.91; 1 trial) 6 (2, 10) more with tamoxifen
4 tamoxifen vs placebo RCTs; 28,421	Low	No Inconsistency	Indirect	Precise	1.93 (1.41, 2.64; 4 trials) 4 (2, 9) more than placebo
8 raloxifene vs placebo RCTs; 19,774	Low	No inconsistency	Indirect	Precise	1.60 (1.15, 2.23; 2 trials) 7 (2, 15) more than placebo
3 tibolone vs placebo RCT; 6,051	Low	Unknown (single study)	Indirect	Imprecise*	0.57 (0.19, 1.69; 1 trial)
Coronary Heart Disease Events					**High for raloxifene and tamoxifen; low for tibolone**
Head to head RCT 1 raloxifene vs tamoxifen; 19,747	Low	Unknown (single study)	Direct	Precise	1.10 (0.85, 1.43; 1 trial))
4 tamoxifen vs placebo RCTs; 28,421	Low	No inconsistency	Indirect	Precise	1.00 (0.79, 1.27; 4 trials)
2 raloxifene vs placebo RCTs; 17,806	Low	No inconsistency	Indirect	Precise	0.95 (0.84, 1.06; 2 trials)
2 tibolone vs. placebo RCTs; 4,902	Low	Unknown	Indirect	Imprecise*	1.37 (0.77, 2.45; 1 trial)
Stroke					**Moderate for tamoxifen, raloxifene, and tibolone**
Head to head RCT 1 raloxifene vs tamoxifen; 19,747	Low	Unknown (single study)	Direct	Imprecise*	0.96 (0.64, 1.43; 1 trial)
4 tamoxifen vs placebo RCTs; 28,421	Low	No inconsistency	Indirect	Imprecise*	1.36 (0.89, 2.08; 4 trials)
2 raloxifene vs placebo RCTs; 15,314	Low	Inconsistent§	Indirect	Precise	0.96 (0.67, 1.38; 2 trials)
1 tibolone vs placebo RCT; 4,506	Low	Unknown (single study)	Indirect	Precise	2.19 (1.14, 4.23; 1 trial) 11 (1, 36) more with tibolone

Number of studies; number of subjects	Domains Pertaining to Strength of Evidence				Strength of Evidence and Magnitude of Effect
	Risk of Bias	Consistency	Directness	Precision	Risk Ratio (95% CI; number of trials) Number of events reduced or increased per 1000 women years assuming 5 years of use (95% CI)
Endometrial cancer					**High for tamoxifen; moderate for raloxifene; insufficient for tibolone**
Head to head RCT 1 raloxifene vs tamoxifen; 19,747	Low	Unknown (single study)	Direct	Imprecise*	0.62 (0.35, 1.08; 1 trial)
3 tamoxifen vs placebo RCTs;15,401	Low	No inconsistency	Indirect	Precise	2.13 (1.3, 3.32; 3 trials) 4 (1, 10) more with tamoxifen
2 raloxifene vs placebo RCTs; 13,741	Low	No inconsistency	Indirect	Imprecise*	1.14 (0.65, 1.98; 2 trials)
2 tibolone vs placebo RCTs; 4,385	Low	Unknown	Indirect	Not estimable‡	0 cases tibolone vs. 4 placebo; p=0.06 in LIFT trial
1 tibolone observational study; 28,028	**High‖**	**Unknown**	**Indirect**	**Precise**	**1.79 (1.43, 2.25; 1 study)**
Cataracts					**High for raloxifene; low for tamoxifen; insufficient for tibolone**
Head to head RCT 1 raloxifene vs tamoxifen; 19,747	Low	Unknown (single study)	Direct	Precise	0.79 (0.68, 0.92; 1 trial) 13 (5, 21) more with tamoxifen
3 tamoxifen vs placebo RCTs; 21,857	Low	Inconsistent†	Indirect	Imprecise*	1.25 (0.93, 1.67; 3 trials)
2 raloxifene vs placebo RCTs; 17,717	Low	No inconsistency	Indirect	Precise	0.93 (0.84, 1.04; 2 trials)

*Estimates indicating no statistically significant differences between comparators with confidence intervals wider than 0.67 to 1.50 are considered imprecise because they could be compatible with different clinical conclusions.
†Results of the NSABP P-1 trial differ from results of the meta-analysis.
‡Low number of events and short duration of treatment and follow-up (2.8 years) limit this outcome measure from the LIFT trial.
§Point estimates are inconsistent and may reflect population heterogeneity between the MORE and RUTH trials for this outcome.
‖Tibolone users in this study are highly selected introducing bias for this outcome.
See appendix and text for definitions of terms used in this table.

Table 13. Estimates of number needed to treat or harm for tamoxifen

Outcomes	RR (95% CI)	Trials	Placebo Rate (SE)*	Number of Events Reduced/Increased (95% CI)†	Number Needed to Treat/Harm (95% CI)‡
Breast cancer reduced					
All breast cancer§	0.72 (0.61, 0.86)	4	5.54 (1.32)	8 (3, 15)	129 (72, 286)
Invasive	0.70 (0.59, 0.82)	4	4.70 (1.02)	7 (4, 12)	142 (84, 280)
Estrogen receptor +	0.58 (0.42, 0.79)	4	3.67 (0.78)	8 (3, 13)	130 (76, 294)
Fractures reduced					
Vertebral	0.75 (0.48, 1.15)	1			
Nonvertebral	0.66 (0.45, 0.98)	1	1.55 (0.20)	3 (0.2, 5)	380 (196, 1798)
Thromboembolic events increased‖	1.93 (1.41, 2.64)	4	0.91 (0.19)	4 (2, 9)	236 (117, 578)
Deep vein thrombosis	1.45 (0.89, 2.37)	2			
Pulmonary embolus	2.69 (1.12, 6.47)	2	0.19 (0.07)	2 (0.1, 6)	623 (127, 5405)
Stroke	1.36 (0.89, 2.08)	4			
Endometrial cancer increased	2.13 (1.36, 3.32)	3	0.75 (0.15)¶	4 (1, 10)	236 (104, 771)
Cataracts**	1.25 (0.93, 1.67)	3			

*Per 1000 women-years. Estimated from a meta-analysis of rates from the placebo groups from the same trials included in the combined RR.

†Numbers of events reduced/increased are calculated by assuming 1000 women take tamoxifen for 5 years.

‡Numbers needed to treat/harm are calculated by assuming each woman takes tamoxifen for 5 years.

§RR for noninvasive breast cancer was significantly reduced in the NSABP P-1 trial (60 vs 93 events; RR=0.63; 0.45, 0.89).

‖Includes deep vein thrombosis and pulmonary embolus.

¶Estimated from two trials that reported rates from the placebo groups (Fisher, 1998 and Cuzik, 2007).

**RR for cataracts was significantly increased in the NSABP P-1 trial (574 vs 507 events; RR=1.14; 1.01, 1.29).

Table 14. Estimates of number needed to treat or harm for raloxifene

Outcomes	RR (95% CI)	Trials	Placebo Rate (SE)*	Number of Events Reduced/Increased (95% CI)†	Number Needed to Treat/Harm (95% CI)‡
Breast cancer reduced					
All breast cancer	0.53 (0.34, 0.84)	2	3.53 (0.69)	8 (3, 14)	121 (70, 340)
Invasive	0.44 (0.27, 0.71)	2	3.19 (0.59)	9 (4, 14)	112 (71, 236)
Estrogen receptor +	0.33 (0.18, 0.61)	2	2.45 (0.42)	8 (4, 12)	122 (81, 226)
Fractures reduced					
Vertebral	0.61 (0.54, 0.69)	2	3.45 (0.35)§	7 (5, 9)	149 (115, 201)
Nonvertebral	0.97 (0.87, 1.09)	2			
Thromboembolic events increased‖	1.60 (1.15, 2.23)	2	2.34 (0.25)	7 (2, 15)	142 (66, 553)
Deep vein thrombosis	1.91 (0.87, 4.23)	2			
Pulmonary embolus	2.19 (0.97, 4.97)	2			
Stroke	0.96 (0.67, 1.38)	2			
Endometrial cancer	1.14 (0.65, 1.98)	2			
Cataracts	0.93 (0.84, 1.04)	2			

*Per 1000 women-years. Estimated from a meta-analysis of rates from the placebo groups from the same trials included in the combined RR.
†Numbers of events reduced/increased are calculated by assuming 1000 women take raloxifene for 5 years.

‡Numbers needed to treat/harm are calculated by assuming each woman takes raloxifene for 5 years.

§Estimated from the placebo group of RUTH (Barrett-Connor, 2006).

‖includes deep vein thrombosis and pulmonary embolus.

Table 15. Estimates of number needed to treat or harm for tibolone from the LIFT trial

Outcomes	RR (95% CI)	Trials	Placebo Rate (SE)*	Number of Events Reduced/Increased (95% CI)†	Number Needed to Treat/Harm (95% CI)‡
Breast cancer reduced					
All breast cancer					
Invasive	0.32 (0.13, 0.80)	1	2.80 (0.66)	10 (3, 17)	105 (58, 302)
Estrogen receptor +					
Fractures reduced					
Vertebral	0.55 (0.41, 0.74)	1	19.60 (1.75)	44 (25, 61)	23 (16, 40)
Nonvertebral	0.74 (0.58, 0.93)	1	26.30 (2.04)	34 (8, 56)	29 (17, 104)
Thromboembolic events increased§	0.57 (0.19, 1.69)	1			
Deep vein thrombosis					
Pulmonary embolus					
Stroke increased	2.19 (1.14, 4.23)	1	1.90 (0.53)	11 (1, 36)	88 (25, 584)
Endometrial cancer					
Cataracts					

*Per 1000 women-years.

†Numbers of events reduced/increased are calculated by assuming 1000 women take tibolone for 5 years.

‡Numbers needed to treat/harm are calculated by assuming each woman takes tibolone for 5 years.

§Includes deep vein thrombosis and pulmonary embolus.

Table 16. Results of STAR

Outcomes	RR (95% CI)	Raloxifene Rate*	Tamoxifen Rate*	Number of Events Reduced/Increased (95% CI)†
Breast cancer reduced				
Invasive	1.02 (0.82, 1.28)	4.41	4.30	
Estrogen receptor +	0.93 (0.72, 1.24)	2.86	3.04	
Noninvasive	1.40 (0.98, 2.00)	2.11	1.51	
Fractures reduced				
Vertebral	0.98 (0.65, 1.60)	1.35	1.39	
Hip	0.88 (0.48, 1.60)	0.60	0.68	
Wrist	0.85 (0.46, 1.53)	0.60	0.71	
Thromboembolic events increased	0.70 (0.54, 0.91)	2.61	3.71	5.5 more with tamoxifen
Deep vein thrombosis	0.74 (0.53, 1.03)	1.69	2.29	
Pulmonary embolus	0.64 (0.41, 1.00)	0.91	1.41	
Stroke	0.96 (0.64, 1.43)	1.33	1.39	
Endometrial cancer‡	0.62 (0.35, 1.08)	1.25	2.00	
Cataracts increased	0.79 (0.68, 0.92)	9.72	12.30	13 more with tamoxifen

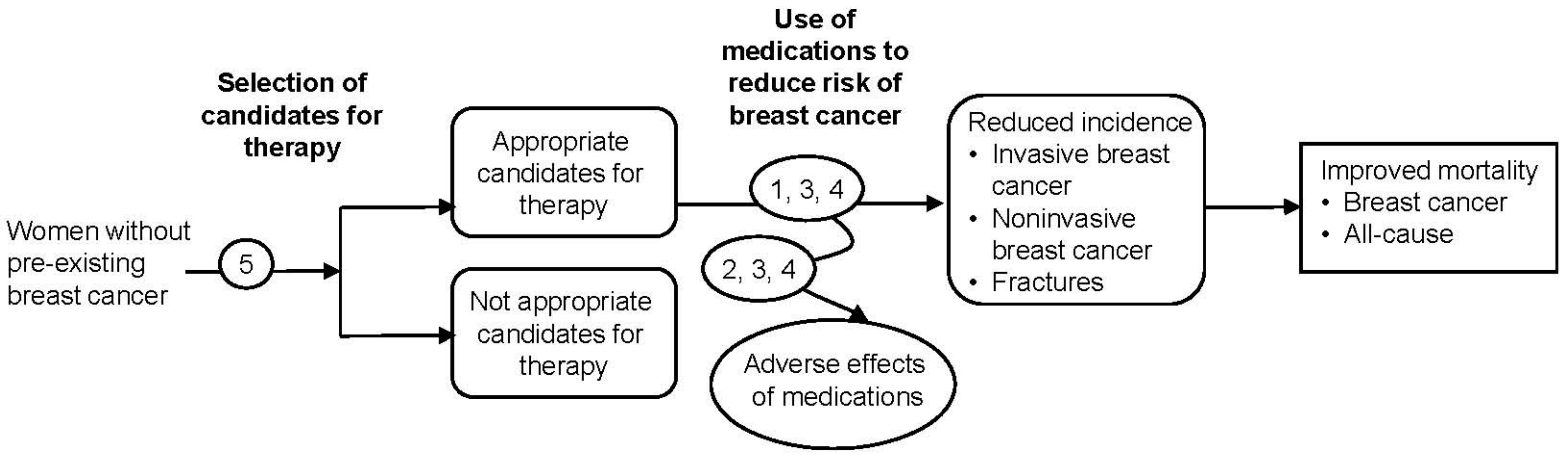
*Per 1000 women-years.

†Numbers of events reduced/increased are calculated by assuming 1000 women take the medication for 5 years.

‡Hyperplasia and hysterectomy rates are higher with tamoxifen among those not diagnosed with uterine cancer.

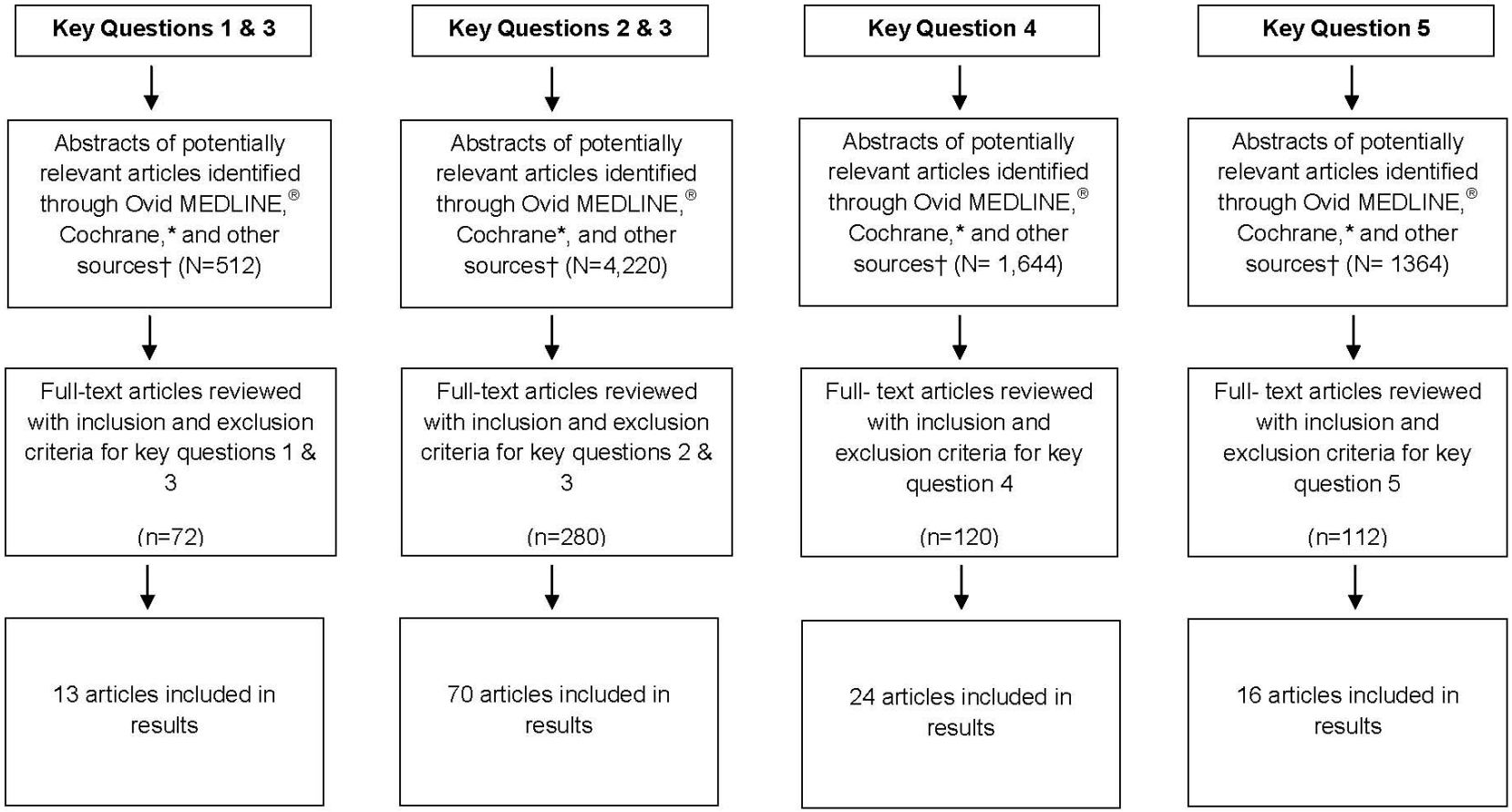
Figures

Figure 1. Analytic framework



Note: Numbers refer to key questions.

Figure 2. Literature flow diagram



*Cochrane databases include the Cochrane Central Register of Controlled Trials and the Cochrane Database of Systematic Reviews
†Identified from reference lists, suggested by experts, etc.
Note: Some abstracts and articles were considered for more than one key question.

Figure 3. Meta-analysis results for all breast cancer outcomes

Trials	No. of Participants Treatment	Placebo	Duration (Mean/Median yrs) Intended Treatment	Total Follow-up	All Breast Cancer Treatment No.	Rate*	Placebo No.	Rate*	Risk Ratio (95% CI)
Tamoxifen									
Fisher, 2005	6681	6707	5	6.1	205	5.02	343	8.44	0.59 (0.50, 0.71)
Cuzik, 2007	3579	3575	5	8.0	142	4.97	195	6.82	0.73 (0.58, 0.91)
Powles, 2007	1238	1233	8	13.2	96	5.60	113	6.60	0.84 (0.64, 1.10)
Veronesi, 2007	2700	2708	4#	11.2	62	2.07	74	2.48	0.84 (0.60, 1.17)
Combined (Test of heterogeneity: Q= 6.4, I^2 = 53.2%; df = 3, P = 0.093)									**0.72 (0.61, 0.86)**
Raloxifene									
Martino, 2004	5129	2576	4 or 8&	5.4&	56	1.96	65	4.71	0.42 (0.29, 0.60)
Barrett-Connor, 2006	5044	5057	5.1#	5.6	52	1.85	76	2.70	0.67 (0.47, 0.96)
Combined (Test of heterogeneity: Q= 3.2, I^2 = 69.0%; df = 1, P = 0.072)									**0.53 (0.34, 0.84)**

* per 1,000 women-years

Veronesi, 2007 and Barrett-Connor, 2006 reported mean or median duration of the actual treatment period.

& The analysis included data from both MORE and CORE. Participants from MORE had 4-year treatment and those who continued in CORE had 4 additional years of treatment. The total follow-up time is averaged over both MORE and CORE for 7705 participants.

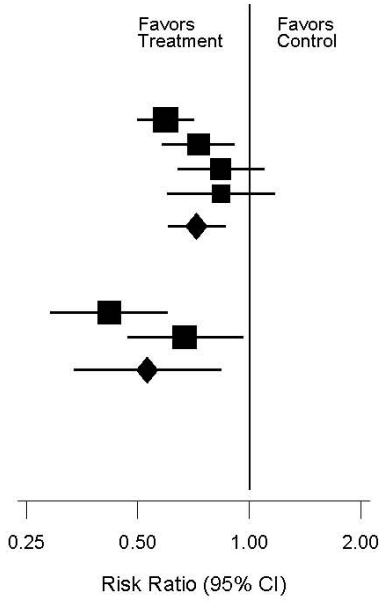


Figure 4. Meta-analysis results for invasive breast cancer

Trials	No. of Participants		Duration (Mean/Median yrs)		Invasive Breast Cancer				Risk Ratio (95% CI)
					Treatment		Placebo		
	Treatment	Placebo	Intended Treatment	Total Follow-up	No.	Rate*	No.	Rate*	
Tamoxifen									
Fisher, 2005	6681	6707	5	6.1	145	3.59	250	6.29	0.57 (0.46, 0.70)
Cuzik, 2007	3579	3575	5	8.0	124	4.34	168	5.88	0.74 (0.58, 0.94)
Powles, 2007	1238	1233	8	13.2	82	4.80	104	6.10	0.78 (0.58, 1.04)
Veronesi, 2007	2700	2708	4#	11.2	53	1.77	66	2.21	0.80 (0.56, 1.15)
Combined (Test of heterogeneity: Q= 4.8, I^2 = 37.6%; df = 3, P = 0.186)									**0.70 (0.59, 0.82)**
Raloxifene									
Martino, 2004	5129	2576	4 or 8&	5.4&	40	1.40	58	4.20	0.34 (0.22, 0.50)
Barrett-Connor, 2006	5044	5057	5.1#	5.6	40	1.43	70	2.49	0.56 (0.38, 0.83)
Combined (Test of heterogeneity: Q= 3.0, I^2 = 66.7%; df = 1, P = 0.084)									**0.44 (0.27, 0.71)**
Tibolone									
Cummings, 2008	2249	2257	2.8	2.8	6	0.90	19	2.80	0.32 (0.13, 0.80)

* per 1,000 women-years

Veronesi, 2007 and Barrett-Connor, 2006 reported mean or median duration of the actual treatment period.

& The analysis included data from both MORE and CORE. Participants from MORE had 4-year treatment and those who continued in CORE had 4 additional years of treatment. The total follow-up time is averaged over both MORE and CORE for 7705 participants.

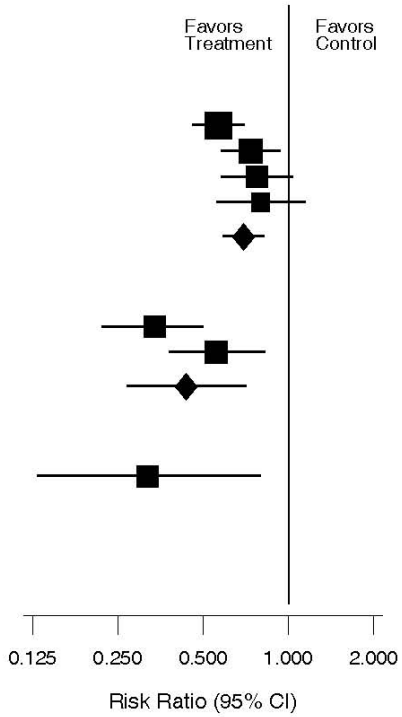


Figure 5. Meta-analysis results for estrogen receptor positive and negative breast cancer

Trials	No. of Participants: Treatment	No. of Participants: Placebo	Duration (Mean/Median yrs): Intended Treatment	Duration (Mean/Median yrs): Total Follow-up	ER Positive Cancer: Treatment No.	Treatment Rate*	Placebo No.	Placebo Rate*	Risk Ratio (95% CI)
Tamoxifen									
Fisher, 2005	6681	6707	5	6.1	70	1.74	182	4.58	0.38 (0.28, 0.50)
Cuzik, 2007	3579	3575	5	8.0	87	3.05	132	4.62	0.66 (0.50, 0.87)
Powles, 2007	1238	1233	8	13.2	53	3.10	86	5.10	0.61 (0.43, 0.86)
Veronesi, 2007	2700	2708	4#	11.2	40	1.34	52	1.74	0.77 (0.51, 1.16)
Combined (Test of heterogeneity: Q= 10.8, I^2 = 72.1%; df = 3, P = 0.013)									**0.58 (0.42, 0.79)**
Raloxifene									
Martino, 2004	5129	2576	4 or 8&	5.4&	22	0.80	44	3.20	0.24 (0.15, 0.40)
Barrett-Connor, 2006	5044	5057	5.1#	5.6	25	0.89	55	1.96	0.45 (0.28, 0.72)
Combined (Test of heterogeneity: Q= 3.3, I^2 = 69.7%; df = 1, P = 0.070)									**0.33 (0.18, 0.61)**
					ER Negative Cancer				
Tamoxifen									
Fisher, 2005	6681	6707	5	6.1	56	1.39	42	1.06	1.31 (0.86, 2.01)
Cuzik, 2007	3579	3575	5	8.0	35	1.23	35	1.23	1.00 (0.61, 1.65)
Powles, 2007	1238	1233	8	13.2	24	1.40	17	1.00	1.40 (0.70, 2.60)
Veronesi, 2007	2700	2708	4#	11.2	21	0.70	19	0.64	1.10 (0.59, 2.05)
Combined (Test of heterogeneity: Q= 1.0, I^2 = 0.0%; df = 3, P = 0.810)									**1.19 (0.92, 1.55)**
Raloxifene									
Martino, 2004	5129	2576	4 or 8&	5.4&	15	0.53	7	0.51	1.06 (0.43, 2.59)
Barrett-Connor, 2006	5044	5057	5.1#	5.6	13	0.46	9	0.32	1.44 (0.61, 3.36)
Combined (Test of heterogeneity: Q= 0.2, I^2 = 0.0%; df = 1, P = 0.628)									**1.25 (0.67, 2.31)**

* per 1,000 women-years
# Veronesi, 2007 and Barrett-Connor, 2006 reported mean or median duration of the actual treatment period.
& The analysis included data from both MORE and CORE. Participants from MORE had 4-year treatment and those who continued in CORE had 4 additional years of treatment. The total follow-up time is averaged over both MORE and CORE for 7705 participants.

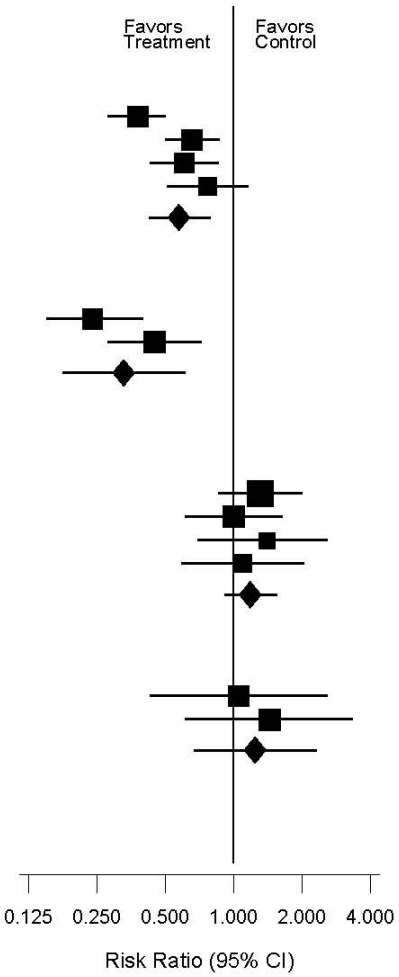


Figure 6. Meta-analysis results for invasive and estrogen receptor positive breast cancer—active and post treatment

Trials	No. of Participants: Treatment	No. of Participants: Placebo	Duration (Mean/Median yrs)	Invasive Breast Cancer: Treatment No.	Invasive Breast Cancer: Treatment Rate*	Invasive Breast Cancer: Placebo No.	Invasive Breast Cancer: Placebo Rate*	Risk Ratio (95% CI)
Tamoxifen - Active treatment								
Cuzik, 2007	3579	3575	5	73	4.09	99	5.54	0.74 (0.55, 0.998)
Powles, 2007	1238	1233	8	44	4.50	48	5.00	0.91 (0.61, 1.37)
Combined (Test of heterogeneity: Q= 0.7, I^2 = 0.0% ; df = 1, P = 0.416)								**0.80 (0.62, 1.01)**
Tamoxifen - Post treatment								
Cuzik, 2007	NR	NR	3	49	4.58	68	6.35	0.72 (0.50, 1.04)
Powles, 2007	NR	NR	5.2	38	5.10	56	7.60	0.67 (0.44, 1.01)
Combined (Test of heterogeneity: Q= 0.1, I^2 = 0.0% ; df = 1, P = 0.795)								**0.70 (0.53, 0.92)**
				ER-positive Breast Cancer				
Tamoxifen - Active treatment								
Cuzik, 2007	3579	3575	5	54	3.03	73	4.09	0.74 (0.51, 1.07)
Powles, 2007	1238	1233	8	30	3.10	39	4.00	0.77 (0.48, 1.23)
Combined (Test of heterogeneity: Q= 0.1, I^2 = 0.0% ; df = 1, P = 0.897)								**0.75 (0.56, 1.01)**
Tamoxifen - Post treatment								
Cuzik, 2007	NR	NR	3	33	3.08	59	5.51	0.56 (0.35, 0.87)
Powles, 2007	NR	NR	5.2	23	3.10	47	6.40	0.48 (0.29, 0.79)
Combined (Test of heterogeneity: Q= 0.2, I^2 = 0.0% ; df = 1, P = 0.655)								**0.52 (0.37, 0.73)**

* per 1,000 women-years

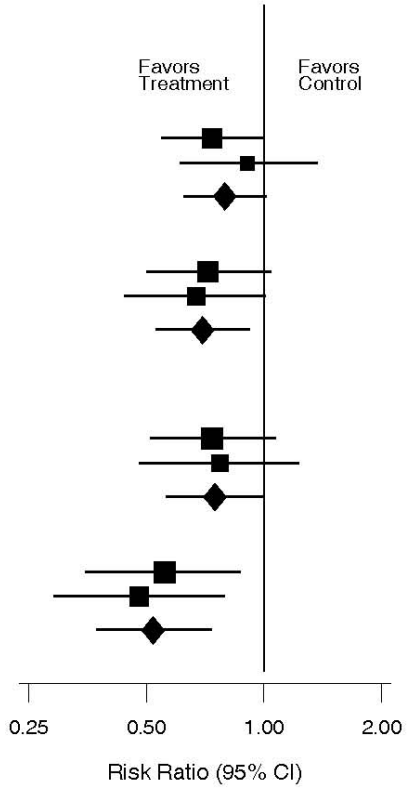


Figure 7. Meta-analysis results for noninvasive breast cancer

Trials	No. of Participants		Duration (Mean/Median yrs)		Non-invasive Breast Cancer				Risk Ratio (95% CI)
					Treatment		Placebo		
	Treatment	Placebo	Intended Treatment	Total Follow-up	No.	Rate*	No.	Rate*	
Tamoxifen									
Fisher, 2005	6681	6707	5	6.1	60	1.47	93	2.29	0.63 (0.45, 0.89)
Cuzik, 2007	3579	3575	5	8.0	17	0.60	27	0.94	0.63 (0.32, 1.20)
Powles, 2007	1238	1233	8	13.2	14	0.80	9	0.50	1.55 (0.67, 3.57)
Veronesi, 2007	2700	2708	4#	11.2	9	0.30	6	0.20	1.50 (0.53, 4.20)
Combined (Test of heterogeneity: Q= 5.9, I^2 = 48.9%; df = 3, P = 0.118)									**0.85 (0.54, 1.35)**
Raloxifene									
Martino, 2004	5129	2576	4 or 8&	5.4&	16	0.56	7	0.51	1.12 (0.46, 2.73)
Barrett-Connor, 2006	5044	5057	5.1#	5.6	11	0.39	5	0.18	2.17 (0.75, 6.24)
Combined (Test of heterogeneity: Q= 0.9, I^2 = 0.0%; df = 1, P = 0.349)									**1.47 (0.75, 2.91)**

* per 1,000 women-years

Veronesi, 2007 and Barrett-Connor, 2006 reported mean or median duration of the actual treatment period.

& The analysis included data from both MORE and CORE. Participants from MORE had 4-year treatment and those who continued in CORE had 4 additional years of treatment. The total follow-up time is averaged over both MORE and CORE for 7705 participants.

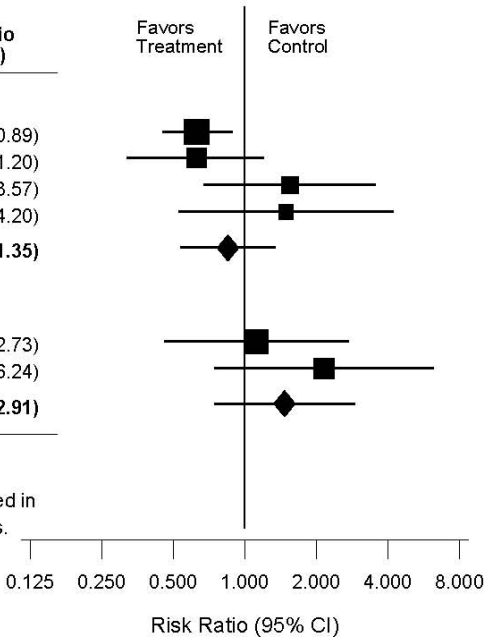


Figure 8. Meta-analysis results for all-cause and breast cancer death

Trials	No. of Participants: Treatment	No. of Participants: Placebo	Duration (Mean/Median yrs): Intended Treatment	Duration (Mean/Median yrs): Total Follow-up	Breast Cancer Death: Treatment No.	Breast Cancer Death: Treatment Rate*	Breast Cancer Death: Placebo No.	Breast Cancer Death: Placebo Rate*	Risk Ratio (95% CI)
Tamoxifen									
Fisher, 2005	6681	6707	5	6.1	12	0.29	11	0.27	1.09 (0.48, 2.46)
Cuzik, 2007	3579	3575	5	8.0	11	0.39	13	0.45	0.85 (0.34, 2.05)
Powles, 2007	1238	1233	8	13.2	12	0.70	9	0.53	1.33 (0.56, 3.16)
Veronesi, 2007	2700	2708	4#	11.2	2	0.07	2	0.07	1.00 (0.14, 7.10)
Combined (Test of heterogeneity: Q= 0.5, I^2 = 0.0%; df = 3, P = 0.919)									**1.07 (0.66, 1.74)**
Raloxifene									
Martino, 2004 (CORE)	2725	1286	4	3.2	0		0		/
					All Cause Death				
Tamoxifen									
Fisher, 2005	6681	6707	5	6.1	126	3.08	114	2.80	1.10 (0.85, 1.43)
Cuzik, 2007	3579	3575	5	8.0	65	2.28	55	1.92	1.18 (0.81, 1.73)
Powles, 2007	1238	1233	8	13.2	54	3.15	54	3.15	0.99 (0.68, 1.44)
Veronesi, 2007	2700	2708	4#	11.2	36	1.46	38	1.54	0.95 (0.60, 1.49)
Combined (Test of heterogeneity: Q= 0.7, I^2 = 0.0%; df = 3, P = 0.867)									**1.07 (0.90, 1.27)**
Raloxifene									
Martino, 2004 (CORE)	2725	1286	4	3.2	47	5.41	29	7.07	0.75 (0.47, 1.19)
Barrett-Connor, 2006	5044	5057	5.1#	5.6	554	2.07	595	2.25	0.92 (0.82, 1.03)
Combined (Test of heterogeneity: Q= 0.7, I^2 = 0.0%; df = 1, P = 0.402)									**0.91 (0.81, 1.02)**

* per 1,000 women-years
# Veronesi, 2007 and Barrett-Connor, 2006 reported mean or median duration of the actual treatment period.

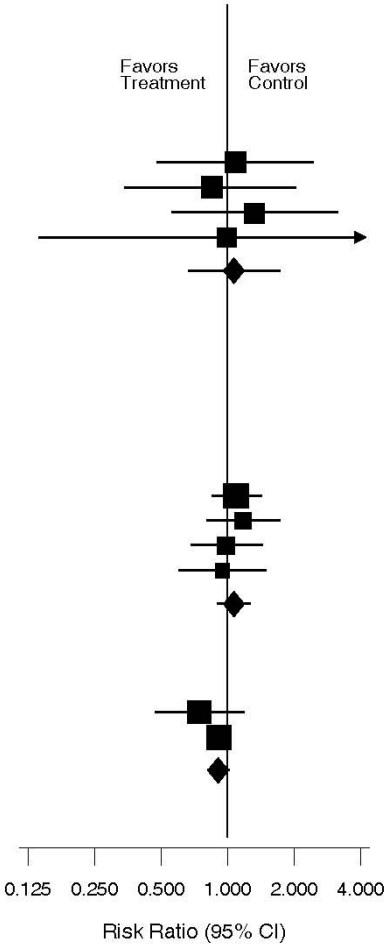


Figure 9. Meta-analysis results for all fractures and osteoporotic site fractures

Trials	No. of Participants		Duration (Mean or Median yrs)		All Fractures				Risk Ratio (95% CI)
					Treatment		Placebo		
	Treatment	Placebo	Intended Treatment	Total Follow-up	No.	Rate*	No.	Rate*	
Tamoxifen									
Cuzik, 2007	3579	3575	5	8.0	121	6.78	142	8.08	0.84 (0.66, 1.07)
Powles, 2007	1238	1233	8	13.2	19	1.94	22	2.29	0.85 (0.46, 1.57)
Combined (Test of heterogeneity: Q= 0.001, I^2 = 0.0%; df = 1, P = 0.977)									**0.84 (0.67, 1.05)**
					Osteoporotic Sites Fracture				
Fisher, 2005	6681	6707	5	6.1	80	1.97	116	2.88	0.68 (0.51, 0.92)
Cuzik, 2007	3579	3575	5	8.0	45	2.52	44	2.50	1.01 (0.67, 1.53)
Combined (Test of heterogeneity: Q= 2.3, I^2 = 56.3%; df = 1, P = 0.130)									**0.81 (0.55, 1.18)**

* per 1,000 women-years. Results are from the active treatment period except for Fisher, 2005 that includes data from the total length of follow-up.

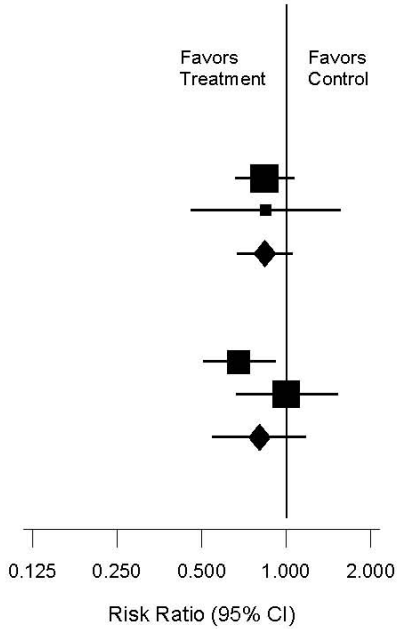


Figure 10. Meta-analysis results for vertebral fractures

Trials	No. of Participants		Duration (Mean or Median yrs)		Vertebral Fracture				Risk Ratio (95% CI)
					Treatment		Placebo		
	Treatment	Placebo	Intended Treatment	Total Follow-up	No.	Rate*	No.	Rate*	
Tamoxifen									
Fisher, 2005	6681	6707	5	6.1	40	0.98	53	1.31	0.75 (0.48, 1.15)
Raloxifene									
Delmas, 2002	4536	2292	4.0	NR	NR	NR	NR	NR	0.60 (0.53, 0.69)
Barrett-Connor, 2006	5044	5057	5.1#	5.6	64	2.28	97	3.45	0.65 (0.47, 0.89)
Combined (Test of heterogeneity: Q= 0.18, I^2 = 0.0%; df = 1, P = 0.676)									**0.61 (0.54, 0.69)**
Tibolone									
Cummings, 2008	2059	2087	2.8	2.8	70	10.9	126	19.6	0.55 (0.41, 0.74)

* per 1,000 women-years
# Barrett-Connor, 2006 reported median duration of the actual treatment period.

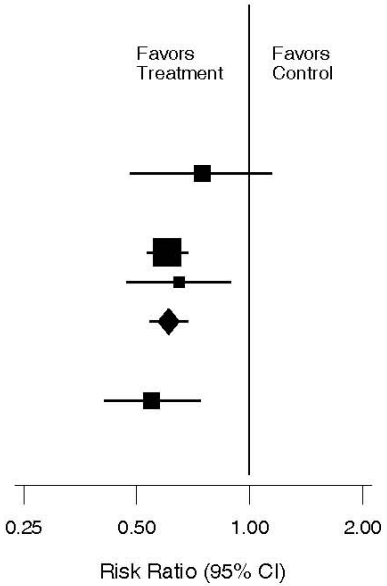


Figure 11. Meta-analysis results for nonvertebral fractures

Trials	No. of Participants		Duration (Mean or Median yrs)		Non-vertebral Fracture				Risk Ratio (95% CI)
					Treatment		Placebo		
	Treatment	Placebo	Intended Treatment	Total Follow-up	No.	Rate*	No.	Rate*	
Tamoxifen									
Fisher, 2005	6681	6707	5	6.1	42	1.03#	63	1.55#	0.66 (0.45, 0.98)
Raloxifene									
Siris, 2005	2725	1286	8	7.9	621	NR	292	NR	1.00 (0.82, 1.21)
Barrett-Connor, 2006	5044	5057	5.1&	5.6	428	15.3	438	15.6	0.96 (0.84, 1.10)
Combined (Test of heterogeneity: Q= 0.11, I^2 = 0.0%; df = 1, P = 0.735)									**0.97 (0.87, 1.09)**
Tibolone									
Cummings, 2008	2249	2257	2.8	2.8	122	19.5	166	26.3	0.74 (0.58, 0.93)

* per 1,000 women-years
# Only hip and radius fractures were included.
& Barrett-Connor, 2006 reported median duration of the actual treatment period.

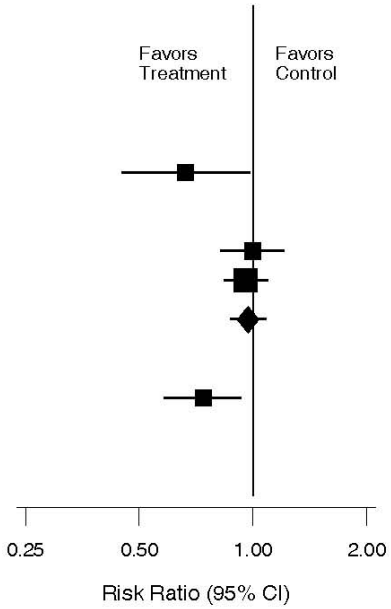


Figure 12. Meta-analysis results for venous thromboembolism

Trials	No. of Participants Treatment	Placebo	Duration (Mean/Median yrs)	Venous Thromboembolism Treatment No.	Rate*	Placebo No.	Rate*	Risk Ratio (95% CI)
Tamoxifen - Active treatment#								
Fisher, 1998	6681	6707	4.0	53	2.03	28	1.07	1.90 (1.20, 3.00)
Decensi, 2005	2700	2708	5.0	10	1.02	9	0.94	1.09 (0.44, 2.68)
Cuzik, 2007	3579	3575	5.0	52	2.91	23	1.29	2.26 (1.36, 3.87)
Powles, 2007	1238	1233	7.8	8	0.82	3	0.31	2.62 (0.69, 9.87)
Combined (Test of heterogeneity: Q= 2.0, I^2 = 0.0% ; df = 3, P = 0.565)								**1.93 (1.41, 2.64)**
Tamoxifen - Post treatment#								
Cuzik, 2007	3449	3489	3.0	16	1.49	14	1.31	1.14 (0.52, 2.53)
Powles, 2007	1079	1034	NR&	5	NR	6	NR	0.80 (0.24, 2.61)
Combined (Test of heterogeneity: Q= 0.2, I^2 = 0.0% ; df = 1, P = 0.688)								**1.02 (0.53, 2.97)**
Raloxifene#								
Grady, 2004	5129	2576	3.3	59	3.50	14	1.70	2.10 (1.20, 3.80)
Barrett-Connor, 2006	5044	5057	5.6	103	3.67	71	2.53	1.44 (1.06, 1.95)
Combined (Test of heterogeneity: Q= 1.3, I^2 = 22.3% ; df = 1, P = 0.257)								**1.60 (1.15, 2.23)**
Tibolone#								
Cummings, 2008	2249	2257	2.8	5	0.80	9	1.30	0.57 (0.19, 1.69)

* per 1,000 women-years

For tamoxifen trials, venous thromboembolic events include deep-vein thrombosis (DVT) and pulmonary embolism (PE) only. For other trials, other thrombosis such as retinal vein thrombosis may be included, depending on the reported overall category.

& Events were reported from at least 3 months after treatment was stopped until the end of follow-up.

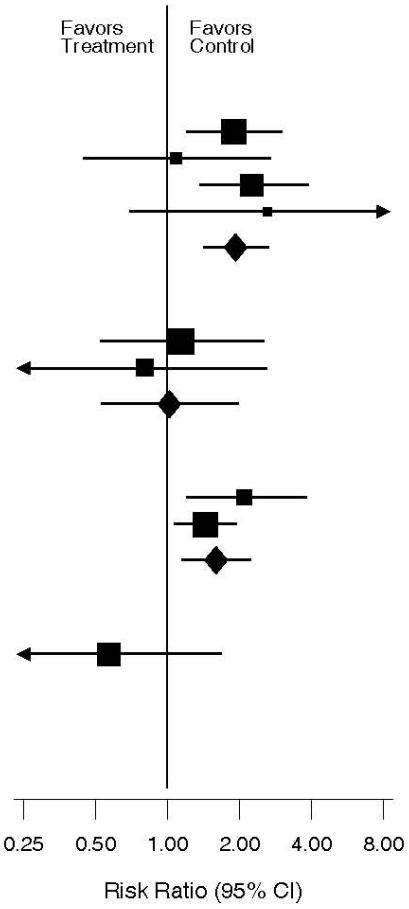


Figure 13. Meta-analysis results for deep vein thrombosis and pulmonary embolism

Trials	No. of Participants Treatment	Placebo	Duration (Mean/Median yrs)	Treatment No.	Treatment Rate*	Placebo No.	Placebo Rate*	Risk Ratio (95% CI)
DVT								
Tamoxifen - Active treatment								
Fisher, 1998	6681	6707	4.0	35	1.34	22	0.84	1.60 (0.91, 2.86)
Decensi, 2005	2700	2708	5.0	9	0.92	8	0.83	1.10 (0.43, 2.86)
Combined (Test of heterogeneity: Q= 0.4, I^2 = 0.0% ; df = 1, P = 0.513)								**1.45 (0.89, 2.37)**
Raloxifene								
Grady, 2004	5129	2576	3.3	43	2.50	7	0.80	3.13 (1.41, 6.95)
Barrett-Connor, 2006	5044	5057	5.6	65	2.32	47	1.67	1.37 (0.94, 1.99)
Combined (Test of heterogeneity: Q= 3.4, I^2 = 70.1% ; df = 1, P = 0.067)								**1.91 (0.87, 4.23)**
PE								
Tamoxifen - Active treatment								
Fisher, 1998	6681	6707	4.0	18	0.69	6	0.23	3.01 (1.15, 9.27)
Decensi, 2005	2700	2708	5.0	1	0.10	1	0.10	0.98 (0.06, 15.70)
Combined (Test of heterogeneity: Q= 0.6, I^2 = 0.0% ; df = 1, P = 0.452)								**2.69 (1.12, 6.47)**
Raloxifene								
Grady, 2004	5129	2576	3.3	18	1.05	2	0.23	3.45 (1.71, 6.94)
Barrett-Connor, 2006	5044	5057	5.6	36	1.28	24	0.85	1.49 (0.89, 2.49)
Combined (Test of heterogeneity: Q= 3.6, I^2 = 72.1% ; df = 1, P = 0.058)								**2.19 (0.97, 4.97)**
Superficial Phlebitis								
Tamoxifen - Active treatment								
Decensi, 2005	2700	2708	5.0	34	3.48	7	1.77	1.96 (1.10, 3.51)
Cuzik, 2007	3579	3575	5.0	17	0.95	6	0.34	2.84 (1.07, 8.78)
Combined (Test of heterogeneity: Q= 0.4, I^2 = 0.0% ; df = 1, P = 0.548)								**2.14 (1.29, 3.56)**

* per 1,000 women-years

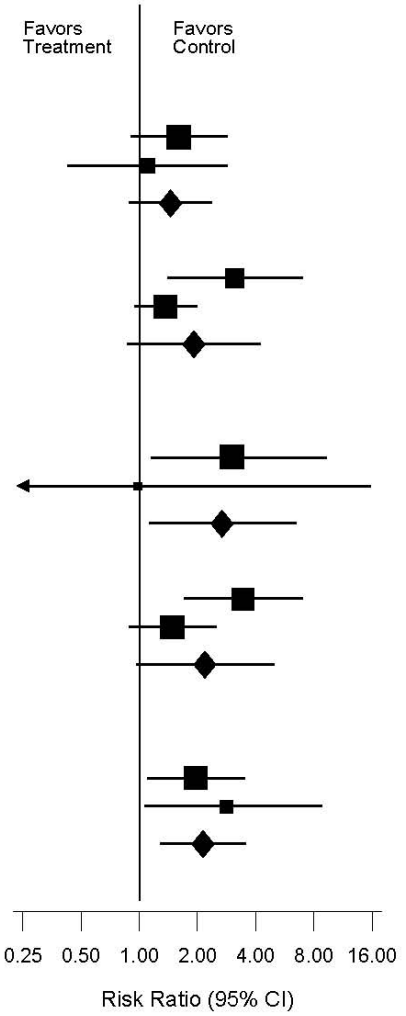


Figure 14. Meta-analysis results for coronary heart disease events

Trials	No. of Participants Treatment	Placebo	Duration (Mean/Median yrs)	CHD events# Treatment No.	Rate*	Placebo No.	Rate*	Risk Ratio (95% CI)
Tamoxifen - Active treatment								
Fisher, 1998	6681	6707	4.0	71	2.73	62	2.37	1.15 (0.81, 1.64)
Cuzik, 2007	3579	3575	5.0	64	3.59	71	3.98	0.90 (0.63, 1.28)
Powles, 2007	1238	1233	7.8	10	1.02	12	1.25	0.82 (0.35, 1.89)
Veronesi, 2007	2700	2708	4.0	5	0.49	5	0.48	1.04 (0.30, 3.58)
Combined (Test of heterogeneity: Q= 1.2, I^2 = 0.0% ; df = 3, P = 0.761)								**1.00 (0.79, 1.27)**
Raloxifene								
Barrett-Connor, 2002	5129	2576	3.4	101	5.75	55	6.29	0.92 (0.66, 1.27)
Barrett-Connor, 2006	5044	5057	5.6	533	19.0	533	19.0	0.95 (0.84, 1.07)
Combined (Test of heterogeneity: Q= 0.04, I^2 = 0.0% ; df = 1, P = 0.837)								**0.95 (0.84, 1.06)**
Tibolone								
Cummings, 2008	2249	2257	2.8	27	4.10	20	3.00	1.37 (0.77, 2.45)

* per 1,000 women-years

CHD events includes any reported coronary heart disearse, such as myocardial infarction, angina, acute ischemic syndrome and other CHD events.

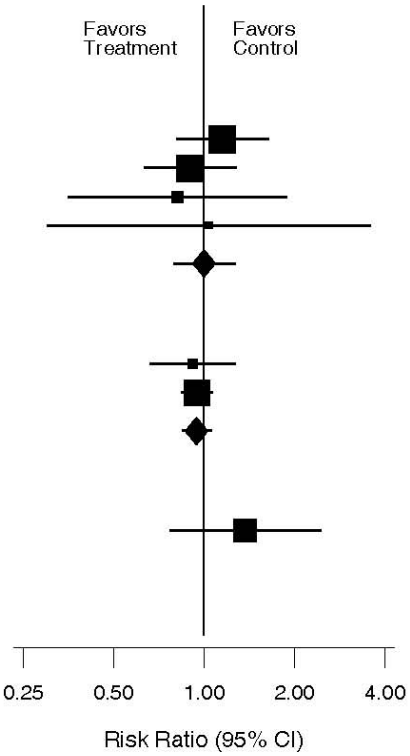


Figure 15. Meta-analysis results for myocardial infarction

Trials	No. of Participants Treatment	Placebo	Duration (Mean/Median yrs)	Myocardial infarction Treatment No.	Rate*	Placebo No.	Rate*	Risk Ratio (95% CI)
Tamoxifen - Active treatment								
Fisher, 1998	6681	6707	4.0	31	1.19	28	1.07	1.11 (0.65, 1.92)
Cuzik, 2007	3579	3575	5.0	2	0.11	7	0.39	0.29 (0.03, 1.50)
Veronesi, 2007	2700	2708	4.0	5	0.49	5	0.48	1.04 (0.30, 3.58)
Combined (Test of heterogeneity: Q= 1.7, I^2 = 0.0% ; df = 2, P = 0.431)								**1.01 (0.63, 1.64)**

* per 1,000 women-years

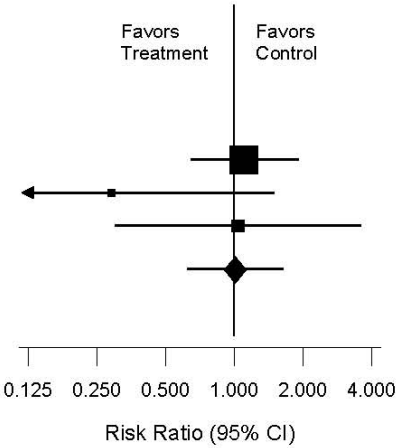


Figure 16. Meta-analysis results for stroke

Trials	No. of Participants Treatment	Placebo	Duration (Mean/Median yrs)	Stroke Treatment No.	Treatment Rate*	Placebo No.	Placebo Rate*	Risk Ratio (95% CI)
Tamoxifen - Active treatment								
Fisher, 1998	6681	6707	4.0	38	1.45	24	0.91	1.59 (0.93, 2.77)
Cuzik, 2007	3579	3575	5.0	8	0.45	8	0.45	1.00 (0.33, 3.06)
Powles, 2007	1238	1233	7.8	7	0.72	9	0.94	0.76 (0.28, 2.05)
Veronesi, 2007	2700	2708	4.0	6	0.59	2	0.19	3.11 (0.63, 15.4)
Combined (Test of heterogeneity: Q= 2.9, I^2 = 0.0% ; df = 3, P = 0.400)								**1.36 (0.89, 2.08)**
Tamoxifen - Post treatment								
Cuzik, 2007	3449	3489	3.0	7	0.65	4	0.37	1.75 (0.45, 8.16)
Powles, 2007	1079	1034	NR#	3	NR	7	NR	0.41 (0.11, 1.58)
Combined (Test of heterogeneity: Q= 1.9, I^2 = 46.8% ; df = 1, P = 0.170)								**0.83 (0.20, 3.42)**
Raloxifene								
Barrett-Connor, 2002	3510	1703	3.4	48	2.73	32	3.66	0.75 (0.48, 1.17)
Barrett-Connor, 2006	5044	5057	5.6	249	8.88	224	7.97	1.10 (0.92, 1.32)
Combined (Test of heterogeneity: Q= 2.5, I^2 = 59.3% ; df = 1, P = 0.117)								**0.96 (0.67, 1.38)**
Tibolone								
Cummings, 2008	2249	2257	2.8	28	4.30	13	1.90	2.19 (1.14, 4.23)

* per 1,000 women-years

Events were reported from at least 3 months after treatment was stopped until the end of follow-up.

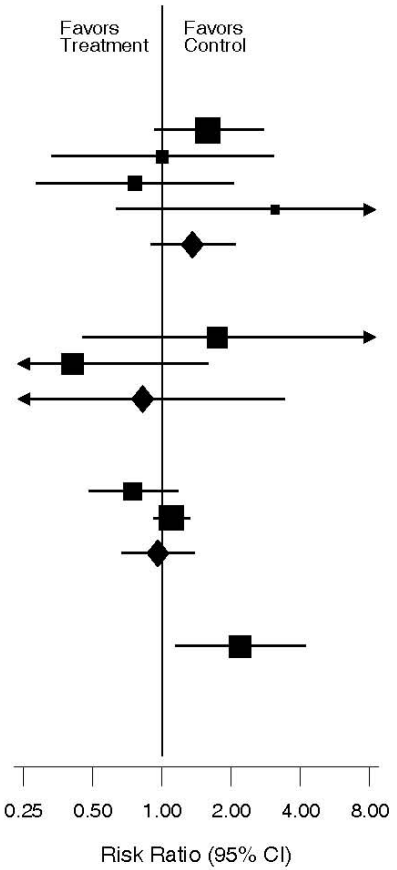


Figure 17. Meta-analysis results for transient ischemic attack

Trials	No. of Participants Treatment	Placebo	Duration (Mean/Median yrs)	Transient ischemic attack Treatment No.	Rate*	Placebo No.	Rate*	Risk Ratio (95% CI)
Tamoxifen - Active treatment								
Fisher, 1998	6681	6707	4.0	19	0.73	25	0.95	0.76 (0.40, 1.44)
Cuzik, 2007	3579	3575	5.0	4	0.22	9	0.50	0.44 (0.10, 1.59)
Veronesi, 2007	2700	2708	4.0	6	0.59	5	0.48	1.24 (0.38, 4.08)
Combined (Test of heterogeneity: Q= 1.2, I^2 = 0.0% ; df = 2, P = 0.535)								**0.77 (0.46, 1.30)**

* per 1,000 women-years

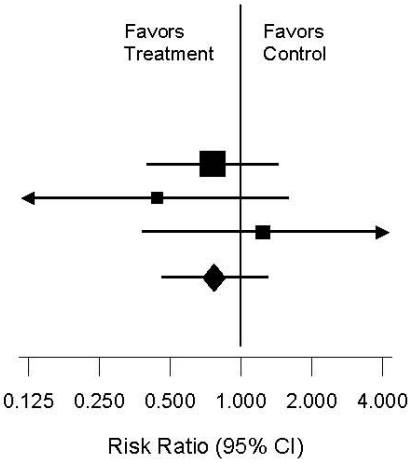


Figure 18. Meta-analysis results for endometrial cancer

Trials	No. of Participants Treatment	No. of Participants Placebo	Duration (Mean/Median yrs)	Endometrial Cancer Treatment No.	Endometrial Cancer Treatment Rate*	Endometrial Cancer Placebo No.	Endometrial Cancer Placebo Rate*	Risk Ratio (95% CI)
Tamoxifen - Active treatment								
Fisher, 1998	4097	4194	4.0	36	2.30	15	0.91	2.53 (1.35, 4.97)
Cuzik, 2007	2347	2292	8.0	17	0.91#	11	0.60#	1.51 (0.71, 3.23)#
Powles, 2007	1238&	1233&	13.2	13	NR	5	NR	2.59 (0.93, 7.24)&
Combined (Test of heterogeneity: Q= 1.2, I^2 = 0.0% ; df = 2, P = 0.551)								**2.13 (1.36, 3.32)**
Raloxifene								
Grady, 2004	3960	1999	3.3	9	NR	5	NR	0.90 (0.30, 2.70)
Barrett-Connor, 2006	3900	3882	5.6	21	0.97	17	0.79	1.23 (0.65, 2.33)
Combined (Test of heterogeneity: Q= 0.2, I^2 = 0.0% ; df = 1, P = 0.630)								**1.14 (0.65, 1.98)**
Tibolone								
Cummings, 2008	1746	1773	2.8	4	0.80	0	0.00	

* per 1,000 women-years, based on number of women with an intact uterus

The rate and RR were recalculated based on the number of women at risk (non-hysterectomized). The values reported in the paper were based on all randomized subjects.

& The number of women at risk (non-hysterectomized) was not reported and risk ratio is calculated based on the number of randomized subjects at baseline.

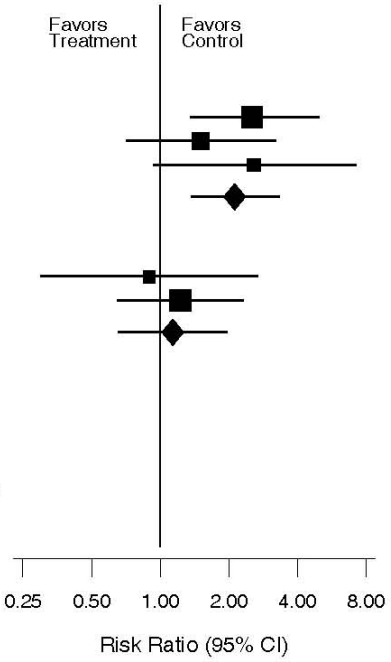


Figure 19. Meta-analysis results for cataracts

Trials	No. of Participants Treatment	Placebo	Duration (Mean/Median yrs)	Cataracts Treatment No.	Treatment Rate*	Placebo No.	Placebo Rate*	Risk Ratio (95% CI)
Tamoxifen								
Fisher, 1998	6101	6131	4.0	574	21.72	507	24.82	1.14 (1.01, 1.29)
Cuzik, 2007	3579	3575	8.0	67	2.35	54	1.89	1.24 (0.87, 1.77)
Powles, 2007	1238	1233	13.2	12	0.70	3	0.18	3.99 (1.13, 14.14)
Combined (Test of heterogeneity: Q= 3.9, I^2 = 48.5%; df = 2, P = 0.144)								**1.25 (0.93, 1.67)**
Raloxifene								
Grady, 2004	5073	2543	3.3	291	NR	160	NR	0.90 (0.80, 1.10)
Barrett-Connor, 2006	5044	5057	5.6	374	13.34	391	13.91	0.96 (0.83, 1.11)
Combined (Test of heterogeneity: Q= 0.3, I^2 = 0.0%; df = 1, P = 0.557)								**0.93 (0.84, 1.04)**

* per 1,000 women-years

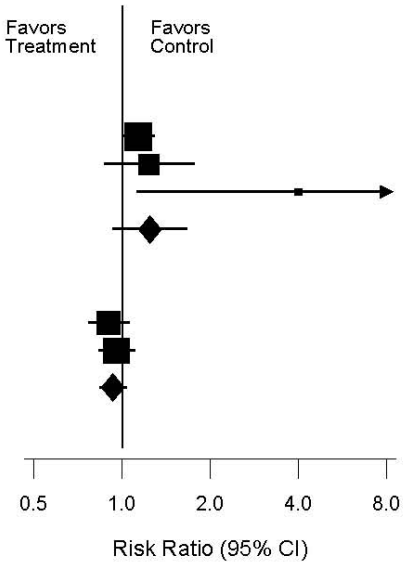


Figure 20. Subgroup analysis by age

Trials	No. of Participants Treatment	No. of Participants Placebo	Total Follow-up (Mean/Median yrs)	Breast Cancer Outcome	Breast Cancer Treatment No.	Breast Cancer Treatment Rate*	Breast Cancer Placebo No.	Breast Cancer Placebo Rate*	Risk Ratio (95% CI)
Tamoxifen - Age <= 50 yrs									
Fisher, 2005	2589	2600	6.1	Invasive	63	4.04	98	6.32	0.64 (0.46, 0.89)
Cuzik, 2007	NR	NR	8.0	All	56	3.64	87	5.63	0.65 (0.45, 0.91)
Veronesi, 2007	1062	1011	11.2	All	22	1.87	22	1.98	0.95 (0.52, 1.71)
Combined (Test of heterogeneity: Q= 1.4, I^2 = 0.0%; df = 2, P = 0.495)									**0.68 (0.55, 0.85)**
Tamoxifen - Age > 50 yrs									
Fisher, 2005	4008	4010	6.1	Invasive	82	3.32	152	6.27	0.53 (0.40, 0.69)
Cuzik, 2007	NR	NR	8.0	All	86	6.54	108	8.24	0.79 (0.59, 1.06)
Veronesi, 2007	1638	1697	11.2	All	38	2.00	56	2.78	0.79 (0.53, 1.20)
Combined (Test of heterogeneity: Q= 4.8, I^2 = 58.6%; df = 2, P = 0.090)									**0.68 (0.51, 0.90)**
Raloxifene									
Lippman, 2006 (<65yrs)	2058	1026	5.4[&]	Invasive	25	1.50	41	5.10	0.42 (0.21, 0.85)
Grady, 2008 (<60 yrs)[#]	NR	NR	5.6	Invasive	4	NR	8	NR	0.49 (0.15, 1.64)
Raloxifene									
Lippman, 2006 (>=65yrs)	2563	1550	5.4[&]	Invasive	15	1.30	17	3.00	0.30 (0.18, 0.49)
Grady, 2008 (>=60 yrs)[#]	NR	NR	5.6	Invasive	36	NR	62	NR	0.57 (0.38, 0.86)

* per 1,000 women-years

& The total follow-up time is averaged over both MORE and CORE for the 7705 participants.

For Grady 2008, total n = 1670 for age < 60 yrs and 8431 for age >= 60 yrs.

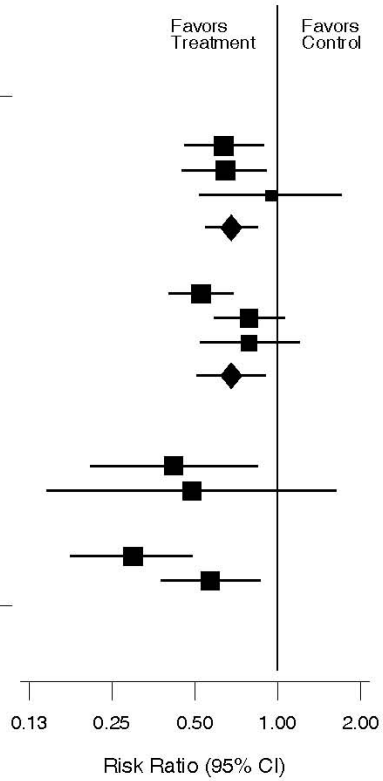


Figure 21. Subgroup analysis by menopausal status

Trials	No. of Participants Treatment	Placebo	Total Follow-up (Mean/Median yrs)	Breast Cancer Outcome	Breast Cancer Treatment No.	Treatment Rate*	Placebo No.	Placebo Rate*	Risk Ratio (95% CI)
Tamoxifen - Premenopausal									
Cuzik, 2007	1644	1653	8.0	All	58	4.20	88	6.25	0.67 (0.47, 0.95)
Powles, 2007	801	798	13.2	ER-positive	14	2.80	28	5.60	0.50 (0.26, 0.95)
Combined (Test of heterogeneity: Q= 0.6, I^2 = 0.0%; df = 1, P = 0.437)									**0.63 (0.46, 0.85)**
Tamoxifen - Postmenopausal									
Cuzik, 2007	1935	1922	8.0	All	84	5.86	107	7.58	0.77 (0.57, 1.04)
Powles, 2007	388	392	13.2	ER-positive	9	3.70	19	8.10	0.46 (0.21, 1.02)
Combined (Test of heterogeneity: Q= 1.4, I^2 = 29.9%; df = 1, P = 0.232)									**0.68 (0.44, 1.05)**

* per 1,000 women-years

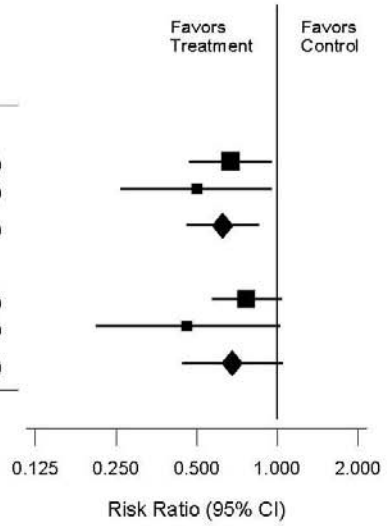


Figure 22. Subgroup analysis by estrogen use

Trials	No. of Participants Treatment	No. of Participants Placebo	Total Follow-up (Mean/Median yrs)	Breast Cancer Outcome	Breast Cancer Treatment No.	Breast Cancer Treatment Rate*	Breast Cancer Placebo No.	Breast Cancer Placebo Rate*	Risk Ratio (95% CI)
Tamoxifen - HRT use = Yes#									
Cuzik, 2007	1462	1414	8.0	All	66	5.52	69	6.00	0.92 (0.65, 1.31)
Powles, 2007	450	464	13.2	ER-positive	12	3.60	25	7.90	0.46 (0.23, 0.91)
Veronesi, 2007	311	289	11.2	All	6	1.71	6	1.82	0.94 (0.30, 2.92)
Combined (Test of heterogeneity: Q= 3.2, I^2 = 37.0%; df = 2, P = 0.205)									**0.75 (0.47, 1.20)**
Tamoxifen - HRT use = No#									
Cuzik, 2007	2114	2141	8.0	All	76	4.58	126	7.38	0.62 (0.46, 0.83)
Powles, 2007	788	769	13.2	ER-positive	11	2.70	22	5.30	0.51 (0.25, 1.05)
Veronesi, 2007	2419	2389	11.2	All	56	2.12	68	2.56	0.83 (0.58, 1.18)
Combined (Test of heterogeneity: Q= 2.2, I^2 = 9.0%; df = 2, P = 0.335)									**0.68 (0.54, 0.86)**
Raloxifene - HRT use = Yes#									
Lippman, 2006	1497	738	5.4&	Invasive	12	1.50	20	5.40	0.29 (0.14, 0.59)
Grady, 2008	979	1025	5.6	Invasive	14	2.57	21	3.68	0.70 (0.35, 1.37)
Combined (Test of heterogeneity: Q= 3.0, I^2 = 67.1%; df = 1, P = 0.081)									**0.45 (0.19, 1.07)**
Raloxifene - HRT use = No#									
Lippman, 2006	3614	1833	5.4&	Invasive	28	1.40	38	3.80	0.36 (0.22, 0.59)
Grady, 2008	4065	4032	5.6	Invasive	26	1.15	48	2.14	0.54 (0.33, 0.87)
Combined (Test of heterogeneity: Q= 1.3, I^2 = 23.5%; df = 1, P = 0.253)									**0.44 (0.30, 0.65)**

* per 1,000 women-years

For tamoxifen trials, hormone replacement therapy (HRT) use refers to HRT use during the trial period only. For raloxifene trials, HRT use refers to prior HRT use.

& The total follow-up time is averaged over both MORE and CORE for the 7705 participants.

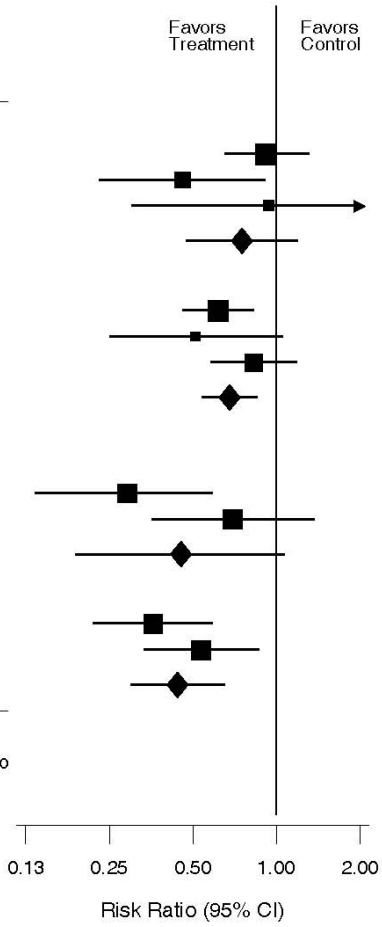


Figure 23. Subgroup analysis by family history of breast cancer

Trials	No. of Participants Treatment	Placebo	Total Follow-up (Mean/Median yrs)	Breast Cancer Outcome	Breast Cancer Treatment No.	Treatment Rate*	Placebo No.	Placebo Rate*	Risk Ratio (95% CI)
Tamoxifen - Without FH#									
Fisher, 2005	1548	1597	6.1	Invasive	33	3.48	62	6.47	0.54 (0.34, 0.83)
Veronesi, 2007	2359	2407	11.2	All	46	1.75	64	2.41	0.73 (0.50, 1.06)
Combined (Test of heterogeneity: Q= 1.03, I^2 = 2.5%; df = 1, P = 0.311)									**0.64 (0.48, 0.86)**
Tamoxifen - With FH#									
Fisher, 2005	5049	5013	6.1	Invasive	12	3.62	188	6.23	0.58 (0.46, 0.73)
Veronesi, 2007	341	301	11.2	All	16	4.29	10	3.00	1.43 (0.65, 3.15)
Raloxifene - Without FH#									
Lippman, 2006	4373	2196	5.4&	Invasive	36	1.50	42	3.50	0.42 (0.27, 0.66)
Grady, 2008	4592	4612	5.6	Invasive	29	1.14	53	2.07	0.53 (0.34, 0.84)
Combined (Test of heterogeneity: Q= 0.51, I^2 = 0.0%; df = 1, P = 0.473)									**0.47 (0.34, 0.65)**
Raloxifene - With FH#									
Lippman, 2006	636	313	5.4&	Invasive	3	0.90	13	8.10	0.11 (0.03, 0.38)
Grady, 2008	452	445	5.6	Invasive	8	3.18	9	3.64	0.89 (0.34, 2.31)

* per 1,000 women-years
# With family history (FH) is defined as having at least one first-degree relative with breast cancer, and otherwise it is without FH.
& The total follow-up time is averaged over both MORE and CORE for the 7705 participants.

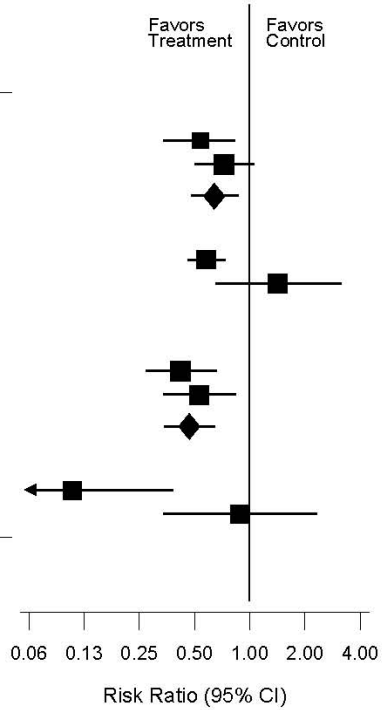


Figure 24. Subgroup analysis by body mass index

Trials	No. of Participants Treatment	Placebo	Total Follow-up (Mean/Median yrs)	Breast Cancer Outcome	Breast Cancer Treatment No.	Rate*	Placebo No.	Rate*	Risk Ratio (95% CI)
Raloxifene - BMI <= 25									
Lippman, 2006	2701	1334	5.4[&]	Invasive	16	1.10	26	3.60	0.29 (0.16, 0.55)
Grady, 2008[#]	NR	NR	5.6	Invasive	9	NR	11	NR	0.84 (0.35, 2.03)
Combined (Test of heterogeneity: Q= 3.8, I^2 = 73.4%; df = 1, P = 0.052)									**0.47 (0.17, 1.33)**
Raloxifene - BMI > 25									
Lippman, 2006	2427	1241	5.4[&]	Invasive	26	1.80	32	4.90	0.37 (0.22, 0.63)
Grady, 2008[#]	NR	NR	5.6	Invasive	31	NR	58	NR	0.52 (0.30, 0.90)
Combined (Test of heterogeneity: Q= 0.7, I^2 = 0.0%; df = 1, P = 0.394)									**0.43 (0.30, 0.63)**

* per 1,000 women-years

For Grady 2008, total n = 2416 for BMI <=25, and 7655 for BMI > 25.

& The total follow-up time is averaged over both MORE and CORE for the 7705 participants.

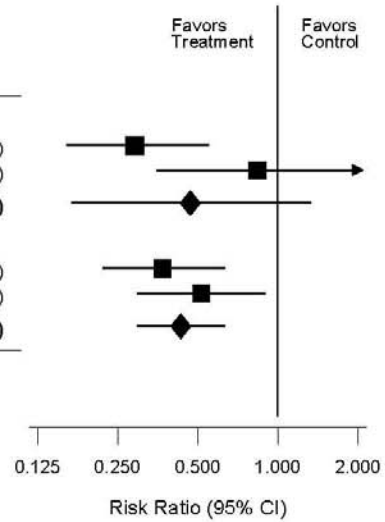
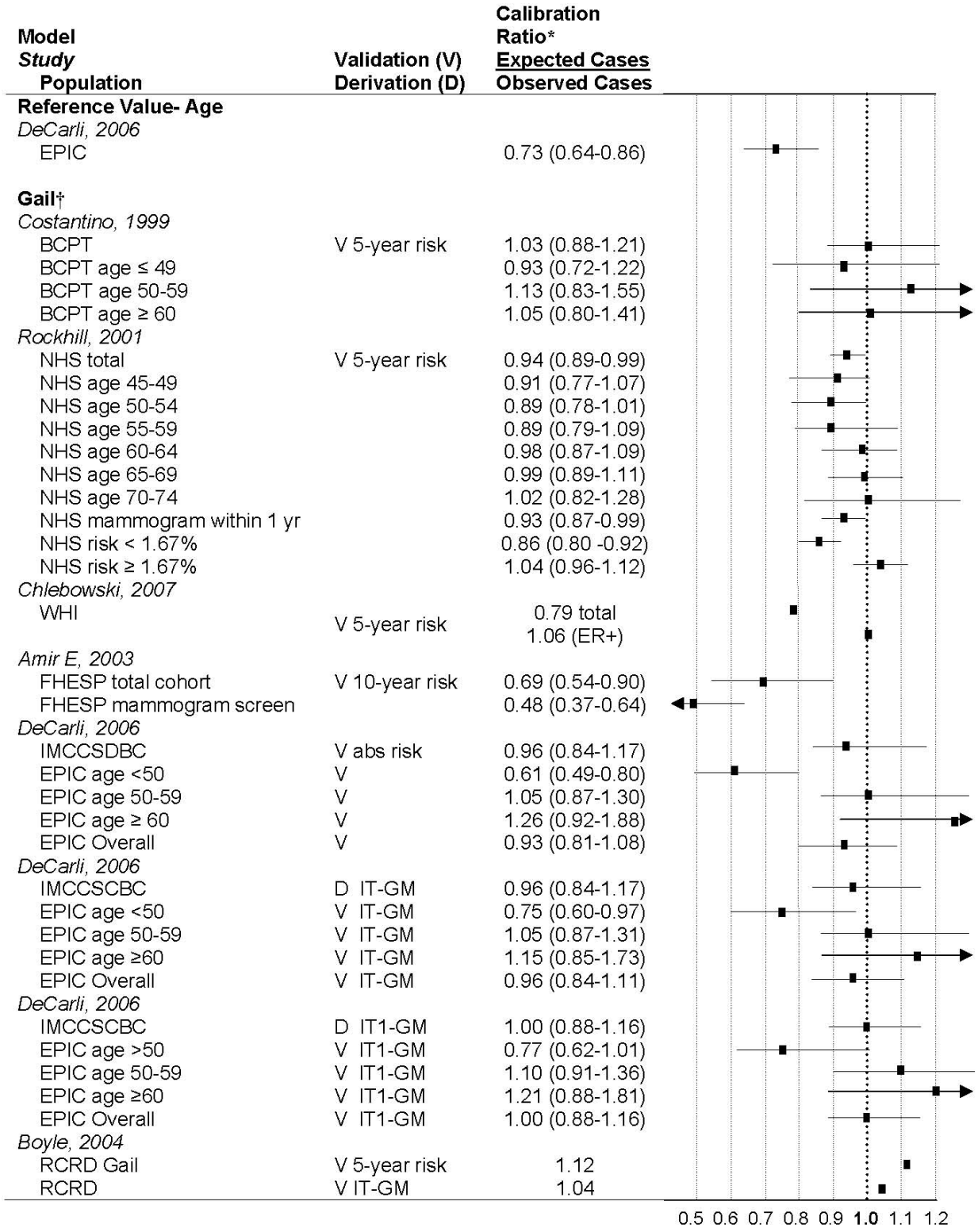
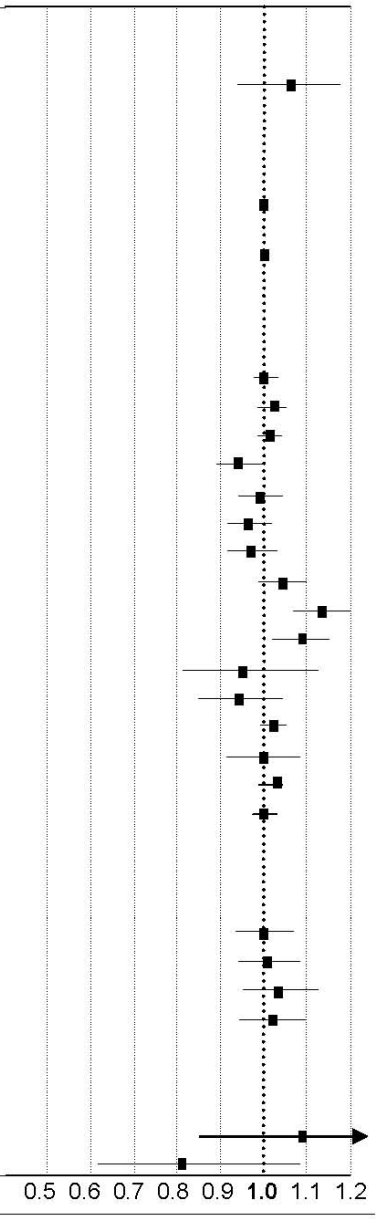


Figure 25. Calibration of breast cancer risk models



Model *Study* Population	Validation (V) Derivation (D)	Calibration Ratio Expected Cases Observed Cases
Gail- African American		
Gail M, 2007		
CARE	D 5-year risk	1.08 (0.97-1.20)
BCSC- Barlow‡		
Barlow W, 2006		
BCSC premenopausal	D(75%)+V(25%) 1-year risk	1.00
BCSC postmenopausal	D(75%)+V(25%) 1-year risk	1.01
BCSC- Tice‡		
Tice JA, 2008		
BCSC 60%	D 5-year risk	1.00 (0.98-1.03)
BCSC 40%	V 5-year risk	1.03 (0.99-1.06)
BCSC total		1.01 (0.99-1.03)
BCSC age 40-44		0.94 (0.89-1.00)
BCSC age 45-49		0.99 (0.94-1.04)
BCSC age 50-54		0.96 (0.92-1.01)
BCSC age 55-59		0.97 (0.92-1.02)
BCSC age 60-64		1.04 (0.98-1.10)
BCSC age 65-69		1.13 (1.07-1.20)
BCSC age 70-74		1.08 (1.02-1.15)
BCSC Asian		0.95 (0.81-1.12)
BCSC Hispanic		0.94 (0.85-1.04)
BCSC White		1.02 (0.99-1.04)
BCSC Black		1.00 (0.91-1.09)
BCSC GAIL≥1.67%		1.03 (0.98-1.03)
BCSC GAIL<1.67%		1.00 (0.98-1.03)
Colditz Rosner		
Rockhill B, 2003		
NHS model 1 total	V 5-year risk	1.00 (0.93-1.07)
NHS model 2 total		1.01 (0.94-1.09)
NHS model 1 postmenopausal		1.03 (0.95-1.11)
NHS model 2 postmenopausal		1.02 (0.94-1.10)
Tyrer- Cuzick§		
Amir E, 2003		
FHESP total	V 10-year risk	1.09 (0.85-1.41)
FHESP screened		0.81 (0.62-1.08)



*Chen and Chlebowski Models did not report Calibration Ratio for the models developed in the studies.

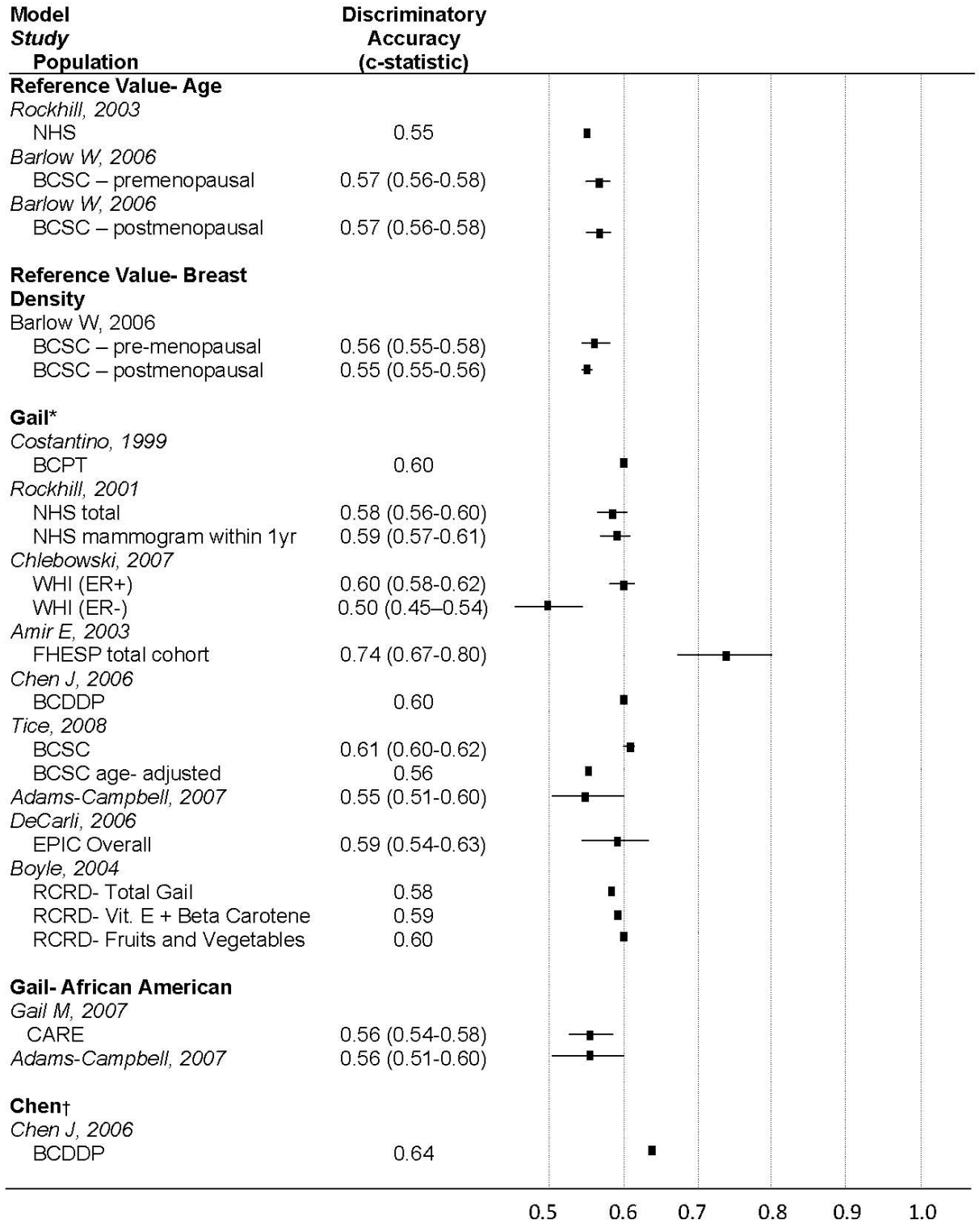
† Gail model used to determine inclusion for P-1, P-2 trials; measured in RUTH, MORE.

‡ Models including breast density as risk factor.

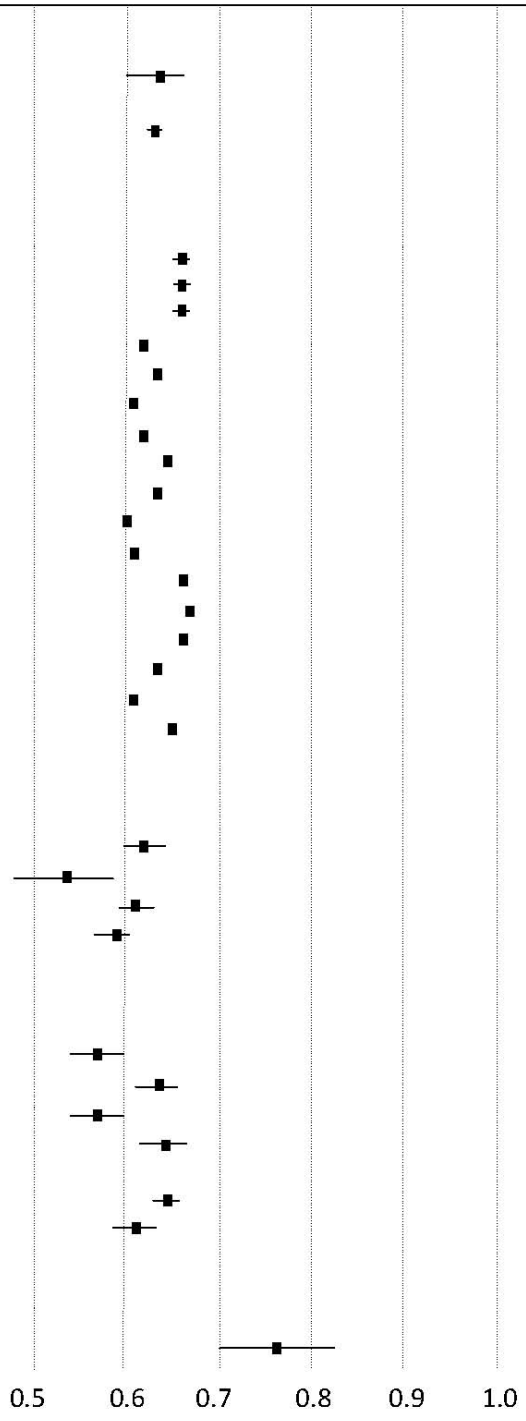
§ Cuzick model used to determine inclusion for IBIS trial.

Abbreviations: BCPT, Breast Cancer Prevention Trial; NHS, Nurses' Health Study; WHI, Women's Health Initiative; BCSC, Breast Cancer Surveillance Consortium; IMCCSDBC, Italian Multicenter Case-control Study of Diet and Breast Cancer; EPIC, European Prospective Investigation into Cancer and Nutrition; CARE, Women's Contraceptive and Reproductive Experiences; FHESP, Family History and Evaluation Screening Program at University Hospital of South Manchester; RCRD, Regional Cancer Registry Data.

Figure 26. Discriminatory accuracy of breast cancer risk models



Model *Study* Population	Discriminatory Accuracy (c-statistic)
BCSC- Barlow†	
Barlow W, 2006	
BCSC premenopausal	0.63 (0.60-0.66)
Barlow W, 2006	
BCSC postmenopausal	0.62 (0.62-0.63)
BCSC- Tice†	
Tice JA, 2008	
BCSC – 60%	0.66 (0.65-0.67)
BCSC – 40%	0.66 (0.66-0.67)
BCSC - total	0.66 (0.65-0.66)
BCSC – age adjusted	0.62
BCSC – ages 40-44	0.63
BCSC – age 45-49	0.61
BCSC – age 50-54	0.62
BCSC – age 55-59	0.64
BCSC – age 60-64	0.63
BCSC – age 65-69	0.60
BCSC – age 70-74	0.61
BCSC – Asian	0.66
BCSC – Hispanic	0.67
BCSC – White	0.66
BCSC – Black	0.63
BCSC – GAIL≥1.67%	0.61
BCSC – GAIL<1.67%	0.65
Chlebowski	
Chlebowski, 2007	
WHI expanded (ER+)	0.62 (0.60-0.64)
WHI expanded (ER-)	0.53 (0.47-0.58)
WHI expanded (total)	0.61 (0.59-0.63)
WHI simplified (ER+)	0.58 (0.56-0.60)
Colditz- Rosner	
Rockhill B, 2003	
NHS model 1 – total	0.57 (0.55-0.59)
NHS model 2 – total	0.63 (0.61-0.65)
NHS model 1 postmenopausal	0.57 (0.55-0.59)
NHS model 2 postmenopausal	0.64 (0.62-0.66)
Colditz GA, 2004	
NHS	0.64 (0.63-0.66)
NHS	0.61 (0.58-0.64)
Tyrer- Cuzick‡	
Amir E, 2003	
FHESP total	0.76 (0.70-0.82)



*Gail model used to determine inclusion for P-1, P-2 trials; measured in RUTH, MORE.
†Models including breast density as risk factor.
‡Cuzick model used to determine inclusion for IBIS trial

Abbreviations: BCPT, Breast Cancer Prevention Trial; NHS, Nurses' Health Study; WHI, Women's Health Initiative; BCSC, Breast Cancer Surveillance Consortium; IMCCSDBC, Italian Multicenter Case-control Study of Diet and Breast Cancer; EPIC, European Prospective Investigation into Cancer and Nutrition; CARE, Women's Contraceptive and Reproductive Experiences; FHESP, Family History and Evaluation Screening Program at University Hospital of South Manchester; BCDDP, Breast Cancer Detection and Demonstration Project, RCRD, Regional Cancer Registry Data.