

Technology

Report

Issue 50

January 2005

**Raloxifene for Primary
and Secondary
Prevention of
Osteoporotic Fractures
in Postmenopausal
Women: A Systematic
Review of Efficacy and
Safety Evidence**

Publications can be requested from:

CCOHTA
600-865 Carling Avenue
Ottawa ON Canada K1S 5S8
Tel. (613) 226-2553
Fax. (613) 226-5392
Email: pubs@ccohta.ca

or download from CCOHTA's web site:
<http://www.ccohta.ca>

Cite as: Schachter HM, Clifford TJ, Cranney A, Barrowman NJ, Moher D. *Raloxifene for primary and secondary prevention of osteoporotic fractures in postmenopausal women: a systematic review of efficacy and safety evidence* [Technology report no 50]. Ottawa: Canadian Coordinating Office for Health Technology Assessment; 2005.

Reproduction of this document for non-commercial purposes is permitted provided appropriate credit is given to CCOHTA.

CCOHTA is a non-profit organization funded by the federal, provincial and territorial governments.

Legal Deposit – 2005
National Library of Canada
ISBN: 1-894978-21-8 (print)
ISBN: 1-894978-22-6 (online)

PUBLICATIONS MAIL AGREEMENT NO: 40026386
RETURN UNDELIVERABLE CANADIAN ADDRESSES TO
CANADIAN COORDINATING OFFICE FOR HEALTH TECHNOLOGY ASSESSMENT
600-865 CARLING AVENUE
OTTAWA ON K1S 5S8

Canadian Coordinating Office for Health Technology Assessment

**Raloxifene for Primary and Secondary Prevention of
Osteoporotic Fractures in Postmenopausal Women:
A Systematic Review of Efficacy
and Safety Evidence**

Howard M Schachter PhD¹
Tammy J Clifford PhD¹
Ann Cranney MBBCh²
Nicholas J Barrowman PhD^{1,3,4}
David Moher MSc^{1,4,5}

January 2005

¹ Chalmers Research Group, CHEO Research Institute, Ottawa
² Clinical Epidemiology Program, Ottawa Health Research Institute, Ottawa
³ School of Mathematics and Statistics, Carleton University, Ottawa
⁴ Department of Pediatrics, University of Ottawa, Ottawa
⁵ Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa

Reviewers

These individuals kindly provided comments on this report.

External Reviewers

Allan Mills, PharmD CGP
Pharmacy Department
Baycrest Centre for Geriatric Care
3560 Bathurst Street
Toronto ON

CCOHTA Scientific Advisory Panel Reviewers

Jeffrey Barkun, MD CM FRCSC FACS MSc
(Epidemiology)
Head, Hepatobiliary and Transplantation Unit
Director of Clinical Research in Transplantation
McGill University Health Centre
Associate Professor of Surgery
McGill University
Montreal QC

Jeff Mahon, MD FRCPC
Associate Professor of Medicine
London HSC, University Campus Site
University of Western Ontario
London ON

This report is a review of existing public literature, studies, materials and other information and documentation (collectively the “source documentation”) which are available to CCOHTA. The accuracy of the contents of the source documentation on which this report is based is not warranted, assured or represented in any way by CCOHTA and CCOHTA does not assume responsibility for the quality, propriety, inaccuracies or reasonableness of any statements, information or conclusions contained in the source documentation.

CCOHTA takes sole responsibility for the final form and content of this report. The statements and conclusions in this report are those of CCOHTA and not of its Panel members or reviewers.

Authorship

H.M. Schachter helped refine the research questions; developed and implemented methods used; interpreted the results; wrote all drafts of the report; and managed the review team.

T.J. Clifford, with H.M. Schachter, contributed to the conception and design of the review, including the eligibility criteria data abstraction forms; assisted in data abstraction; and critically reviewed drafts of the report.

A. Cranney provided content expertise, assisted with selection of relevant studies from search results and provided input on drafts of the report.

N.J. Barrowman contributed to the data synthesis plan including imputation strategy, statistical heterogeneity testing and assessment of publication bias; performed quantitative data synthesis; interpreted heterogeneity statistics; and participated in drafting and revising the report.

D. Moher contributed methodological expertise.

Acknowledgements

The authors thank Ms. Margaret Sampson for conducting a feasibility search and for reviewing the final search strategy, Mr. Jagadish Rangrej for helping to organize several of the analyses, Ms. Janet Joyce for performing electronic searches, Ms. Kaitryn Campbell for reference verification and Mr. Ken Bassett for providing his valuable input in writing the final draft.

Conflicts of Interest

H.M. Schachter, T.J. Clifford, N.J. Barrowman, J. Joyce and D. Moher disclosed no conflicts. A. Cranney has received consulting fees from Merck Frosst, Procter and Gamble and Eli Lilly.



Raloxifene for Primary and Secondary Prevention of Osteoporotic Fractures in Postmenopausal Women: A Systematic Review of Efficacy and Safety Evidence

Technology Name

Raloxifene (Evista®)

Disease or Condition

Osteoporosis is a progressive condition of decreased bone mass. This leads to fragile bones, which are prone to fracture. In women, osteoporotic fractures of the vertebrae can result in back pain, decreased quality of life and an increased risk of spinal fractures.

Technology Description

Raloxifene is a selective estrogen receptor modulator (SERM) drug. Licensed in Canada in 1999 for the prevention of osteoporosis in postmenopausal women, raloxifene does not heal existing vertebral deformity or fractures, but it is used to prevent subsequent fractures or to delay structural bone changes.

The Issue

With an aging Canadian population, vertebral fractures will become more common. There is a need to assess raloxifene's efficacy and safety compared with other drugs approved for the prevention of osteoporosis.

Assessment Objectives

- To systematically assess, using a meta-analysis if appropriate, raloxifene's safety and efficacy to prevent osteoporotic vertebral fractures in postmenopausal women.
- To compare raloxifene against placebo and other drug treatments, including estrogen alone

in women with hysterectomies, estrogen-progestin combination therapies, bisphosphonates (alendronate, risedronate and etidronate) and salmon calcitonin.

Methods

An extensive literature search identified randomized controlled trials that assessed the safety and efficacy of raloxifene in postmenopausal women. The main outcome measure was vertebral fracture, with the secondary outcome measure being bone mineral density. A total of 17 studies met the inclusion criteria for the review.

Conclusions

- Raloxifene has no effect on the incidence of non-vertebral fractures such as hip fractures.
- Because of trial diversity, the planned meta-analysis to compare raloxifene's efficacy with other drugs used to prevent osteoporosis could not be accomplished.
- Compared with placebo, raloxifene's primary clinical benefit is a reduction in vertebral fracture in older postmenopausal women, particularly if some vertebral fractures are present at baseline.
- This benefit is offset by a similar increase in serious adverse events due to venous thrombo-embolic disease.
- Raloxifene causes significantly more mild to moderate adverse effects such as hot flashes and leg cramps compared with placebo, but significantly less vaginal bleeding than estrogen-progestin combination therapy.

This summary is based on a comprehensive health technology assessment available from CCOHTA's web site (www.ccohta.ca): Schachter HM, Clifford TJ, Cranney A, Barrowman NJ, Moher D. *Raloxifene for primary and secondary prevention of osteoporotic fractures in postmenopausal women: a systematic review of efficacy and safety evidence.*

Canadian Coordinating Office for Health Technology Assessment (CCOHTA)

600-865 Carling Avenue, Ottawa ON Canada K1S 5S8 Tel: 613-226-2553 Fax: 613-226-5392 www.ccohta.ca

CCOHTA is an independent, not-for-profit organization that supports informed health care decision-making by providing unbiased, reliable information about health technologies.

EXECUTIVE SUMMARY

The Issue

Osteoporosis is a progressive systemic disease characterized by low bone density and deterioration of bone tissue, with bones becoming fragile and prone to fracture. In women, symptomatic and asymptomatic osteoporotic vertebral fractures often occur 10 to 15 years after menopause, which may result in back pain, decreased quality of life and an increased risk of spinal fractures. With an aging Canadian population, vertebral fractures will become more common. In 1993, the cost of illness in Canada due to osteoporosis and its complications was estimated at \$1.3 billion. In 1999, raloxifene was licensed in Canada for the prevention of osteoporosis in postmenopausal women. Raloxifene does not heal existing fractures or vertebral fracture, but is used to delay structural bone changes and prevent future fractures. Primary prevention is aimed at women who have no baseline radiologic evidence of vertebral fracture, while secondary prevention targets women with baseline evidence of vertebral fracture.

Objectives

The purpose of this systematic review is to evaluate, using a meta-analysis if appropriate, the safety and efficacy of raloxifene compared with placebo and other drugs approved in Canada for the primary and secondary prevention of osteoporotic vertebral fractures in postmenopausal women. These include estrogen alone in women with hysterectomies, estrogen-progestin therapies, bisphosphonates (alendronate, risedronate and etidronate) and salmon calcitonin. Vertebral fracture was identified *a priori* as the primary outcome measure of efficacy, with bone mineral density (BMD) a secondary outcome measure. A meta-analysis was planned to evaluate these outcomes and adverse event data. Several subgroup and sensitivity analyses will evaluate the robustness and validity of the trials.

Methods

Published and unpublished randomized controlled trials (RCTs) were identified via electronic and manual searches. Trials were evaluated for methods-related quality and synthesized qualitatively and quantitatively.

Results

Seventeen RCTs met the eligibility criteria, including six that compared raloxifene to placebo, four versus estrogen-progestin therapies; two versus estrogen alone (women with hysterectomies), one versus alendronate, one versus tamoxifen and three versus estrogen alone (no hysterectomy). No RCT compared raloxifene to etidronate, risedronate or salmon calcitonin. Tamoxifen and estrogen alone in women without hysterectomy are not approved for osteoporosis prevention.

Postmenopausal women in secondary prevention studies were older and had spent more years post-menopause than those in primary prevention studies. Most trials used a dose of 60 mg/day, which is the recommended dosage in Canada. According to the Jadad scale, the methodological quality of the studies ranged from low to moderately high. The largest study was the Multiple

Outcomes of Raloxifene Evaluation (MORE) trial, which included primary and secondary prevention arms.

In all trials reporting the outcomes of interest, raloxifene was compared with placebo. Adverse event data were obtained from trials comparing raloxifene with placebo, estrogen or combined estrogen-progestin therapy.

In the MORE trial, raloxifene had no effect on the incidence of non-vertebral fractures such as hip fracture.

Two studies reported data on the primary outcome of vertebral fracture, although there were differences in its definition and assessment, in patient populations, in study size and in duration of treatment. As a result, the planned meta-analysis of vertebral fracture data was inappropriate because of issues related to the choice of statistical model and trial heterogeneity.

Conclusion

The best estimate of overall benefit and harm from raloxifene is derived from the results of the three-year, placebo-controlled MORE trial. In this trial, the absolute risk reduction for symptomatic vertebral fractures is 0.8% (1.4% in placebo and 0.6% in raloxifene groups, number needed to treat=125 over three years). This benefit needs to be balanced by an increase in the number of serious adverse events due to venous thrombo-embolic disease of 0.65% (0.31% in placebo and 0.96% in raloxifene, number needed to harm=154 over three years).

The largest observed incidence of vertebral fracture occurs among older postmenopausal women with vertebral fractures present at baseline. This is also the group in which raloxifene has the greatest impact. The group includes women with a range of BMD levels. The impact of raloxifene to prevent vertebral fractures in younger postmenopausal women or women with higher BMD levels but no pre-existing vertebral fractures, has not been established.

The impact of raloxifene on morbidity and mortality due to cardiovascular disease or cancer (breast or uterine) has not been established, although the evidence suggests a small decrease in the incidence of early breast cancer. A subgroup analysis of cardiovascular adverse events reported in the MORE trial indicates no significant increase in morbidity and mortality in raloxifene users as compared to placebo, after three years.

TABLE OF CONTENTS

| | |
|---|-----------|
| EXECUTIVE SUMMARY | iv |
| 1 INTRODUCTION..... | 1 |
| 1.1 Clinical Background | 1 |
| 1.2 Technology..... | 2 |
| 2 OBJECTIVES | 3 |
| 3 CLINICAL EFFECTIVENESS REVIEW | 3 |
| 3.1 Methods..... | 3 |
| 3.1.1 Trial identification | 3 |
| 3.1.2 Eligibility criteria | 4 |
| 3.1.3 Selection process..... | 4 |
| 3.1.4 Data abstraction | 4 |
| 3.1.5 Assessment of methodological quality of trials from reports | 5 |
| 3.1.6 Data synthesis and analysis of efficacy and safety data | 5 |
| 3.1.7 Publication bias..... | 5 |
| 3.2 Results..... | 6 |
| 3.2.1 Qualitative results | 6 |
| 3.2.2 Quantitative results | 13 |
| 4 DISCUSSION | 25 |
| 5 CONCLUSIONS | 29 |
| 6 REFERENCES..... | 30 |
| Appendix 1: Literature Search Strategy..... | 39 |
| Appendix 2: Additional Details Concerning Methods..... | 41 |
| Appendix 3: Tools to Assess Quality of Randomized Controlled Trials..... | 43 |
| Appendix 4: Raloxifene for Primary and Secondary Prevention of Osteoporosis in Postmenopausal Women..... | 44 |
| Appendix 5: Classification of Trials by Bone Mineral Density of Participants at Baseline..... | 49 |

1 INTRODUCTION

1.1 Clinical Background

Osteoporosis is defined as “a progressive systemic disease characterized by low bone density and micro-architectural deterioration of bone tissue, with a consequent increase in bone fragility and susceptibility to fracture.”¹ Its impact is a great personal and collective cost to Canadians.²⁻⁴ In 1996, 9,382 patients were hospitalized in Ontario for hip fracture.³ In 1993, the cost of illness in Canada due to osteoporosis and its complications was estimated at \$1.3 billion.⁴ As of 2001, the annual burden of care related to hip fractures in Canada was estimated at \$650 million.⁵ Hip fracture rates increase exponentially with age. The estimated annual incidence is 4.8 per 10,000 Canadian women aged 70, increasing to 12.5 per 10,000 of those aged 80.⁶

The focus of our research is symptomatic (clinical) and asymptomatic vertebral fracture, which is the most common morphometric change noted on x-rays of postmenopausal women. These changes are often seen 10 to 15 years after natural or surgical menopause.^{1,3} The population with vertebral deformities tends to be younger on average than women who suffer hip fractures, which usually occur after age 75.

The clinical impact of vertebral fracture varies, with some people experiencing mild discomfort and some who are asymptomatic.⁷ Asymptomatic vertebral fracture is often identified after radiographic confirmation of a subsequent symptomatic vertebral fracture.^{7,8} Vertebral fractures can result in back pain and decreased quality of life.⁹⁻¹¹ Women who experience vertebral fractures are at high risk to sustain another within the next 12 months.¹² In addition, vertebral fractures are a marker or risk factor for future hip fractures.¹³ With an aging Canadian population, vertebral fractures will become more common in the future.¹⁴

The crucial clinical manifestations of osteoporosis are symptomatic vertebral fracture and symptomatic non-vertebral fractures, particularly hip fracture. Reflecting a combination of bone density and bone quality (i.e., bone geometry; bone architecture and connectivity; remodelling status; accumulated micro-damage), bone strength influences the propensity to fracture.¹⁵ Since it is impossible to directly assess bone quality, bone mineral density (BMD) is typically used to assess the fracture risk associated with bone aging,¹⁵ most commonly measured with dual energy x-ray absorptiometry (DXA).¹⁶

BMD measurements are often interpreted according to guidelines produced by a study group on densitometry hosted by the World Health Organization (WHO) in 1993.¹⁷ The guidelines define osteoporosis as a BMD T-score (i.e., a standardized score estimated from sample parameters) of greater than 2.5 standard deviations (SD) below the average value for peak young adult bone mass. In this model, women with very low BMD are considered to be in need of treatment, presumably for their low BMD level. Osteopenia is defined as a T-score between 2.5 and 1 SD below peak bone mass; and “healthy” as a BMD value less than 1 SD below this reference point.^{17,18}

Labelling women as either having or not having osteoporosis, based on BMD measurements alone, is problematic, given the potential for misclassification. Other risk factors for osteoporotic fractures include low body weight, family history of osteoporotic fracture, history of fragility fracture, history of falls and increasing age, which is the dominant clinical risk factor.¹⁹⁻²¹ BMD is an important risk factor when it is combined with age over 65.¹⁹ All risk factors, including loss of bone mass, increase with age, but their association with fractures increases more rapidly starting at menopause.^{22,23} The absolute risk of the most clinically important fragility fracture – hip fracture – is low at menopause and occurs almost exclusively in women over age 70.^{19,23,24}

Raloxifene is prescribed to women identified as having osteoporosis through an incident non-vertebral fracture, a symptomatic vertebral fracture, an asymptomatic vertebral fracture found on a routine x-ray, a low BMD level or a combination of risk factors. Raloxifene is not used to treat the incident fracture or vertebral fracture. Rather, it is used to prevent the next incident vertebral fracture. Thus, the term “treatment” is not used in this report, because raloxifene and other anti-resorptive drugs for osteoporosis do not contribute to the healing of fractures. Raloxifene therapy is characterized as therapy for the “primary” or “secondary” prevention of fractures. Primary prevention is used to categorize patients if no vertebral fracture is apparent at baseline. Secondary prevention is used if vertebral fracture is present.

1.2 Technology

According to the federal government’s approved product monograph, raloxifene (Evista[®]: Eli Lilly) is a member of the benzothiophene class of compounds known as the selective estrogen receptor modulators (SERM). According to the product monograph, raloxifene has estrogen agonist effects on bone and lipid metabolism concomitant with an estrogen antagonist effect on uterine and breast tissue.²⁵⁻²⁷ SERMs can decrease the resorption of bone tissue by osteoclasts.²² This inhibits the process by which bone is lost.

According to the product monograph, raloxifene exerts various effects on the three-dimensional conformation of the ligand-bound receptor. This accounts for its agonist properties on some tissues and its antagonistic properties with respect to others. It acts antagonistically as a competitive inhibitor of 17-beta-estradiol for the estrogen receptor. It forms estrogen receptor-beta ligand complexes distinct from estrogen, with the estrogen alpha-receptor binding to a raloxifene response element; and codes for transforming growth factor (TGF)-beta, which is a regulator of bone remodelling.

Raloxifene is promoted as a long-term approach to primary and secondary prevention of osteoporosis in postmenopausal women.²⁶ It has been approved in Canada for these two purposes (described as prevention and treatment of osteoporosis in the product monograph),^{16,27,28} with a recommended dose of 60 mg/day.

2 OBJECTIVES

The focus of this systematic review is to evaluate the safety and efficacy of raloxifene for the primary and secondary prevention of symptomatic and asymptomatic osteoporotic vertebral fractures. The required level of evidence is the randomized trial (RCT), given its status as the gold standard for research on the efficacy of health care interventions.²⁹ No restrictions are placed on the type of control or comparator, thereby allowing the possible elucidation of raloxifene's absolute (versus placebo) and relative (versus other standard treatments) utility.

A meta-analysis was planned to evaluate vertebral fracture (i.e., primary outcome), BMD (secondary outcome) and adverse event data. We also planned to use a meta-analysis of primary outcome data to investigate possible population (e.g., age; number of years post-menopause; race; geographic location as a surrogate index of vitamin D or sunlight exposure) and intervention sources of clinical heterogeneity (i.e., raloxifene dose; intervention length; use of supplements). We aimed to appraise the robustness and validity of raloxifene's effect in light of trial quality, study design and publication bias.

3 CLINICAL EFFECTIVENESS REVIEW

3.1 Methods

3.1.1 Trial identification

Without restriction on the publication status or language of reports or the year of publication, several electronic databases were searched: MEDLINE[®] (1966 to February 2002 inclusive), EMBASE (1974 to July 2001 inclusive), Current Contents Search[®] (1990 to February 2002), Adis LMS Drug Alerts (1983 to April 2001), Pascal (1973 to July 2001), HealthSTAR (1975 to October 2000), Unlisted Drugs (1984 to July 1994), TOXLINE[®] (1965 to December 2000), Inside Conferences (1993 to April 1999), BioBusiness (1985 to August 1998), BIOSIS Previews[®] (1993 to April 1999) and International Pharmaceutical Abstracts (1970 to March 1999). These searches included a study design filter to capture RCTs. The topic was identified by focusing on terms in the titles, abstracts and subject fields of all records. These terms were contained in the MEDLINE search strategy described in Appendix 1. The searches also included a standard filter to capture systematic reviews and meta-analyses.

In a surrogate hand search, the Cochrane Library's Controlled Trials Register was consulted (2002, Issue 1). The Cochrane Library was also searched for extant systematic reviews and meta-analyses.

Through contact with content experts and the manufacturer of raloxifene and through trial registries such as Current Controlled Trials (2000 to March 6, 2002), attempts were made to identify published and unpublished studies (e.g., grey literature reports such as conference

proceedings or ongoing studies that are “unpublished, have limited distribution, and are not included in bibliographic retrieval systems”³⁰). The Eli Lilly Canada representative provided trial reports, trial information and data, including guidance as to which reports referred to which unique studies. The reference lists from textbooks, reviews and reports of relevant primary studies were examined in an effort to identify additional trial material.

Information was sought from the *Journal of Bone and Mineral Research* for the American Society for Bone and Mineral Research conference proceedings (1995 to 2001 inclusive); *Osteoporosis International* for material presented at the World Congress on Osteoporosis (1995 to 2001); *Maturitas* (1995 to 2001) for the European Congress on Menopause; and the proceedings of American Society of Clinical Oncology meetings (1995 to 2001). Journals such as the *Journal of Clinical Oncology* were also searched manually for potentially relevant reports (1995 to March 2002).

3.1.2 Eligibility criteria

A trial was eligible for inclusion if it met each of the following criteria:

- an RCT involving the use of raloxifene of any dose compared to other drugs or placebo for postmenopausal women with low BMD (i.e., a T-score of more than 2.5 SD below the average value for peak young adult bone mass) or women with higher BMD levels (above the 2.5 SD threshold)
- the inclusion of at least one of the following outcomes: incident radiographically confirmed vertebral fractures (primary outcome); BMD (secondary outcome), because of its key role in the clinical diagnosis of osteoporosis and despite its poor predictive value for future fractures;²⁴ all reported adverse events (e.g., breast cancer, venous thromboembolic events, hot flashes).

All definitions of “postmenopausal” were included, e.g., natural or surgical (post-hysterectomy). We accepted trials that involved co-therapies such as calcium, vitamin D or exercise.

We excluded studies whose focus was the impact of raloxifene on biochemical markers of bone turnover, bone remodelling kinetics or bone histomorphometry; quality of life; cardiovascular risk (e.g., C-reactive protein, serum lipoprotein or plasma homocysteine levels); cognitive function; and neuromuscular function (e.g., balance or falls).

3.1.3 Selection process

The selection process by which evidence was organized and evaluated for relevance involved many steps,³¹ which are described in Appendix 2.

3.1.4 Data abstraction

After a calibration exercise involving five studies, two independent reviewers (HMS, TC) abstracted the content of each included trial using a form that focused on the report (language of publication, year of publication, published versus unpublished status and sources of funding); trial (design, number of centres); population (sample size, age, years post-menopause, country in which the study was conducted); intervention characteristics (intervention length; raloxifene

dosing schemes; types, magnitudes and frequencies of control or comparator interventions); and co-interventions (e.g., supplements). Primary (vertebral fracture incidence), secondary (BMD) and all reported adverse event outcome data (e.g., breast cancer, venous thromboembolic events, hot flashes)²⁵⁻²⁷ were also documented. Consensus was used to resolve disagreements. The original reports were not masked, as there was conflicting evidence regarding the benefit of this practice.³²

3.1.5 Assessment of methodological quality of trials from reports

The strategy by which the methodological quality of trials was assessed^{33,34} is described in Appendix 2. The instruments are listed in Appendix 3.

3.1.6 Data synthesis and analysis of efficacy and safety data

Report and trial characteristics were summarized qualitatively. Qualitative and quantitative evaluations were performed separately when different controls or comparators were involved.

Statistical analyses followed the intention-to-treat principle, focusing on data collected during the last follow-up visit at which participants were receiving the intervention. Conventions relating to data synthesis and analysis³⁵⁻⁴⁰ are presented in Appendix 2.

Sensitivity and subgroup analyses were planned to investigate possible sources of clinical heterogeneity in primary efficacy data. The following potential effect modifiers were to be tested: age; number of years post-menopause; race; geographic location of a study as a surrogate index of vitamin D (i.e., sunlight) exposure; dosage (i.e., low: ≤ 60 mg/day; high: ≥ 120 mg/day; combined: all doses); duration (e.g., with 12 months considered to be the minimum length needed to test efficacy and safety) and the use of vitamin D and calcium supplements.^{16,28,41} We also intended a priori to evaluate whether the features of the Multiple Outcomes of Raloxifene Evaluation (MORE) trial changed the meta-analytic picture of raloxifene's efficacy: i.e., changing the therapeutic guidelines after year 3 (the primary trial endpoint). In year 4, participants were allowed to take additional bone-active agents other than oral estrogen.

Sensitivity analyses planned for primary outcome data included trial design and trial quality defined using the Jadad score. The Jadad score is a numerical quantity assigned to a trial based on the presence of defined characteristics (Appendix 3) (i.e., high versus low) and on the adequacy of concealment of allocation to trial arms. The only planned subgroup analysis involving safety data focused on the impact of dosage. Additional conventions are described in Appendix 2.

3.1.7 Publication bias

Publication bias is the tendency to preferentially publish statistically positive results.⁴² Methods^{43,44} to deal with it are described in Appendix 2.

3.2 Results

3.2.1 Qualitative results

Appendix 4 lists every report regardless of publication status; or whether they are duplications or describing non-overlapping information or data (e.g., efficacy versus safety). More than one report can refer to a given study (Figure 1). Qualitative summaries of reports and trials are included in Tables 1 to 3. All studies use a parallel group design.

a) Report characteristics

Of 486 citations included after screening, 262 were excluded and 224 were considered potentially relevant. The full reports for the latter were retrieved and assessed for relevance (Figure 1).⁴⁵ Of these, 159 reports were excluded, leaving 65 reports describing 17 unique RCTs that were entered into a qualitative assessment and eligible for the meta-analysis. Reasons for exclusion at each stage of the systematic review are presented in Figure 1. One report was never found.⁴⁶

Fifteen reports were excluded because they integrated data from at least one trial already included in the systematic review; or from at least one trial whose identity and key population or intervention parameters were insufficiently defined to permit the extrapolation of its results. One included report⁴⁷ involved two identical trials⁴⁸ and described one study undertaken at two sites.

Most trials were described in at least one published report [70.6% (12/17)] (Table 3) and all reports were found in English-language journals between 1996 and March 2002 inclusive. Eli Lilly, Canada's representative confirmed that the company had sponsored 12 of the 17 included trials.

Given the multiple reports for several trials, we refer to each trial by using the first author's name (Appendix 4) (for example, Lufkin or Meunier). The exception is the Multiple Outcomes of Raloxifene Evaluation, which is known as the MORE trial.

b) Population and trial characteristics

The studies were categorized according to baseline BMD as defined a priori:

- category 1: trials with all participants with very low BMD that is more than 2.5 SD below peak young adult level (n=2) (MORE,⁴⁹⁻⁵⁸ Lufkin⁵⁹)⁶⁰⁻⁶³
- category 2: trials with fewer than all participants with very low BMD (i.e., some had BMD less than 2.5 SD below the peak levels) (n=6)⁴⁹⁻⁷⁵
- category 3: trials with all participants with higher BMD (i.e., less than 2.5 SD below peak levels) (n=11)^{47,48,75-110}
- category 4: trials with fewer than all participants identified as having higher BMD that is less than 2.5 SD below peak levels (i.e., some women had very low BMD) (n=15)^{47,48,64-75,75-110}
- category 5: mixed trials with participants having a range of BMD levels (n=4) (Appendix 4, Tables 1 to 3).⁶⁴

Figure 1: Progress through stages of systematic review

(More than one report of any publication status can refer to a given trial; RCT=randomized trial.)

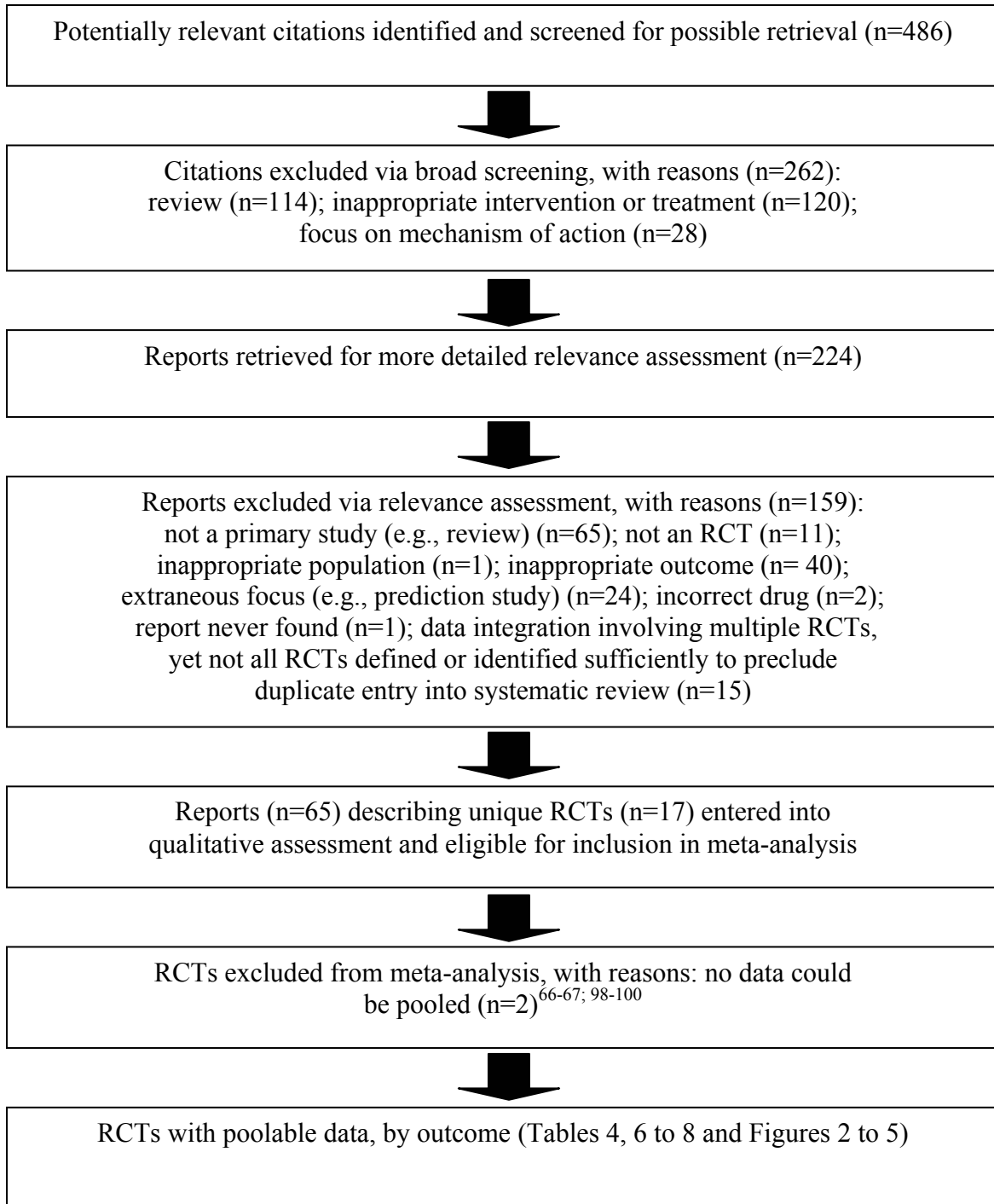


Table 1: Population characteristic defining the included raloxifene studies by trial type based on prevalent vertebral fractures and BMD at baseline

| | Sample Size, Mean (SD), Range | Age, Mean (SD), Range in Years | Number of Years Post- Menopause Mean (SD), Range | Primary Diagnostic Criterion Used to Identify Population (Mode) | Geographic Distribution |
|---|--|---|---|--|------------------------------------|
| Secondary prevention trials, each with all pts osteoporotic* (n=2) ⁴⁹⁻⁶³ | 3,924 (5,347) 143 to 7,705 | 67.45 (1.34) 66.5 to 68.4 | 20.95 (2.33) 19.3 to 22.6 | BMD or VFx (2) | North America and Europe |
| Secondary prevention trials, some with fewer than all pts osteoporotic* (n=6) ⁴⁹⁻⁷⁵ | 1,408.5 (3,086.1) 51 to 7,705 | 65.1 (6.6) 56 to 75 (n=5/6) | 13.6 (7.9) 2 to 22.6 (n=5/6) | BMD or VFx (3), BMD (3) | North America and Europe |
| Primary prevention trials, each with all pts osteopenic** or “healthy” (n=11) ^{47,48,75-110} | 316.4 (326.9) 33 to 1,145 | 56.0 (4.4) 52.9 to 59.3 (n=10/11) | 5.9 (3.1) 2 to 11 (n=7/11) | “Healthy” (7) | North America, Europe and Asia |
| Primary prevention trials, some with fewer than all pts osteopenic** or “healthy” (n=15) ^{47,48,64-75,75-110} | 272.2 (292.2) 33 to 1,145 | 57.0 (4.8) 52.9 to 75 (n=13/15) | 7.2 (4.8) 2 to 18.2 (n=10/15) | “Healthy” (7), BMD (7) | North America, Europe and Asia |
| Mixed primary and secondary prevention trials, each with osteoporotic and non-osteoporotic (i.e. osteopenic or “healthy”) pts (n=4) ⁶⁴⁻⁷⁵ | 150.8 (123.7) 51 to 330 | 63.9 (8.2) 56 to 75 (n=3/4) | 9.9 (7.0) 2 to 18.2 (n=3/4) | BMD (4) | Europe and Asia |

SD=standard deviation; pts=participants; BMD=bone mineral density values (e.g. WHO definitions); VFx=prevalent vertebral fracture.

*Osteoporotic defined as BMD<2.5 SD below young adult mean.

**Osteopenic defined as 2.5 SD<BMD<1 SD below mean.

Table 2: Intervention characteristics defining included raloxifene studies, by trial type

| | Intervention Length: Weeks Mean (SD), Range* Trials Lasting ≥12 months | Single RLX Dose mg/day (Mode) | RLX Dose Level Contrast mg/day (Mode) | Comparator (Mode) | Intervention: RLX (mg/day) versus Control Contrast (Mode) | Supplements Mandated Number (%); Type (Mode) |
|--|---|--|--|-----------------------------------|--|---|
| Secondary prevention trials, each with all pts osteoporotic* (n=2) ⁴⁹⁻⁶³ | 111.4 (83.9) 52 to 170.7* n=2 | RLX 60 (2), RLX 120 (2) | RLX 60 versus RLX 120 (2) | PB (2) | RLX 60 versus RLX 120 versus PB (2) | 2 (100) Calcium and vitamin D (2) |
| Secondary prevention trials, some with fewer than all pts osteoporotic (n=6) ⁴⁹⁻⁷⁵ | 69.5 (58.9) 12 to 170.7* n=4 | RLX 60 (4) | RLX 60 versus RLX 120 (2) | PB (7) | RLX 60 versus RLX 120 versus PB (2) | 5 (83.3) Calcium and vitamin D (4) |
| Primary prevention trials, each with all pts osteopenic** or “healthy” (n=11) ^{47,48,75-110} | 62.6 (47.9) 8 to 156* n=7 | RLX 60 (9) | RLX 60 versus RLX 150 (4) | PB (6) ERT (5) | RLX 60 versus RLX 150 versus ERT (2) | 5 (45.5) Calcium alone (5) |
| Primary prevention trials, some with fewer than all pts osteopenic or “healthy” (n=15) ^{47,48,64-75,75-110} | 58.9 (45.1) 8 to 156* n=9 | RLX 60 (12) | RLX 60 versus RLX 150 (5) | PB (9) ERT (6) | RLX 60 versus RLX 150 versus ERT (2) | 8 (53.3) Calcium alone (6) |
| Mixed primary and secondary prevention trials, each with osteoporotic and non-osteoporotic (i.e. osteopenic or “healthy”) pts (n=4) ⁶⁴⁻⁷⁵ | 48.5 (40.5) 12 to 104* n=2 | RLX 60 (3) | All unique | PB (3) | All unique | 3 (75) Calcium and vitamin D (2) |

SD=standard deviation; pts=participants; RLX=raloxifene; PB=placebo; ERT=estrogen replacement therapy (CEE 0.625 mg/day); mode=measure of central tendency.

*Osteoporotic defined as BMD<2.5 SD below young adult mean.

**Osteopenic defined as 2.5 SD<BMD<1 SD below mean.

Table 3: Trial characteristics defining included raloxifene studies, by trial type

| | Jadad Total Trial Quality Score, Mean (SD) Range | Trials with Jadad Total Trial Quality Score of 0 to 2 (Low) Number (%) | Trials with Maximum Jadad Total Trial Quality Score (=5), Number (%) | Trials with “Adequate” Allocation Concealment, Number (%) | Trials with “Unclear” Allocation Concealment, Number (%) | Trials Described by ≥1 Published Report,* Number (%) | Confirmed Eli Lilly Trials,** Number (%) |
|--|---|---|---|--|---|---|---|
| Secondary prevention trials, each with all pts osteoporotic* (n=2) ⁴⁹⁻⁶³ | 3.5 (0.71) 3 to 4 | 0 (0) | 0 (0) | 1 (50) | 1 (50) | 2 (100) | 2 (100) |
| Secondary prevention trials, some with fewer than all pts osteoporotic (n=6) ⁴⁹⁻⁷⁵ | 2.7 (1.03) 1 to 4 | 2 (33.3) | 0 (0) | 1 (16.7) | 5 (83.3) | 4 (66.7) | 5 (83.3) |
| Primary prevention trials, each with all pts osteopenic** or “healthy” (n=11) ^{47,48,75-110} | 3.3 (1.10) 2 to 5 | 3 (27.3) | 2 (18.2) | 1 (9.1) | 10 (90.9) | 8 (72.7) | 8 (72.7) |
| Primary prevention trials, some with fewer than all pts osteopenic or “healthy” (n=15) ^{47,48,64-75,75-110} | 3.0 (1.13) 1 to 5 | 5 (33.3) | 2 (13.3) | 1 (6.7) | 14 (93.3) | 10 (66.7) | 11 (73.3) |
| Mixed primary and secondary prevention trials, each with osteoporotic and non-osteoporotic (i.e. osteopenic or “healthy”) pts (n=4) ⁶⁴⁻⁷⁵ | 2.3 (0.96) 1 to 3 | 2 (50) | 0 (0) | 0 (0) | 4 (100) | 2 (50) | 3 (75) |

SD=standard deviation; pts=participants; *more than one report can refer to a given trial (Table 1); **confirmed by Eli Lilly’s Canadian representative.

*Osteoporotic defined as BMD<2.5 SD below young adult mean.

**Osteopenic defined as 2.5 SD<BMD<1 SD below mean.

These categories were not mutually exclusive. A trial could be included in more than one category, and categories 2 and 5 are subsets of categories 1 and 4 respectively.

For four trials, the patient populations were too heterogeneous to classify according to BMD criteria. One trial⁶⁴⁻⁶⁷ had 40% of participants at baseline with BMD levels greater than the threshold (2.5 SD below the peak young adult level). In two other studies, the BMD inclusion criterion was a level >2 SD below the mean value for young adults,^{68-71,74} As a result, both studies included women who were unclassifiable in this scheme. One study⁷³⁻⁷⁵ was not classifiable using the above criteria, because the BMD inclusion criterion was established between 3 SD below and 1 SD above the mean value.

Details of the similarities and differences in the trials for the five BMD categories are presented in Appendix 5. Because of the a priori plan to categorize trials by baseline BMD, the tables retain an organization based on five categories of baseline BMD. This system is not used in the main analysis, because most of the trials, with the exception of MORE and Luftkin, do not report vertebral fractures as an outcome. BMD, however, is retained as it is an important baseline patient characteristic.

Nine of the 17 trials that were identified⁴⁷⁻⁸³ included efficacy or safety data that could be pooled while the remaining eight^{75,84-110} reported efficacy data outside the focus of this review while providing relevant safety data. Each of the latter was a study of women with higher BMD levels. One study^{82,83} excluded explicit inclusion criteria and baseline population data.

Trials enrolling women with very low BMD differed from those that included women with higher BMD levels in terms of age and number of years post-menopause. Regarding age, there were trials with all women having very low BMD (mean=67.5 years) or fewer than all (65.1) participants with very low BMD versus trials with all women having higher BMD (56.0) or fewer than all with higher BMD (57.0). Regarding number of years post-menopause, there were trials with all women having very low BMD (mean=21 years) or fewer than all (13.6) participants with low BMD versus trials with all women having higher BMD (5.9) or fewer than all (7.2) participants with higher BMD (Table 1).

In addition to population-defined clinical heterogeneity, two^{68-70,101} and five studies^{68-70,81-83,101,108-110} respectively did not report data on the age of participants or the number of years post-menopause. Mixed BMD studies involved participants who were almost as old (63.9 years) as those in very low BMD studies; and who fell between purely low and purely higher BMD studies in the number of years post-menopause (9.9 years).

Both trials that reported vertebral fracture, the necessary outcome for inclusion in the meta-analysis, enrolled patients based on its prevalence and BMD levels. Three of the trials that included fewer than all participants with very low BMD (3/6) enrolled patients based on prevalent vertebral fracture, but they did not report it as an outcome. Three studies with mixed high and low BMD (3/6) used only BMD as an entry criterion (Table 1).

Most trials occurred in North America and Europe (Appendix 4, Table 1). There was insufficient information that could be used to summarize the racial composition of the samples.

c) Intervention characteristics

Length of follow-up

Data from the MORE trial described three and four years of follow-up. The primary predetermined endpoint was three years.⁴⁹ Overall, 11 trials lasted at least 12 months.^{47-70,75-93,97-100,108-110} The trial type with the largest proportion of studies lasting more than 12 months was type I, which included those with all participants having very low BMD (2/2). Trials that included all (mean=111.4 weeks) or fewer than all participants (69.5 weeks) with very low BMD entailed longer therapy durations than both types of studies with higher BMD (62.6, 58.9 and 48.5 weeks for all, fewer than all and mixed BMD levels respectively) (Table 2).

Dose

There was consistency across all trials regarding the most popular single dose of raloxifene: 60 mg/day (Table 2), which is the approved dose in Canada and elsewhere. The most popular range of raloxifene doses was 60 versus 120 mg/day (mode for each=2) and 60 versus 150 mg/day for low BMD and higher BMD trials (modal values=4 and 5) respectively. Each trial with mixed BMD levels among participants at baseline used a unique range of raloxifene doses.

Comparator

The 17 identified RCTs include six versus placebo alone, four versus estrogen-progestin therapies; two versus estrogen alone (women with hysterectomies); one versus alendronate, one versus tamoxifen and three versus estrogen alone (no hysterectomy). No RCTs compared raloxifene to etidronate, risedronate or salmon calcitonin. Two active comparators tamoxifen and estrogen alone in women without hysterectomy are unapproved for osteoporosis prevention. Estrogen therapy was often used as a comparator in studies of women with higher BMD at baseline.

Supplements

Studies varied in terms of the types of mandated supplements. Calcium (range 500 to 1,000 mg/day) with vitamin D (range 300 to 800 IU/day) and calcium alone (range 400 to 600 mg/day) were given most often (Table 2). In one study, participants were asked not to take supplements.^{82,83} One trial involving patients with very low BMD⁷³⁻⁷⁵ and five trials involving women with higher BMD^{75,84-90,97-107} did not report whether supplements were used.

d) Trial quality

The mean Jadad total quality scores were comparable for trials with all participants diagnosed with very low BMD (mean=3.5) and those with low and higher BMD (mean=3.3) respectively. These indicated moderately high trial quality. Trials with mixed BMD levels had total scores that barely exceeded low quality (2.3). This trial type received a low Jadad total quality score (0 to 2) 50% (2/4) of the time. Trials involving women with higher BMD in either category were the only ones receiving a maximum or high (5) quality score. Two trials provided evidence that allocation to the intervention arms had been performed adequately.^{49-58,91-93} Most trials (15/17)

were rated as having unclear allocation concealment. Of the two trials reporting vertebral fracture as an outcome, the MORE trial was larger and exhibited higher methodological quality than the Lufkin trial (Appendix 4).

3.2.2 Quantitative results

The trial populations varied in terms of age and necessarily, the number of years post-menopause. Based on these two variables, we had planned subgroup analyses of efficacy (i.e., clinical fractures, vertebral fracture, BMD) and safety (mortality, serious adverse events, other adverse events). While 17 relevant trials were included in this systematic review, two provided data that were suitable for the planned meta-analysis, two provided data on the primary outcome of vertebral fracture, while others provided data on mild to moderate adverse effects.

a) Results of Main Analyses of Primary Outcome Data Symptomatic and asymptomatic vertebral fractures

Three studies provided radiological and clinical data on vertebral fracture, all comparing raloxifene to placebo: the Meunier, MORE and Lufkin trials (Appendix 4, Table 4, Figure 2). The Meunier trial⁶⁴⁻⁶⁷ failed to report any vertebral fracture events, so it was excluded from the meta-analysis, leaving the MORE and Lufkin trials.

Using the meta-analyzed vertebral fracture results (MORE and Lufkin data), no benefit was seen in raloxifene recipients (for combined, low and high doses), even though a significant risk reduction was observed for all dose definitions when the MORE trial's data at 40 or at 47.4 months were analyzed separately. Results obtained exclusively from the MORE trial showed a significant relative risk reduction (RRR) after three years [0.59 (95% CI 0.51, 0.70)] and four years of intervention [0.60 (0.52, 0.69)].

Given that the results of only two trials could be pooled, likely yielding an unreliable estimate of variation (and there was evidence of significant statistical heterogeneity), the random effects model was chosen as the correct statistical model. This model smoothed the weightings so that the larger trial's contribution to the pooled result carried less weight.

While we report our findings in keeping with the pre-determined methodology, there is little justification for using a random effects model or for undertaking meta-analysis. The two trials that can be meta-analyzed differ in terms of number of randomized participants, patterns of incident vertebral fractures and trial quality when defined in two ways (Appendix 4). Thus, by virtue of its greater sample size and higher quality, the MORE study should "drive" the meta-analysis via a fixed effects model. Statistically, by using a fixed effects model, the larger trial (MORE: n=7,705) receives a larger weighting than a smaller trial (Lufkin: n=143). As a result, evidence from the MORE trial regarding raloxifene's placebo-controlled efficacy in the prevention of vertebral fractures¹⁴ is revealed.

Certain factors may have played a role in producing the discordance across the vertebral fracture-related results between the Lufkin and MORE trials, with the most important factor being the definition of the vertebral fracture outcome.¹⁴ In the MORE trial, a vertebral fracture required confirmation via at least two of three types of assessments: two independent semi-quantitative readings (i.e., grade 0=none, 1=mild, 2=moderate, 3=severe) and one quantitative assessment

(i.e., fracture=decrease of at least 20% and 4 mm via radiographic assessment). The Lufkin study defined an incident vertebral fracture as either a $\geq 15\%$ or a $\geq 30\%$ decrease in vertical height respectively. The results were not significant at the $\geq 15\%$ cutoff.

Table 4: Primary outcome and subgroup analyses of vertebral fracture data against placebo

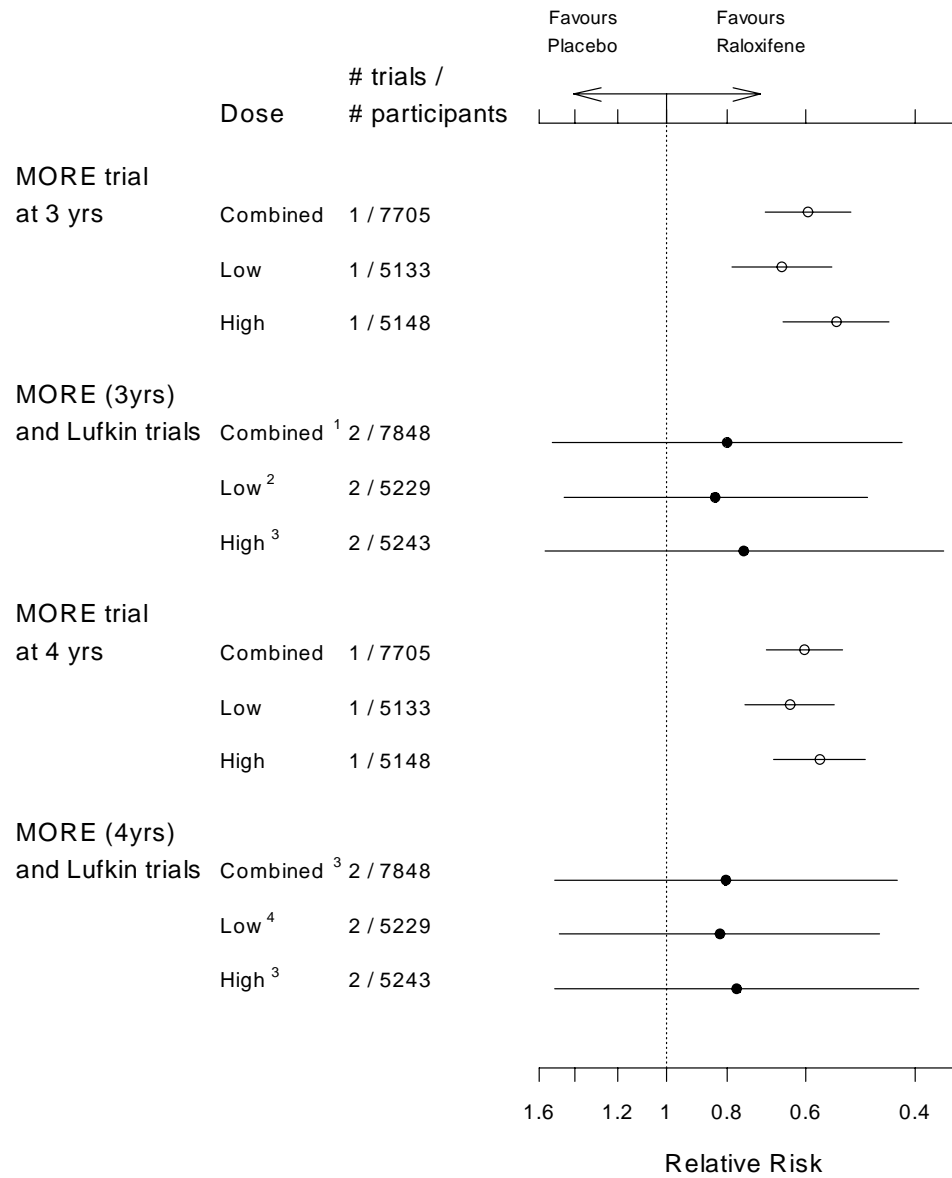
| Dose | Number of Trials/ Number of Participants | Relative Risk of (95% CI)* Vertebral Fractures | Statistical Heterogeneity |
|----------|--|--|------------------------------|
| Combined | 1/7,705 ¹ | 0.59 (0.51, 0.70)**** | n/a |
| Low | 1/5,133 ¹ | 0.65 (0.54, 0.79)**** | n/a |
| High | 1/5,148 ¹ | 0.54 (0.44, 0.65)**** | n/a |
| Combined | 2/7,848 ² | 0.80 (0.42, 1.53)***** | p=0.00 |
| Low | 2/5,229 ² | 0.84 (0.48, 1.46)***** | p=0.03 |
| High | 2/5,243 ² | 0.75 (0.36, 1.57)***** | p=0.01 |
| Combined | 1/7,705 ³ | 0.60 (0.52, 0.69)**** | n/a |
| Low | 1/5,133 ³ | 0.64 (0.54, 0.75)**** | n/a |
| High | 1/5,148 ³ | 0.57 (0.48, 0.68)**** | n/a |
| Combined | 2/7,848 ⁴ | 0.81 (0.43, 1.51)***** | p=0.01 |
| Low | 2/5,229 ⁴ | 0.82 (0.46, 1.48)***** | p=0.02 |
| High | 2/5,243 ⁴ | 0.77 (0.40, 1.51)***** | p=0.01 |

n/a=not applicable; CI=confidence interval; combined=(low dose=RLX \leq 60 mg/day) plus (high dose=RLX \geq 120 mg/day); *for relative risk, a CI encompassing a value of 1 indicates statistically non-significant result; **random effects model; ***excluding Meunier trial⁶⁴⁻⁶⁷ since each arm of study had zero events; ****fixed effects model; ¹MORE trial at three years; ²MORE trial at three years plus Lufkin trial; ³MORE trial at four years; ⁴MORE trial at four years plus Lufkin trial.

The investigators claimed that with a $\geq 30\%$ cutoff, there was a dose-related reduction in vertebral fracture for the raloxifene groups versus placebo. Neither the Lufkin trial reports nor this systematic review team's contact at the manufacturer provided $\geq 30\%$ cutoff data expressed in a manner (i.e., number of participants with at least one incident vertebral fracture) amenable to pooling with MORE data. The $\geq 30\%$ cutoff seemed to be a post hoc measure of questionable validity, given the lack of difference observed at the predefined $\geq 15\%$ cutoff.

A second factor that likely led to discordance between the vertebral fracture results in the Lufkin and MORE trials was the differences in study populations (Appendix 4). The subjects in the Lufkin trial were older than the women in the largest subgroup enrolled in the MORE trial: women without baseline vertebral fracture (mean of 68 years in the Lufkin trial versus 65 years in this subgroup). There was little difference in age between the patients in the Lufkin trial and those in the MORE subgroup with pre-existing vertebral fractures. Women in the MORE trial with vertebral fractures at baseline were significantly more likely to experience two or more incident vertebral fractures over three years. This supported the view that a history of vertebral fractures was an independent risk factor for future fractures, regardless of BMD.^{12,17,20,111-114} The Lufkin trial did not report on the proportion of women with baseline vertebral fracture, so it was impossible to know whether the populations were similar in this respect.

Figure 2: Risk of vertebral fracture against placebo



Combined=(low dose=RLX≤60 mg/day) plus (high dose=RLX≥120 mg/day); open circle=fixed effects model; solid circle=random effects model; p values for statistical heterogeneity: ¹p=0.00, ²p=0.03, ³p=0.01, ⁴p=0.02.

Both studies used a low dose defined as 60 mg/day and a high dose not exceeding 120 mg/day. The higher doses did not provide an advantage over what was observed at the 60 mg/day dose. Both trials included calcium and vitamin D supplements, but these were provided at different doses (Table 2).

In light of these definitions of clinical heterogeneity and the observation that the MORE trial was larger than and exhibited higher methods-related quality than did the Lufkin trial (Appendix 4), additional attention was paid to MORE results and additional data pertaining to the primary outcome were described, including findings reported by reviewers at the US Food and Drug Administration (FDA).

Key efficacy outcomes: MORE trial

The MORE trial was of sufficient size and duration so that it could assess the clinical outcomes of raloxifene therapy, including symptomatic vertebral fracture and safety outcomes, such as mortality and serious adverse events. Table 5 shows the results concerning vertebral fracture.

The MORE trial alone showed a significant relative risk reduction at three years for 60 mg raloxifene per day [0.65 (0.54, 0.79)] and 120 mg raloxifene [0.54 (0.44, 0.65)] compared with placebo; and at four years for 60 mg raloxifene [0.64 (0.54, 0.75)] and 120 mg raloxifene [0.57 (0.48, 0.68)] respectively. The greater risk reduction, compared with placebo, was associated with the higher dose of 120 mg/day. The risk reduction in both groups did not differ significantly. As with the main analyses, all dose-related analyses of pooled results were associated with statistical heterogeneity (p=0.01 to 0.03).

Table 5: Vertebral effects of raloxifene versus placebo over three years of follow-up

| | Placebo Absolute Risk (N*=2,576) | Raloxifene (pooled) Absolute Risk (N=5,129) | ARR | NNT/NNH |
|---|---|--|--------------|----------------------|
| Symptomatic (clinical) vertebral fracture (published trial report)* | 35 to 36*(1.4%) | 29 to 30 (0.6%) | 0.8% | ~125 (NNT) |
| Symptomatic (clinical) vertebral fracture (FDA medical review)† | 3.1% | 1.8% 60 mg 1.5% 120 mg | 1.3% 1.6% | 77 (NNT) 63 (NNT) |

*Published report states that total number of women with symptomatic fractures was 65, RR=0.4 (95% CI 0.3 to 0.7) for raloxifene users. Using a 2X2 table in Epi Info, this is estimated to be 35 to 36 cases in placebo group; 29 to 30 in pooled raloxifene groups (number per treatment arm not presented in report).

†FDA review reports on vertebral fracture in 2,292 women on placebo, 2,259 on raloxifene 60 mg and 2,277 on raloxifene 120 mg, reflecting number of women with x-rays at baseline and endpoint.

ARR=absolute risk reduction; ARI=absolute risk increase; NN=numbers needed to treat; NNH=numbers needed to harm.

In the MORE trial, women with prevalent vertebral fractures at baseline (study group 2) had higher rates of incident vertebral fracture than women with BMD ≥ 2.5 SD below the mean. This difference was greater than that observed between placebo and raloxifene users; at three years of follow-up, 4.5% of women with BMD ≥ 2.5 SD below the mean on placebo had experienced incident (asymptomatic or symptomatic) vertebral fracture, compared with 14.7% of women on raloxifene 60 mg/day with vertebral fracture at baseline. The magnitude of this difference lent weight to the importance of the prevalence of vertebral fracture as a predictor of future incident vertebral fractures.

There were no statistically significant differences between placebo and raloxifene users in the proportion of patients with at least one incident non-vertebral fractures: 9.3% on placebo versus 8.5% on raloxifene.

Key efficacy outcomes: Lufkin trial

Raloxifene did not significantly alter the probability of incident morphometric vertebral fracture (defined as $\geq 15\%$ reduction in height). One or more incident vertebral fractures were observed on x-ray in 40% of women on placebo and 48% of women on all doses of raloxifene (49% at 60 mg/day). The high frequency of incident vertebral fractures in one year, on placebo and on raloxifene, suggested that this patient population also had a high prevalence of vertebral fractures at baseline. There were no wrist fractures; one hip fracture occurred in a woman on raloxifene 120 mg/day.

a) Subgroup and sensitivity analyses of primary outcome data

With only two trials providing vertebral fracture data that could be pooled, we could not perform the planned subgroup and sensitivity evaluations of primary outcome data; and we did not pool the Lufkin and MORE data using samples of women who had high and low BMD at baseline.⁴⁹

b) Results of analyses of secondary outcome data

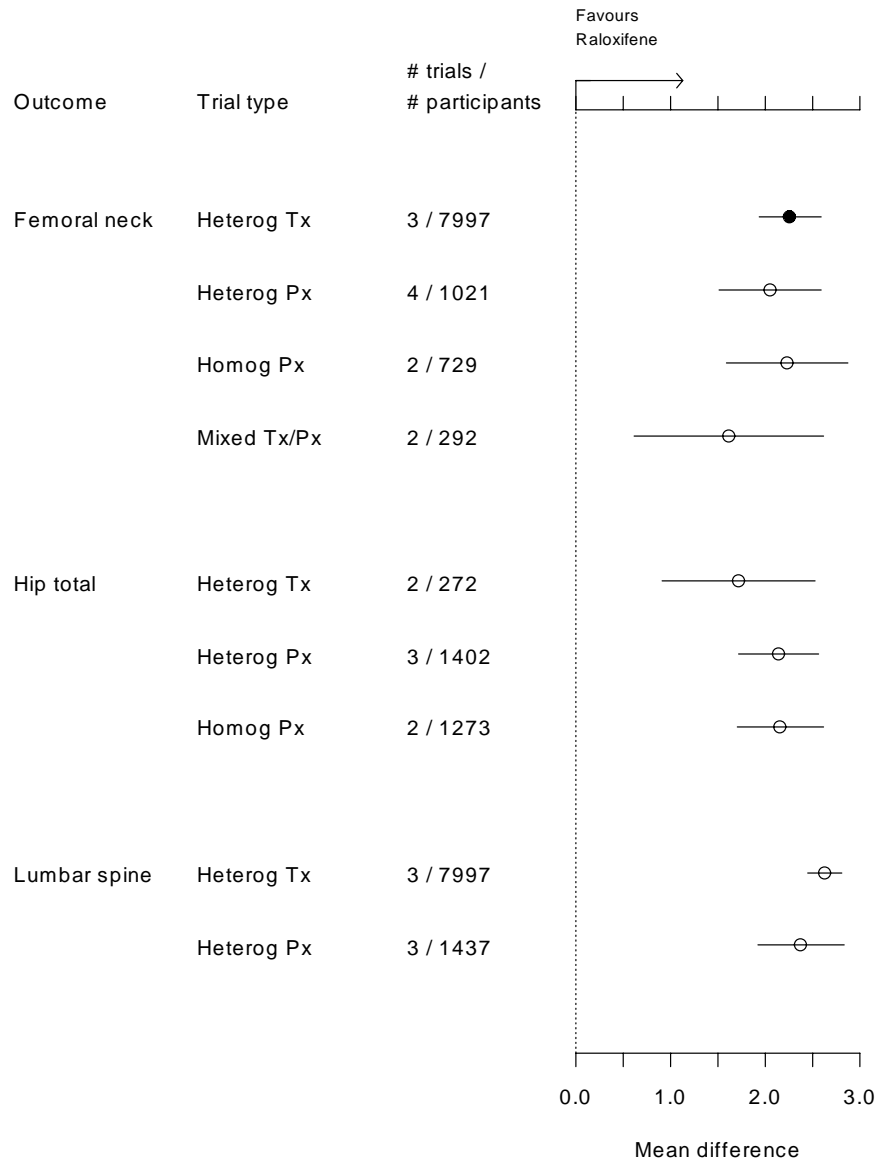
Both primary and secondary prevention studies provided analyzable BMD data. The only comparator was placebo. Raloxifene had a significant and positive impact on BMD, with some variation by degree, for all trials regardless of BMD level at enrolment (Table 6, Figure 3).

Table 6: Secondary outcome analyses of bone mineral density (BMD) data: against placebo, organized by trial type

| BMD Sites Outcome (g/cm ²) | Trial Type**+ | Number of Trials/ Number of Participants ^{trials} | Weighted Mean** (95% CI) | Statistical Heterogeneity |
|--|---------------|--|--------------------------|---------------------------|
| Femoral neck BMD | Heterog Tx | 3/7,997 ^{1,a} | 2.26 (1.93, 2.59) | p=0.34 |
| | Heterog Px | 4/1,021 ^{2,b} | 2.05 (1.51, 2.58)*** | p=0.69 |
| | Homog Px | 2/729 ³ | 2.23 (1.59, 2.86)*** | p=0.66 |
| | Mixed Tx/Px | 2/292 ^{4,b} | 1.62 (0.62, 2.61)*** | p=0.63 |
| Hip total BMD | Heterog Tx | 2/272 ⁵ | 1.72 (0.92, 2.52)*** | p=0.43 |
| | Heterog Px | 3/1,402 ⁶ | 2.13 (1.71, 2.56)*** | p=0.59 |
| | Homog Px | 2/1,273 ⁷ | 2.16 (1.70, 2.62)*** | p=0.32 |
| Lumbar spine BMD | Heterog Tx | 3/7,997 ^{1,a} | 2.62 (2.44, 2.81)*** | p=0.49 |
| | Heterog Px | 3/1,437 ^{8,b} | 2.37 (1.92, 2.82)*** | p=0.57 |

CI=confidence interval; Tx=treatment; Px=prevention; heterog Tx=Tx trials with not all pts diagnosed osteoporotic; homog Tx=Tx trials with all pts diagnosed osteoporotic; heterog Px=Px trials with not all pts diagnosed osteopenic or “healthy”; homog Px=Px trials with all pts diagnosed osteopenic or “healthy”; mixed Tx/Px=heterogeneously composed trials, which when added to trials with homogeneous composition of participants defined via WHO BMD criteria, yields designation of “heterog trials”; ¹combined=all doses=(low dose=RLX \leq 60 mg/day) plus (high dose=RLX \geq 120 mg/day); *trial type=treatment versus prevention trial, based on WHO’s BMD defined typology (i.e. osteoporotic versus osteopenic versus healthy); **difference in mean percent change from baseline; ***with little statistical heterogeneity, random effects estimate is identical to fixed effects estimate; ¹MORE trial at three years plus Meunier⁶⁴⁻⁶⁷ and Johnell⁶⁸⁻⁷⁰ trials; ²Meunier,⁶⁴⁻⁶⁷ Johnell,⁶⁸⁻⁷⁰ Johnston,^{47,48,76-80} and Pavo^{82,83} trials; ³Johnston^{47,48,76-80} and Pavo^{82,83} trials; ⁴Meunier,⁶⁴⁻⁶⁷ and Johnell⁶⁸⁻⁷⁰ trials; ⁵Lufkin⁵⁹⁻⁶³ and Meunier,⁶⁴⁻⁶⁷ trials; ⁶Meunier,⁶⁴⁻⁶⁷ Johnston^{47,48,76-80} and Pavo^{82,83} trials; ⁷Johnston at three years^{47,48,76-80} and Pavo^{82,83} trials; ⁸Meunier,⁶⁴⁻⁶⁷ Johnell⁶⁸⁻⁷⁰ and Johnston,^{47,48,76-80} trials; ^avariance imputation required for MORE trial at three years and Johnell⁶⁸⁻⁷⁰ trial; ^bvariance imputation required for Johnell⁶⁸⁻⁷⁰ trial.

Figure 3: Bone mineral density against placebo, by trial type*



*All analyses involved combined RLX doses; combined=(low dose=RLX≤60 mg/day) plus (high dose=RLX≥120 mg/day); open circle=fixed effects model; solid circle=random effects model; Tx=treatment; Px=prevention; heterog Tx=Tx trials with not all participants diagnosed osteoporotic; homog Tx=Tx trials with all participants diagnosed osteoporotic; heterog Px=Px trials with not all participants diagnosed osteopenic or “healthy”; mixed Tx/Px=heterogeneously composed trials, which, when added to trials with homogeneous composition of participants defined via WHO BMD criteria, yields designation of “heterog” trials.

A positive effect on femoral neck BMD was observed for all trial types, with the greatest increase seen in populations where most or all women had very low BMD at enrolment [2.26 (1.93, 2.59)]. The next greatest benefit occurred in populations with BMD that was uniformly higher than the 2.5 SD threshold [2.23 (1.59, 2.86)]. Trials with mixed BMD levels at baseline showed an intermediary effect [1.62 (0.62, 2.61)].

For hip BMD, raloxifene showed a statistically significant impact. The greatest increase was observed for patient populations with uniformly very low BMD [2.16 (1.70, 2.62)]. The next greatest increase was for the category of trials with low BMD mixed with trials that include higher BMD levels, Px [2.13 (1.71, 2.56)]. The weakest effect was for the heterogeneous BMD trial types [1.72 (0.92, 2.52)].

For lumbar spine BMD, raloxifene exhibited a significant, positive effect, with the strongest impact associated with the category that mixed very low BMD with higher BMD [2.62 (2.44, 2.81)]. The next greatest impact came from trials in the mixed BMD trial types [2.30 (1.27, 3.34)]. These investigations did not reveal instances of significant ($p < 0.10$) statistical heterogeneity.

c) Results of analyses of mild to moderate adverse events (tolerability)

A meta-analysis was planned to assess data on all adverse events. The most complete data were provided by the MORE trial alone. Given its significant place in the literature as defined by its size, methods-related quality and intervention length, MORE safety results were presented separately, after the results that could be pooled. This strategy established a larger perspective on raloxifene's efficacy and safety than was observed by looking exclusively at pooled results.

Meta-analysis was used according to the a priori protocol to evaluate adverse event data from the 17 included trials. The analyzable binary (Table 7, Figure 4) and continuous adverse event data (Table 8, Figure 5) involved three types of comparators (placebo, estrogen and combined hormone therapy) and four of five trial types (i.e., no mixed BMD trials). Only studies in populations of women with higher BMD levels at baseline, i.e., less than 2.5 SD below peak BMD, provided analyzable data when estrogen and combined hormone therapy were the comparators.

On three occasions, a subgroup analysis by dose yielded a finding that changed the consideration of a finding as statistically significant. Each involved binary data, a low raloxifene dose and placebo as the comparator.

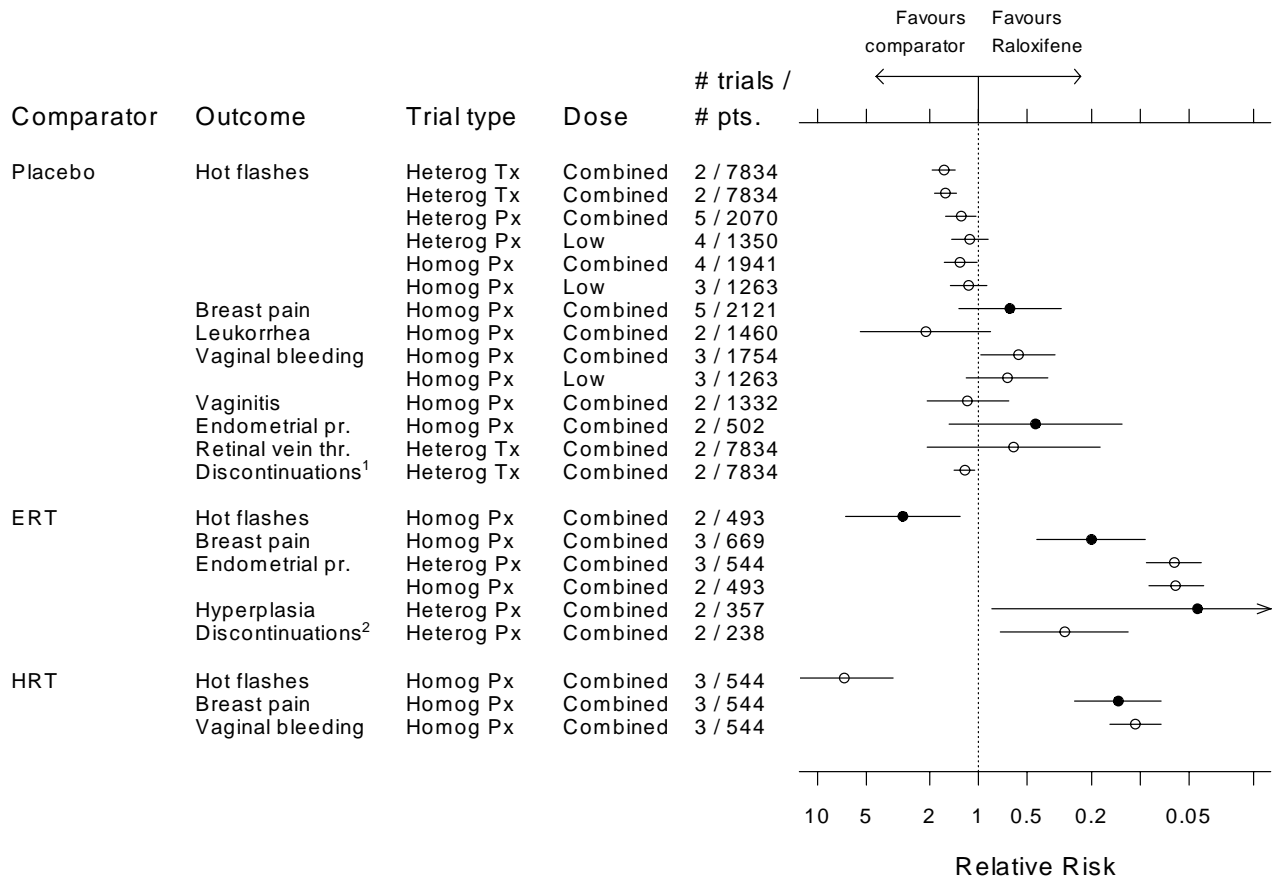
Results were presented by adverse event type; and organized by comparator and trial type. Observations from analyses of binary data preceded those involving continuous data. Assessed against each comparator (i.e., placebo, estrogen, combined hormone therapy), there was a significantly greater risk of hot flashes associated with raloxifene use (Table 7, Figure 4). For example, when compared with placebo, the greatest relative risk was seen in trials with very low BMD. The relative risk at 40 months was 1.65 (1.40, 1.94) and at 47.4 months was 1.61 (1.38, 1.88).

Table 7: Analyses of binary adverse event data;[†] organized by comparator and trial type

| Adverse Event Outcome | Trial Type* | Dose | Number of Trials/ Number of Participants ^{trials} | Relative Risk (95% CI)** | Statistical Heterogeneity |
|---|-------------|------------------|--|-----------------------------|------------------------------|
| Against placebo | | | | | |
| Hot flashes | Heterog Tx | Combined | 2/7,834 ¹ | 1.65 (1.40, 1.94)*** | P=0.57 |
| | Heterog Tx | Combined | 2/7,834 ² | 1.61 (1.38, 1.88)*** | P=0.60 |
| | Heterog Px | Combined | 5/2,070 ³ | 1.29 (1.03, 1.61)*** | P=0.93 |
| | Heterog Px | Low [†] | 4/1,350 ⁴ | 1.14 (0.88, 1.47)*** | P=0.98 |
| | Homog Px | Combined | 4/1,941 ⁵ | 1.30 (1.03, 1.63)*** | P=0.83 |
| | Homog Px | Low [†] | 3/1,263 ⁶ | 1.15 (0.88, 1.49)*** | P=0.96 |
| Breast pain and Soreness | Homog Px | Combined | 5/2,121 ⁷ | 0.64 (0.31, 1.32) | P=0.18 |
| Leukorrhea | Homog Px | Combined | 2/1,460 ⁸ | 2.13 (0.84, 5.39)*** | P=0.67 |
| Vaginal bleeding | Homog Px | Combined | 3/1,754 ⁶ | 0.57 (0.33, 0.97)*** | P=0.69 |
| | Homog Px | Low [†] | 3/1,263 ⁶ | 0.67 (0.37, 1.19)*** | P=0.46 |
| Vaginitis | Homog Px | Combined | 2/1,332 ⁹ | 1.17 (0.65, 2.11)*** | P=0.88 |
| Endometrial proliferation | Homog Px | Combined | 2/502 ¹⁰ | 0.38 (0.19, 0.79) | P=0.17 |
| Retinal vein Thrombosis | Heterog Tx | Combined | 2/7,834 ¹ | 0.61 (0.18, 2.09)*** | P=0.36 |
| Discontinuations (adverse event) | Heterog Tx | Combined | 2/7,834 ¹ | 1.22 (1.05, 1.42)*** | P=0.54 |
| Against estrogen replacement therapy | | | | | |
| Hot flashes | Homog Px | Combined | 2/493 ¹⁰ | 2.96 (1.31, 6.69) | P=0.30 |
| Breast pain and Soreness | Homog Px | Combined | 3/669 ¹¹ | 0.19 (0.11, 0.32) | P=0.24 |
| Discontinuations (no reason reported) | Heterog Px | Combined | 2/238 ¹⁴ | 0.29 (0.12, 0.73)*** | P=0.51 |
| Against hormone replacement therapy | | | | | |
| Hot flashes | Homog Px | Combined | 3/544 ¹⁵ | 6.77 (3.39, 13.52)*** | P=0.71 |
| Breast pain and Soreness | Homog Px | Combined | 3/544 ¹⁵ | 0.14 (0.07, 0.25) | P=0.30 |
| Vaginal bleeding | Homog Px | Combined | 3/544 ¹⁵ | 0.11 (0.07, 0.15)*** | P=0.75 |

[†]includes three subgroup analysis results (each for low dose) where a statistically significant “combined” finding (i.e. without distinguishing by dose magnitude) became a non-significant one when dose magnitude was investigated; CI=confidence interval; pts=participants; Tx=treatment; Px=prevention; heterog Tx=Tx trials with not all pts diagnosed osteoporotic; homog Tx=Tx trials with all pts diagnosed osteoporotic; heterog Px=Px trials with not all pts diagnosed osteopenic or “healthy”; homog Px=Px trials with all pts diagnosed osteopenic or “healthy”; mixed Tx/Px=heterogeneously composed trials, which when added to trials, with homogeneous composition of participants defined via WHO BMD criteria, yields designation of “heterog” trials; combined = all doses = (low dose=RLX≤60 mg/day) plus (high dose=RLX≥120 mg/day); *trial type = treatment versus prevention trial, based on WHO’s BMD defined typology (i.e. osteoporotic versus osteopenic versus healthy); **for relative risk, a CI encompassing a value of “1” indicates statistically non-significant result; ***with little statistical heterogeneity, random effects estimate is identical to fixed effects estimate; ¹MORE trial at three years and Meunier⁶⁴⁻⁶⁷ trial; ²MORE trial at four years and Meunier⁶⁴⁻⁶⁷ trial; ³Meunier,⁶⁴⁻⁶⁷ Johnston,^{47,48,76-80} Goldstein,⁹¹⁻⁹³ Boss⁹⁴⁻⁹⁶ and Walsh¹⁰⁵⁻¹⁰⁷ trials; ⁴Meunier,⁶⁴⁻⁶⁷ Johnston,^{47,48,76-80} Goldstein,⁹¹⁻⁹³ Walsh¹⁰⁵⁻¹⁰⁷ trials; ⁵Johnston,^{47,48,76-80} Goldstein,⁹¹⁻⁹³ Boss⁹⁴⁻⁹⁶ and Walsh¹⁰⁵⁻¹⁰⁷ trials; ⁶Johnston,^{47,48,76-80} Goldstein,⁹¹⁻⁹³ and Walsh¹⁰⁵⁻¹⁰⁷ trials; ⁷Johnston,^{47,48,76-80} Goldstein,⁹¹⁻⁹³ Boss⁹⁴⁻⁹⁶ and Walsh¹⁰⁵⁻¹⁰⁷ trials; ⁸Johnston,^{47,48,76-80} and Goldstein⁹¹⁻⁹³ trials; ⁹Johnston,^{47,48,76-80} and Boss⁹⁴⁻⁹⁶ trials; ¹⁰Goldstein⁹¹⁻⁹³ and Boss⁹⁴⁻⁹⁶ trials; ¹¹Goldstein,⁹¹⁻⁹³ Boss⁹⁴⁻⁹⁶ and Freedman⁹⁷⁻¹⁰⁰ trials; ¹²Prestwood,⁷³⁻⁷⁵ Goldstein⁹¹⁻⁹³ and Boss⁹⁴⁻⁹⁶ trials; ¹³Prestwood⁷³⁻⁷⁵ and Goldstein⁹¹⁻⁹³ trials; ¹⁴Prestwood⁷³⁻⁷⁵ and Boss⁹⁴⁻⁹⁶ trials; ¹⁵Tsai,⁸¹ Fugere^{75,84-90} and Walsh¹⁰⁵⁻¹⁰⁷ trials.

Figure 4: Risk of adverse events (binary), by comparator and trial type



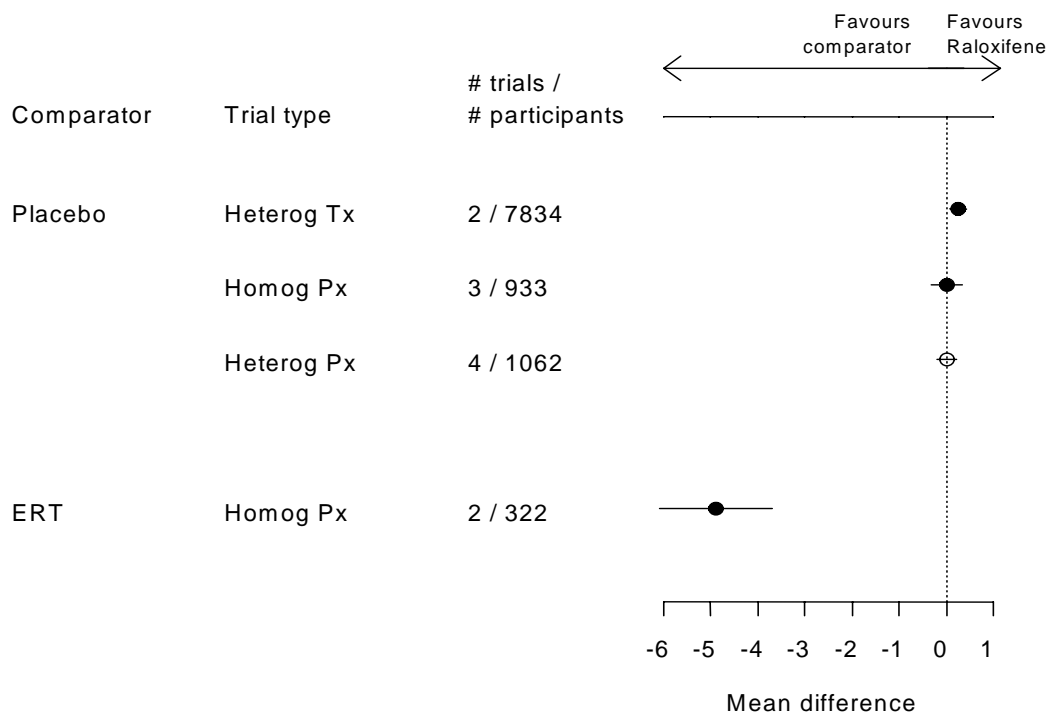
Combined=(low dose=RLX≤60 mg/day) plus (high dose=RLX≥120 mg/day); open circle=fixed effects model; solid circle=random effects model; ERT=estrogen replacement therapy; HRT=hormone replacement therapy; pts=participants; pr=proliferation; thr=thrombosis; heterog Tx=Tx trials with not all participants diagnosed osteoporotic; homog Tx=Tx trials with all participants diagnosed osteoporotic; heterog Px=Px trials with not all participants diagnosed osteopenic or “healthy”; mixed Tx/Px=heterogeneously composed trials, which when added to trials with homogeneous composition of participants defined via WHO BMD criteria, yields designation of “heterog” trials; ¹discontinuations due to adverse events; ²discontinuations with no reason reported.

Table 8: Analyses of continuous adverse event data; organized by comparator and trial type

| Adverse Event Outcome (TVU) | Trial Type* | Dose | Number of Trials/ Number of Participants ^{trials} | Weighted Mean** (95% CI) | Statistical Heterogeneity |
|-----------------------------|-------------|----------|---|------------------------------------|---------------------------|
| Against placebo | | | | | |
| Endometrial thickness | Heterog Tx | Combined | 2/7,834 ^{1, a} | 0.25 (0.08, 0.43) | p=0.27 |
| | Homog Px | Combined | 3/933 ^{2, b} | 0.00 (-0.33, 0.33) | p=0.23 |
| | Heterog Px | Combined | 4/1,062 ^{3, b} | -0.01 (-0.22, 0.20) ^{***} | p=0.40 |

CI=confidence interval; TVU=transvaginal ultrasonography; Tx=treatment; Px=prevention; heterog Tx=Tx trials with not all pts diagnosed osteoporotic; homog Tx=Tx trials with all pts diagnosed osteoporotic; heterog Px=Px trials with not all pts diagnosed osteopenic or “healthy”; homog Px=Px trials with all pts diagnosed osteopenic or “healthy”; mixed Tx/Px=heterogeneously composed trials, which when added to trials with a homogeneous composition of pts defined via WHO BMD criteria, yields a designation of “heterog trials”; combined = all doses = (low dose=RLX≤60 mg/day) plus (high dose=RLX≥120 mg/day); *trial type=treatment versus prevention trial, based on WHO’s BMD defined typology (i.e. osteoporotic versus osteopenic versus healthy); **change from baseline; ***with little statistical heterogeneity, random effects estimate is identical to fixed effects estimate; ¹MORE trial at three years and Meunier⁶⁴⁻⁶⁷ trial; ²Johnston,^{47,48,76-80} Goldstein⁹¹⁻⁹³ and Vardy¹⁰¹ trials; ³Meunier,⁶⁴⁻⁶⁷ Johnston,^{47,48,76-80} Goldstein⁹¹⁻⁹³ and Vardy¹⁰¹ trials; ⁴Goldstein⁹¹⁻⁹³ and Vardy¹⁰¹ trials; ^avariance imputation required for MORE trial at three years; ^bvariance imputation required for Johnston^{47,48,76-80} and Vardy¹⁰¹ trials; ^cvariance imputation required for Vardy¹⁰¹ trial.

Figure 5: Endometrial thickness (continuous) by comparator and trial type*



*All analyses involved combined raloxifene (RLX) doses; combined=(low dose=RLX≤60 mg/day) plus (high dose=RLX ≥120 mg/day); open circle=fixed effects model; solid circle=random effects model; Tx=treatment; Px=prevention; heterog Tx=Tx trials with not all participants diagnosed osteoporotic; homog Tx=Tx trials with all participants diagnosed osteoporotic; heterog Px=Px trials with not all participants diagnosed osteopenic or “healthy”; mixed Tx/Px=heterogeneously composed trials, which when added to trials with homogeneous composition of participants defined via WHO BMD criteria, yields designation of “heterog” trials.

The lowest relative risk of adverse events was associated with trials involving women with higher BMD levels at baseline (younger women with fewer years post-menopause). This effect failed to achieve significance for low doses of raloxifene. Overall, the effect regarding hot flashes was greatest when combined hormone therapy was the comparator [6.77 (3.39, 13.52)]. The relative impact of raloxifene versus estrogen [2.96 (1.31, 6.69)] was closer to the relative effect of raloxifene versus placebo.

When compared with placebo, there was an increased risk of discontinuation due to adverse events associated with raloxifene [1.22 (1.05, 1.42)]. There was a decreased risk of endometrial proliferation associated with raloxifene use when compared with placebo [0.38 (0.19, 0.79)] or when fewer than all the postmenopausal women were diagnosed as either osteopenic or healthy [0.06 (0.04, 0.09)]. A risk reduction for hyperplasia was also observed when raloxifene was compared with estrogen alone in women with intact uteri [0.04 (0.00, 0.83)]. As estrogen alone is contraindicated in women with intact uteri because of the associated risk of endometrial cancer, this comparison provides little assistance in a safety assessment of raloxifene. Increases in endometrial thickness and hyperplasia are measured as potential indicators of risk of endometrial cancer. Thus, estrogen alone is also an inappropriate comparator for these outcomes. The results from the MORE trial suggest that when compared to placebo, raloxifene may be protective.

No difference in the incidence of breast pain and soreness events was observed when placebo and raloxifene were compared. A decreased risk of these events was associated with raloxifene compared with estrogen [0.19 (0.11, 0.32)] or combined hormone therapy [0.14 (0.07, 0.25)]. A decreased risk of episodes of vaginal bleeding was associated with raloxifene versus placebo [0.57 (0.33, 0.97)] or combined hormone therapy [0.11 (0.07, 0.15)]. The larger benefit was observed when combined hormone therapy was the control. Where placebo was the comparator, the initial significant effect disappeared for the low raloxifene dose [0.67 (0.37, 1.19)]. No differences were found with respect to episodes of leukorrhea, vaginitis or retinal vein thrombosis when raloxifene was compared to placebo.

Continuous data (Table 8, Figure 5) revealed that, relative to placebo, raloxifene use was associated with significantly greater increase in endometrial thickness [0.25 (0.08, 0.43)]. This finding occurred in trials where not all participants had very low BMD at baseline.

No evidence of significant statistical heterogeneity was observed for meta-analyses of adverse events.

Adverse events: MORE trial

Table 9 describes the adverse events that occurred in >2% of patients who participated in the MORE trial.

Table 9: Adverse events experienced by at least 2% of participants in MORE trial*

| Adverse Event | Placebo N=2,576 | Raloxifene 60 mg/d N=2,557 | Raloxifene 120 mg/d N=2,572 | Raloxifene (pooled) versus Placebo p |
|---------------------------------------|--------------------|-------------------------------|-----------------------------------|--|
| Influenza syndrome | 293 (11.4%) | 346 (13.5%) | 345 (13.4%) | 0.01 |
| Hot flashes (vasodilation) | 165 (6.4%) | 249 (9.7%) | 269 (11.6%) | <0.001 |
| Leg cramps | 96 (3.7%) | 178 (7.0%) | 178 (6.9%) | <0.001 |
| Peripheral edema | 114 (4.4%) | 134 (5.2%) | 168 (6.5%) | <0.01 |
| Endometrial cavity fluid [§] | 43 (5.7%) | 60 (8.1%) | 66 (8.7%) | 0.02 |
| Diabetes | 14 (0.5%) | 31 (1.2%) | 28 (1.1%) | 0.01 |
| Hypertension | 231 (9.0%) | 177 (6.9%) | 194 (7.5%) | 0.01 |
| Hypercholesterolemia | 121 (4.7%) | 55 (2.2%) | 50 (1.9%) | <0.001 |
| Hematuria | 55 (2.1%) | 35 (1.4%) | 33 (1.3%) | <0.01 |
| Bradycardia | 30 (1.2%) | 13 (0.5%) | 17 (0.7%) | 0.01 |
| Endometrial cancer | 4 (0.2%) | 4 (0.2%) | 2 (< 0.1%) | 0.67 |
| Breast pain | 65 (2.5%) | 61 (2.4%) | 70 (2.7%) | 0.94 |
| Vaginal bleeding | 62 (3.1%) | 67 (3.4%) | 56 (2.8%) | 0.99 |

*40 months of follow-up

[§]among 2,262 women who had transvaginal ultrasonography**d) Results of serious adverse events**

Serious adverse events were reported in the MORE publications and more completely in the medical review by the US FDA¹¹⁵ (Table 10).

Table 10: Adverse effects of raloxifene versus placebo over three years of follow-up

| | Placebo Absolute Risk (N*=2,576) | Raloxifene (pooled) Absolute Risk (N=5,129) | ARR/ARI | NNT/NNH |
|--|---|--|---------|-----------|
| Deaths: total* | 26 (1.0%) | 41 (0.8%) | - | - |
| Deaths: cardiovascular** | 15* (0.6%) | 31 (0.6%) | - | - |
| Serious adverse events (FDA medical review) | 25% | 24% | - | - |
| Venous thromboembolic events | | | | |
| Number of women | 8 (0.31%) | 49 (0.96%) | + .65% | 154 (NNH) |
| Early breast cancer diagnoses *** | 27 (1.05%) | 15 (0.29%) | - .76% | 132 (NNT) |
| Cardiovascular events: nonfatal | 82 (3.2%) | 145 (2.8%) | - | - |

*13 types of non-vertebral fractures were examined individually and none differed significantly (Bonferroni adjustment for 13 comparisons).

**Cardiovascular deaths included fatal myocardial infarction (MI), sudden death, unwitnessed death in absence of other non-coronary causes, death related to coronary artery procedure, fatal stroke.

Non-fatal cardiovascular outcomes included MI, unstable angina, coronary ischemia and cerebrovascular events (stroke or transient ischemic events).

***All but two cases were classified as invasive; invasiveness of remaining two was unknown.

ARR=absolute risk reduction; ARI=absolute risk increase; NNT=numbers needed to treat; NNH=numbers needed to harm

Cardiovascular event rates have been reported.¹¹⁶ The rate of early breast cancer incidence is lower in women on raloxifene than on placebo, 27 (1.05%) versus 15 (0.29%) respectively, an absolute risk decrease of -0.76% (NNT=132).

4 DISCUSSION

In this systematic review, a meta-analysis was planned to pool evidence from RCTs regarding raloxifene's effect on vertebral fracture (primary outcome) and BMD (secondary outcome) in postmenopausal women. We identified 17 RCTs that included vertebral fractures and BMD data. All trials assessing non-vertebral fractures and morphometric and clinical vertebral fractures had placebo as the control intervention. Adverse event data involved placebo, estrogen and combined hormone therapy as comparators.

The planned meta-analysis of vertebral fractures could not be justified for the two trials reporting vertebral fracture data. The central problems were the observed clinical heterogeneity (i.e., definition of vertebral fracture and population characteristics^{17,20,111-114}) and statistical heterogeneity. As a result, there was no justification for using a random effects model to combine MORE and Lufkin data. In addition, vertebral fracture was of clinical importance when it was associated with back pain and disability, but the Lufkin trial did not include any separate reporting of women with symptomatic vertebral fracture.

Given the size, quality and duration of the MORE trial, its results best reflected raloxifene's effects on vertebral fracture, as compared to placebo, in women with very low BMD or vertebral deformity at baseline. The MORE trial tested two doses of raloxifene: 60 mg/day and 120 mg/day. The 120 mg/day dose provided a greater benefit, but the difference between 120 mg/day and 60 mg/day was not statistically significant. In addition, 120 mg/day exceeds the recommended daily dose in Canada (60 mg).

In the MORE trial, the incidence of vertebral fracture differed in the two substudies and was related to the presence of vertebral fracture at baseline (secondary prevention). Vertebral fracture at baseline was a stronger determinant of subsequent vertebral fracture during the study period, than treatment allocation to raloxifene or placebo.

Pooled analysis was used for BMD findings. Raloxifene had a significant positive impact on BMD compared with placebo at various sites. The most pronounced impacts were observed for the lumbar spine in studies where fewer than all participants had very low BMD at baseline (heterogeneous low BMD) or fewer than all patients had higher BMD at baseline (i.e., heterogeneous high BMD).

In the MORE study, the positive BMD impact did not correlate with clinical benefit as measured by non-vertebral fractures or vertebral fracture. Women with low BMD at baseline but no vertebral fracture had lower rates of vertebral fracture during the trial than women with vertebral fracture. The latter group included women enrolled in this group regardless of initial BMD. In addition, data from the MORE trial (reported by the FDA) indicated that women on placebo with higher (i.e., better) lumbar spine BMD also had a higher incidence of vertebral fracture.

Investigators have tried to resolve the "treatment paradox" seen in the MORE study and elsewhere,^{117,118} whereby decreases in vertebral fracture rates occur without commensurate increases in bone density.¹¹⁹⁻¹²¹ Other variables may be responsible for the risk reduction in vertebral fracture that is not reflected in BMD. For example, slowing bone resorption may

increase the biomechanical strength of bone and reduce bone turnover, which in turn alter vertebral fracture risk. One explanation for this finding is that vertebral fractures can artifactually increase BMD, which leads to an increase in lumbar spine BMD measures. It is unclear, however, why this would occur in the placebo group but not in women treated with raloxifene.

The MORE trial was of sufficient size and duration to assess the clinical outcomes of raloxifene therapy (Table 11).

Table 11: Non-vertebral fracture incidence: raloxifene versus placebo over three years of follow-up

| | Placebo Absolute Risk (N*=2,576) | Raloxifene (pooled) Absolute Risk (N=5,129) | ARR/ARI | NNT/NNH |
|------------------------------|---|--|----------------|----------------|
| Hip fractures* | 18 (0.7%) | 40 (0.8%) | - | - |
| All non-vertebral fractures* | 240 (9.3%) | 437 (8.5%) | - | - |

*13 types of non-vertebral fractures were also examined individually and none differed significantly (Bonferroni adjustment for 13 comparisons).

The Lufkin trial also reported on clinical fractures. Total fracture rate was not reported. Three women on placebo (6.6%) and three on raloxifene (3.4%) experienced non-vertebral fractures. No placebo users and one raloxifene user (1.1%) experienced a hip fracture. The differences were not statistically significant. At a rate of 9% for total non-vertebral fractures in the placebo group, the trial had 80% power to detect an absolute difference of 1.9% in fracture rate with a 95% degree of confidence (post hoc power analysis, Epi Info).

In the pooled analysis, raloxifene decreased the risk of several mild to moderate adverse events compared with estrogen or combined hormone therapy (i.e., breast pain and soreness, endometrial proliferation, vaginal bleeding). Two of the active comparators are unapproved for osteoporosis prevention: tamoxifen and estrogen alone in women without hysterectomy (the latter is contraindicated because of increased endometrial cancer risks).^{122,123}

As compared with placebo, raloxifene increases the rate of hot flashes significantly at a higher than approved dose level (120 mg/day). The comparison of raloxifene to hormone therapy is questionable, however, given that raloxifene has no claimed effects in the symptomatic relief of hot flashes, while this is an approved indication for combined estrogen-progestin therapies and for estrogen alone in women with a previous hysterectomy. The comparison to placebo, on the other hand, indicates whether raloxifene would increase the probability of hot flashes as compared to no treatment.

Discontinuations due to an adverse event were more likely for raloxifene users in placebo-controlled, trials of women with very low BMD (generally older women). Breast pain and soreness were less likely to be a problem for raloxifene users than for recipients of a comparator in trials involving women with higher BMD levels.

The impact of raloxifene on lipid metabolism was beyond the focus of this review, but the MORE trial demonstrated that raloxifene significantly decreased levels of LDL cholesterol and total cholesterol relative to placebo (Table 9).

A post hoc assessment of cardiovascular outcomes in the MORE trial found no evidence of an increased rate of cardiovascular events with raloxifene versus placebo. Raloxifene's estrogen antagonist effect on breast tissue could not be pooled because of inadequate breast cancer data from most trials. The MORE and Johnston-Delmas trials included data, but the data could not be combined. In the MORE trial, there was a reduction in early invasive breast cancer, although it was impossible to establish whether raloxifene had a sustained effect on breast cancer progression or mortality.

Raloxifene's effect on uterine tissue cannot be illuminated by the results of the pooled analyses. In placebo-controlled RCTs, raloxifene use is associated with increased endometrial thickness. The observations concerning endometrial proliferation and vaginal bleeding, however, are associated with studies of younger women. Unopposed estrogen therapy is a known cause of endometrial cancer and thus, it is an inappropriate comparator to assess the potential risks associated with cancer from another drug. None of the morphological or anatomical changes in the uterus associated with raloxifene are linked to cancer incidence, morbidity or mortality.

Serious adverse event rates could only be determined from the large MORE trial. Raloxifene significantly increased and decreased the risk of venous thromboembolic events and breast cancer respectively.

Overall, 24.2% of participants in the MORE trial experienced a serious adverse event. The incidence did not differ between raloxifene and placebo. Limited information was available on the type of events [cardiovascular, venous thromboembolic events (VTE), breast cancer incidence, mortality]. The reported serious adverse events accounted for approximately 5% of the total serious adverse events in each category. Underreporting of serious adverse events was seen as typical of the RCT literature.^{124,125}

The MORE trial reports limited data regarding VTEs beyond the total events for each therapeutic group. Details are crucial, because these events typically range from less severe venous obstruction to fatal pulmonary embolism.

The manufacturer provided additional details regarding VTEs to the US FDA. They reported this as the main serious adverse event observed in the approximately 10,000 women participating in all clinical trials sponsored by Lilly (Table 12). No dose-specific reports were available, but the manufacturer stated that the frequency of VTE did not increase with (increasing) doses above 60 mg/day.¹¹⁵

Table 12: Venous thromboembolic events

| Adverse Events: Estimated annual incidence rate (per 1,000 patients) | Treatment Group | | |
|---|----------------------------|-------------------------------|---|
| | Placebo N=3,195 | Raloxifene N=6,681 | RR (95% Confidence Interval) |
| All venous thromboembolic events (VTE) | 1.12 | 3.82 | 3.4 (1.5 to 8.0) |
| All VTE except retinal vein thrombosis | 0.75 | 3.64 | 4.9 (1.8 to 14) |
| Pulmonary emboli | 0.56 | 1.24 | 2.2 (0.6 to 7.8) |

The estimated annual attributable risk of a venous thromboembolic VTE with raloxifene use is 2.7/1,000 women. This is a measure of the incidence in raloxifene users minus the incidence in placebo controls. VTE increases in frequency with age and most women participating in clinical trials are <65. The largest increases in VTE are observed in trials involving older women.¹¹⁵

The most common complication of VTEs is post-thrombotic syndrome (persistent leg pain and swelling with or without ulceration) after the occurrence of a deep vein thrombosis (DVT). Post-thrombotic syndrome is estimated to occur within five years in 60% to 70% of patients who develop proximal deep vein thrombosis and within two years in 16% of the patients who develop a distal deep vein thrombosis.¹²⁶ Siragusa *et al.* found that the incidence of this disorder was 24% compared with 4% in controls, after two to four years, in patients who had subclinical DVT demonstrated by contrast venography after hip or knee surgery, despite at least three months of anticoagulation.¹²⁷

In the MORE trial, raloxifene decreased the incidence of early invasive breast cancer diagnoses. Whether these represented breast cancer prevention or a delay in diagnosis is unknown, as the number of diagnoses was small and follow-up insufficient. The effect on breast cancer mortality was unknown. There were also more cases of treatment-emergent diabetes in raloxifene users versus placebo in the MORE trial. The longer-term clinical implications of these differences remained unclear.

To summarize observed clinical benefit and harm (beyond tolerability), raloxifene was not found to have a statistically significant effect on mortality or overall serious morbidity, nor was any effect observed on non-vertebral fracture rates, notably hip fractures. The MORE trial was adequately powered to determine the reduction of total non-vertebral fractures and no reduction was found except in a post hoc analysis of individuals with severe vertebral fractures.¹²⁸ This raised questions about the product's efficacy in terms of the outcomes that were likely to be most important to patients and clinicians. The MORE trial was inadequately powered to examine hip fractures alone. Larger, long-term studies are needed to determine raloxifene's use for the prevention of hip fractures. The most important observed clinical outcomes were a reduction in symptomatic vertebral fractures, a reduction in early invasive breast cancer in postmenopausal women and an increase in thrombo-embolic events.

There are several research implications. The published results of the Women's Health Initiative trial of estrogen-progestin treatment for disease prevention reinforce the need for evidence that benefits outweigh harms before a strategy for osteoporosis prevention can be recommended to healthy populations.¹²⁹

Other studies are needed to explain why the MORE trial found that raloxifene has a beneficial effect on vertebral fractures, but no effect on non-vertebral fractures. This will help explain the interrelationships among biochemical markers of bone turnover, biomechanical bone strength, BMD, vertebral fractures and clinical fractures.¹⁵⁻¹⁷

The STAR trial, now underway, compares raloxifene and tamoxifen for the prevention of breast cancer. Unfortunately, the study lacks a placebo arm, making it impossible to determine the effect of these drugs on serious adverse events and mortality in an initially healthy population.

5 CONCLUSIONS

The best estimate of overall benefit and harm from raloxifene is derived from the MORE trial's three-year, placebo-controlled results. Raloxifene has no effect on the incidence of non-vertebral fractures or on hip fracture. The absolute risk reduction for symptomatic vertebral fracture is 0.8% (1.4% in placebo and 0.6% in raloxifene groups, number needed to treat=125 over three years). This benefit needs to be balanced by the increase in the number of serious adverse events due to venous thrombo-embolic disease of 0.65% (0.31% in placebo and 0.96% in raloxifene, number needed to harm=154 over three years).

The largest observed incidence of vertebral fracture occurs among older women with vertebral deformity at baseline. This is also the group in which raloxifene has the greatest impact. The group includes women with a range of BMD levels. The impact of raloxifene on vertebral fracture in younger postmenopausal women or women with higher BMD levels has not been established.

The impact of raloxifene on morbidity and mortality due to cardiovascular disease or cancer (breast or uterine) has not been established, although the evidence suggests a small but significant decrease in the incidence of early breast cancer. A subgroup analysis of cardiovascular adverse events reported in the MORE trial indicates no significant increase in morbidity and mortality in raloxifene users as compared to placebo, after three years.

6 REFERENCES

1. Genant HK, Cooper C, Poor G, Reid I, Ehrlich G, Kanis J, et al. Interim report and recommendations of the World Health Organization Task-Force for Osteoporosis. *Osteoporos Int* 1999;10(4):259-64.
2. Melton LJ, Chrischilles EA, Cooper C, Lane AW, Riggs BL. Perspective. How many women have osteoporosis? *J Bone Miner Res* 1992;7(9):1005-10.
3. Jaglal S. Osteoporotic fractures: incidence and impact. In: Badley EM, Williams JI, editors. *Patterns of health care in Ontario: arthritis & related conditions: an ICES practice atlas*. Toronto: Institute for Clinical Evaluative Sciences in Ontario; 1998. p.143-56.
4. Goeree R, O'Brien B, Pettitt D, Cuddy L, Ferraz M, Adachi JD. An assessment of the burden of illness due to osteoporosis in Canada. *J Soc Obstet Gynaecol Can* 1996;18(July Suppl):15-24.
5. Wiktorowicz ME, Goeree R, Papaioannou A, Adachi JD, Papadimitropoulos E. Economic implications of hip fracture: health service use, institutional care and cost in Canada. *Osteoporos Int* 2001;12(4):271-8.
6. Papadimitropoulos EA, Coyte PC, Josse RG, Greenwood CE. Current and projected rates of hip fracture in Canada. *CMAJ* 1997;157(10):1357-63.
7. Cooper C, Atkinson EJ, O'Fallon WM, Melton LJ, III. Incidence of clinically diagnosed vertebral fractures: a population-based study in Rochester, Minnesota, 1985-1989. *J Bone Miner Res* 1992;7(2):221-7.
8. Cummings SR, Melton LJ. Epidemiology and outcomes of osteoporotic fractures. *Lancet* 2002;359(9319):1761-7.
9. Nevitt MC, Ettinger B, Black DM, Stone K, Jamal SA, Ensrud K, et al. The association of radiographically detected vertebral fractures with back pain and function: a prospective study. *Ann Intern Med* 1998;128(10):793-800.
10. Nevitt MC, Thompson DE, Black DM, Rubin SR, Ensrud K, Yates AJ, et al. Effect of alendronate on limited-activity days and bed-disability days caused by back pain in postmenopausal women with existing vertebral fractures. Fracture Intervention Trial Research Group. *Arch Intern Med* 2000;160(1):77-85.
11. Adachi JD, Loannidis G, Berger C, Joseph L, Papaioannou A, Pickard L, et al. The influence of osteoporotic fractures on health-related quality of life in community-dwelling men and women across Canada. *Osteoporos Int* 2001;12(11):903-8.
12. Lindsay R, Silverman SL, Cooper C, Hanley DA, Barton I, Broy SB, et al. Risk of new vertebral fracture in the year following a fracture. *JAMA* 2001;285(3):320-3.
13. Klotzbuecher CM, Ross PD, Landsman PB, Abbott TA, III, Berger M. Patients with prior fractures have an increased risk of future fractures: a summary of the literature and statistical synthesis. *J Bone Miner Res* 2000;15(4):721-39.
14. Blank RD, Bockman RS. A review of clinical trials of therapies for osteoporosis using fracture as an end point. *J Clin Densitom* 1999;2(4):435-52.
15. NIH Consensus Development Panel on Osteoporosis Prevention, Diagnosis, and Therapy. Osteoporosis prevention, diagnosis, and therapy. *JAMA* 2001;285(6):785-95.

16. Brown JP, Fortier M, Khan A, Rowe T. An evidence-based review of the management of osteoporosis. *J Soc Obstet Gynaecol Can* 2001;April:1-7.
17. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Report of a WHO Study Group. *World Health Organ Tech Rep Ser* 1994;843:1-129.
18. Kanis JA, Melton LJ, III, Christiansen C, Johnston CC, Khaltsev N. The diagnosis of osteoporosis. *J Bone Miner Res* 1994;9(8):1137-41.
19. Kanis JA, Johnell O, Oden A, Dawson A, De Laet C, Jonsson B. Ten year probabilities of osteoporotic fractures according to BMD and diagnostic thresholds. *Osteoporos Int* 2001;12(12):989-95.
20. Cummings SR, Nevitt MC, Browner WS, Stone K, Fox KM, Ensrud KE, et al. Risk factors for hip fracture in white women. Study of Osteoporotic Fractures Research Group. *N Engl J Med* 1995;332(12):767-73.
21. Brown JP, Josse RG. 2002 clinical practice guidelines for the diagnosis and management of osteoporosis in Canada. *CMAJ* 2002;167(10 Suppl):S1-34.
22. Mark S, Gupta J. Prevention and treatment of osteoporosis: introduction. *Ann N Y Acad Sci* 2001;949:186-7.
23. Sherman S. Preventing and treating osteoporosis: strategies at the millennium. *Ann N Y Acad Sci* 2001;949:188-97.
24. Green CJ, Bassett K, Foerster V, Kazanjian A. *Bone mineral testing: does the evidence support its selective use in well women?* Vancouver: BC Office of Health Technology Assessment, University of British Columbia; 1997. Available: www.chspr.ubc.ca/bcohta/pdf/bmd.pdf (accessed 2003 Feb 5).
25. Canadian Pharmacists Association. Welbanks L, editor. *CPS 2000: compendium of pharmaceuticals and specialties*. 35th ed. Ottawa: The Association; 2000.
26. Khovidhunkit W, Shoback DM. Clinical effects of raloxifene hydrochloride in women. *Ann Intern Med* 1999;130(5):431-9.
27. *EVISTA [product monograph]*. Eli Lilly Canada; 2000.
28. Ontario Program for Optimal Therapeutics. *Ontario guidelines for the prevention and treatment of osteoporosis*. 1st. Toronto: The Program; 2000. Available: <http://www.opot.org/guidelines/osteoporosis.pdf> (accessed 2003 Feb 5).
29. Jadad AR. *Randomised controlled trials: a user's guide*. 1st. London: BMJ Publishing Group; 1998.
30. *A dictionary of epidemiology*. Last JM, editor. *A dictionary of epidemiology*. 3rd ed. New York: Oxford University Press; 1995.
31. Donner A, Eliasziw M. Sample size requirements for reliability studies. *Stat Med* 1987;6(4):441-8.
32. Berlin JA. Does blinding of readers affect the results of meta-analyses? University of Pennsylvania Meta-analysis Blinding Study Group. *Lancet* 1997;350(9072):185-6.
33. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;17(1):1-12.
34. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273(5):408-12.

35. Greenland S, Robins JM. Estimation of a common effect parameter from sparse follow-up data. *Biometrics* 1985;41(1):55-68.
36. Hardy RJ, Thompson SG. Detecting and describing heterogeneity in meta-analysis. *Stat Med* 1998;17(8):841-56.
37. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3):177-88.
38. Moher D, Jadad AR, Klassen TP. Guides for reading and interpreting systematic reviews: III. How did the authors synthesize the data and make their conclusions? *Arch Pediatr Adolesc Med* 1998;152(9):915-20.
39. Vickers AJ. The use of percentage change from baseline as an outcome in a controlled trial is statistically inefficient: a simulation study. *BMC Med Res Methodol* 2001;1(1):6.
40. Follmann D, Elliott P, Suh I, Cutler J. Variance imputation for overviews of clinical trials with continuous response. *J Clin Epidemiol* 1992;45(7):769-73.
41. Mackerras D, Lumley T. First- and second-year effects in trials of calcium supplementation on the loss of bone density in postmenopausal women. *Bone* 1997;21(6):527-33.
42. Begg CB, Berlin JA. Publication bias: a problem in interpreting medical data. *J R Stat Soc A* 1988;151:419-63.
43. Egger M, Davey SG, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315(7109):629-34.
44. Duval S, Tweedie R. A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *J Am Stat Assoc* 2000;95:89-98.
45. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999;354(9193):1896-900.
46. Erkkola RS, Halonen K, Heikkinen J, Kivinen S, Lammintausta R. Comparison of FC1271a, a novel SERM, and raloxifene in postmenopausal women: a 12-week clinical phase II study [abstract]. 11th International Congress of Endocrinology; 2000 Oct 29; Sydney, Australia.
47. Johnston CC, Bjarnason NH, Cohen FJ, Shah A, Lindsay R, Mitlak BH, 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.
48. Delmas PD, Bjarnason NH, Mitlak BH, Ravoux AC, Shah AS, Huster WJ, 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.
49. Ettinger B, Black DM, Mitlak BH, Knickerbocker RK, Nickelsen T, Genant HK, et al. Reduction of vertebral fracture risk in postmenopausal women with osteoporosis treated with raloxifene: results from a 3-year randomized clinical trial. Multiple Outcomes of Raloxifene Evaluation (MORE) Investigators. *JAMA* 1999;282(7):637-45.
50. Delmas PD, Ensrud KE, Harper K, Wu W, Adachi J, Eastell R, et al. Effects of raloxifene on bone mineral density and biochemical markers of bone turnover in postmenopausal women with osteoporosis: 4-Year results from the MORE trial [abstract]. *J Bone Min Res* 2000;15(Suppl 1):S556.

51. Eastell R, Adachi J, Harper K, Sarkar S, Delmas PD, Ensrud K, et al. The effects of raloxifene on incident vertebral fractures in postmenopausal women with osteoporosis: 4-year results from the MORE trial [abstract]. *J Bone Min Res* 2000;15(Suppl 1):S229.
52. Ensrud K, Black D, Recker R, Harris S, Delmas P, Pols H, et al. The effect of 2 and 3 years of raloxifene on vertebral and non-vertebral fractures in postmenopausal women with osteoporosis [abstract]. *Bone* 1998;23(5 Suppl 1):S174.
53. Ettinger B, Black D, Cummings S, Genant H, Gluër C, Lips P, et al. The effect of raloxifene in the treatment of osteoporosis: 24-month interim results [abstract]. The North American Menopause Society 9th Annual Meeting; 1998 Sep 16; Toronto.
54. Ettinger B, Black D, Cummings S, Genant H, Glüer C, Lips P. Raloxifene reduces the risk of incident vertebral fractures: 24-month interim analyses [abstract]. *Osteoporos Int* 1998;8(Suppl 3):11.
55. Cauley J, Krueger K, Eckert S, Muchmore D, Taylor Y, Scott T. Breast cancer risk in postmenopausal women with osteoporosis is reduced with raloxifene treatment: extension of the MORE trial data to a median follow-up of 52 months [abstract]. American Society of Clinical Oncology 36th Annual Meeting; 2000 May 20; New Orleans.
56. Cummings SR, Eckert S, Krueger KA, Grady D, Powles TJ, Cauley JA, 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.
57. Cummings SR, Norton L, Eckert S, Grady D, Cauley J, Knickerbocker R, et al. Raloxifene reduces the risk of breast cancer and may decrease the risk of endometrial cancer in post-menopausal women. Two-year findings from the multiple outcomes of raloxifene evaluation (MORE) trial [abstract]. *Proc ASCO* 1998;17:2a.
58. Maricic M, Adachi J, Meunier P, Lufkin E, Delmas P, Draper M. Raloxifene 60 mg/day has effects within 12 months in postmenopausal osteoporosis treatment and prevention studies [abstract]. *Arthritis Rheum* 2000;43(Suppl 9):S197.
59. Lufkin EG, Whitaker MD, Nickelsen T, Argueta R, Caplan RH, Knickerbocker RK, et al. Treatment of established postmenopausal osteoporosis with raloxifene: a randomized trial. *J Bone Miner Res* 1998;13(11):1747-54.
60. Lufkin EG, Whitaker MD, Nickelsen T, Arqueta R, Caplan RH, Knickerbocker RK, et al. Raloxifene effects on bone mineral density and safety assessments in postmenopausal women with osteoporosis [abstract]. *Osteoporos Int* 1998;8(Suppl 3):P284.
61. Lufkin E, Whitaker M, Argueta R, Caplan R, Nickelsen T, Riggs B. Raloxifene promising treatment in osteoporosis [abstract]. *Osteoporos Int* 1997;7(Suppl 2):49.
62. Lufkin EG, Whitaker MD, Argueta R, Caplan RH, Nickelsen T, Riggs BL. Raloxifene treatment of postmenopausal osteoporosis [abstract]. *J Bone Miner Res* 1997;12:S150.
63. Nickelsen T, Lufkin EG, Riggs BL, Cox DA, Crook TH. 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.
64. Meunier PJ, Vignot E, Garnero P, Confavreux E, Paris E, Liu-Leage S, et al. Treatment of postmenopausal women with osteoporosis or low bone density with raloxifene. Raloxifene Study Group. *Osteoporos Int* 1999;10(4):330-6.

65. Meunier PJ, Vignot E, Garnero P, Confavreux E, Paris E, Sarkar S, et al. Treatment of postmenopausal osteoporosis with raloxifene [abstract]. *Osteoporos Int* 1998;8(Suppl 3):P304.
66. Meunier PJ, Vignot E, Garnero P, Confavreux E, Paris E, Sarkar S, et al. Postmenopausal women with osteoporosis: effects of raloxifene on bone mineral density in the hip and lumbar spine [abstract]. *Bone* 1998;23(5 Suppl 1):S295.
67. Sarkar S, Meunier PJ, Vignot E, Garnero P, Confavreux E, Liu-Leage S, et al. Changes in bone mineral density and serum cholesterol in osteoporotic postmenopausal women treated with raloxifene [abstract]. The North American Menopause Society 9th Annual Meeting; 1998 Sep 16; Toronto.
68. Johnell O, Lu Y, Seeman E, Reginster J, Scheele WH. Effects of raloxifene and alendronate on bone mineral density and bone turnover markers in postmenopausal women with osteoporosis [abstract]. *Osteoporos Int* 2000;11(Suppl 2):S184.
69. Johnell O, Scheele W, Lu Y, Lakshaman M. Effects of raloxifene (RLX), alendronate (ALN), and RLX+ALN on bone mineral density and biochemical markers of bone turnover in postmenopausal women with osteoporosis [abstract]. *J Bone Miner Res* 1999;14(Suppl 1):S157.
70. Stock J, Johnell O, Scheele W, Lu Y, Lakshaman M. Effects of raloxifene (RLX), alendronate (ALN), and RLX+ALN on bone mineral density and biochemical markers of bone turnover in postmenopausal women with osteoporosis [abstract]. *Arthritis Rheum* 1999;42(Suppl 9):S176.
71. Morii H, Orimo H, Taketani Y, Ishii J, Inoue T, Fujita T, et al. Baseline characteristics do not affect the response to raloxifene hydrochloride in Japanese women with osteoporosis [abstract]. *Maturitas* 1997;27(Suppl 1):70-1.
72. Morii H, Orimo H, Taketani Y, Ishii J, Inoue T, Fujita T. Effect of raloxifene hydrochloride on postmenopausal Japanese women with osteoporosis [abstract]. *J Bone Miner Res* 1996;11(Suppl):S446.
73. Prestwood KM, Guinness M, Muchmore DB, Lu Y, Wong M, Raisz LG. A comparison of the effects of raloxifene and estrogen on bone in postmenopausal women. *J Clin Endocrinol Metab* 2000;85(6):2197-202.
74. Guinness M, Prestwood K, Lu Y, Muchmore DB, Raisz L. Histomorphometric, bone marker, and bone mineral density response to raloxifene HCl and Premarin in postmenopausal women [abstract]. The Endocrine Society 79th Annual Meeting; 1997 Jun 11; Minneapolis.
75. Prestwood K, Muchmore D, Scheele W, Shah A, Lu Y, Fugère P, et al. Raloxifene HCl does not stimulate the endometrium in postmenopausal women [abstract]. The North American Menopause Society 8th Annual Meeting; 1997 Sep 4; Boston.
76. Shah A, Iversen P, Mitlak B. Clinically favorable effects of raloxifene on bone mineral density and serum lipids: population assessment via bivariate analysis. *Prim Care Update Ob/Gyns* 1998;5(4):165-6.
77. Bjarnason NH, Delmas PD, Mitlak BH, Shah AS, Huster WJ, Draper MW, et al. Raloxifene maintains favourable effects on bone mineral density, bone turnover and serum lipids without endometrial stimulation in postmenopausal women. 3-year study results [abstract]. *Osteoporos Int* 1998;8(Suppl 3):11.
78. Delmas PD. Raloxifene reduces the risk of vertebral fracture and improves lipid profiles of postmenopausal women after 3 years [abstract]. *Calcif Tissue Int* 1999;64(Suppl 1):S120.
79. Bjarnason N, Delmas PD, Mitlak B, Shah A, Huster W, Draper MW, et al. Raloxifene HCl: long term efficacy and safety profile [abstract]. *Acta Obstet Gynecol Scand Suppl* 1997;167(Pt 4):36.

80. Delmas PD, Bjarnason NH, Mitlak B, Shah A, Huster W, Draper M, et al. Raloxifene HCl prevents bone loss and lowers serum cholesterol without endometrial stimulation in early postmenopausal women [abstract]. The Endocrine Society 79th Annual Meeting; 1997 Jun 11; Minneapolis.
81. Tsai KS, Yen ML, Pan HA, Wu MH, Cheng WC, Hsu SH, et al. Raloxifene versus continuous combined estrogen/progestin therapy: densitometric and biochemical effects in healthy postmenopausal Taiwanese women. *Osteoporos Int* 2001;12(12):1020-5.
82. Pavo I, Masanauskaite D, Rojinskaya L, Smetnik V, Prilepskaya V, Melnichenko G, et al. Effects of raloxifene (RLX) vs. placebo (PL) on bone mineral density (BMD) in postmenopausal women in the absence of calcium supplementation [abstract]. *J Bone Min Res* 1999;14(Suppl 1):S410.
83. Stock J, Pavo I, Masanauskaite D, Rojinskaya L, Smetnik V, Prilepskaya V, et al. Effects of raloxifene (RLX) vs. placebo (PL) on bone mineral density (BMD) in postmenopausal women in the absence of calcium supplementation [abstract]. *Arthritis Rheum* 1999;42(9 Suppl):S177.
84. Fugère P, Scheele WH, Shah A, Strack TR, Glant MD, Jolly E. Uterine effects of raloxifene in comparison with continuous-combined hormone replacement therapy in postmenopausal women. *Am J Obstet Gynecol* 2000;182(3):568-74.
85. Fugère P, Scheele WH, Srikanth KR, Anglin G, Strack TR, Nauden AB, et al. Raloxifene does not stimulate the uterus in postmenopausal women as compared to continuous combined hormone replacement therapy following 24 months of treatment [abstract]. *Fertil Steril* 1999;72(3 Suppl 1):S182-3.
86. Shah AS, Scheele WH, Glant MD, Fugère P. Raloxifene HCl is not stimulatory in the endometrium as assessed by the blaustein criteria and an estrogenicity scoring system [abstract]. *Prim Care Update Ob/Gyns* 1998;5(4):167.
87. Jolly EE, Scheele WH, Shah AS, Davies GC, Lafortune M, Fugère P. A high dose of raloxifene HCl is not stimulatory in the uterus of healthy postmenopausal women [abstract]. IV European Congress on Menopause; 1997 Oct 8; Paris.
88. Glant M, Scheele WH, Shah A, Davies GC, Fugère P. Raloxifene HCl is not stimulatory in the endometrium as assessed by an estrogenicity scoring system [abstract]. IV European Congress on Menopause; 1997 Oct 8; Paris.
89. Scheele WH, Symanowski SM, Neale S, Shah A, Lafortune M, Fugère P. Raloxifene does not cause stimulatory effects on the uterus in healthy postmenopausal women [abstract]. The Endocrine Society 79th Annual Meeting; 1997 Jun 11; Minneapolis.
90. Scheele WH, Symanowski SM, Neale S, Shah A, Lafortune M, Fugère P. A high dose of raloxifene HCl is not stimulatory in the uterus of healthy postmenopausal women [abstract]. World Congress of Gynecology and Obstetrics annual meeting; 1997 Aug 3; Copenhagen.
91. Goldstein SR, Scheele WH, Rajagopalan SK, Wilkie JL, Walsh BW, Parsons AK. A 12-month comparative study of raloxifene, estrogen, and placebo on the postmenopausal endometrium. *Obstet Gynecol* 2000;95(1):95-103.
92. Thiebaud D, Goldstein S, Srikanth R, Parsons A, Londono J, Kenemans P, et al. Effects of raloxifene, estrogen and placebo on the postmenopausal endometrium: a 12-month comparative study [abstract]. *Climacteric* 1999;2(Suppl 1):330.
93. Goldstein S, Srikanth R, Parsons A, Londono J, Kenemans P, Wilkie J, et al. Effects of raloxifene on the endometrium in healthy postmenopausal women [abstract]. The North American Menopause Society 9th Annual Meeting; 1998 Sep 16; Toronto.

94. Boss SM, Huster WJ, Neild JA, Glant MD, Eisenhut CC, Draper MW. Effects of raloxifene hydrochloride on the endometrium of postmenopausal women. *Am J Obstet Gynecol* 1997;177(6):1458-64.
95. Draper MW, Flowers DE, Huster WJ, Neild JA, Harper KD, Arnaud C. A controlled trial of raloxifene (LY139481) HCl: impact on bone turnover and serum lipid profile in healthy postmenopausal women. *J Bone Miner Res* 1996;11(6):835-42.
96. Huster WJ, Boss SM, Glant MD, Eisenhut CC, Neild JA, Draper MW. Raloxifene does not cause endometrial proliferation in postmenopausal women [abstract]. *J Soc Gynecol Investig* 1996;3(2 Suppl 1):142A.
97. Freedman M, San Martin J, O'Gorman J, Eckert S, Lippman ME, Lo SC, 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.
98. San Martin J, O'Gorman J, Eckert S, Freedman M, Lo SCB, Walls EL, et al. Digitized mammography in postmenopausal women receiving raloxifene or estrogen in a two-year, placebo-controlled, randomized controlled trial [abstract]. American Society of Clinical Oncology 36th Annual Meeting; 2000 May 20; New Orleans.
99. Cohen F, Watts S, Beymer K. Vaginal complaints in healthy postmenopausal women: data from a 3-year trial comparing raloxifene with conjugated estrogens or placebo [abstract]. The North American Menopause Society 9th annual meeting; 1998 Sep 16; Toronto.
100. Lakshaman M, Cohen F, Watts S, Beymer K. Vaginal complaints in healthy postmenopausal women: data from a 3-year trial comparing raloxifene with conjugated estrogens or placebo [abstract]. *Climacteric* 1999;2(Suppl 1):331.
101. Vardy MD, Cosman F, Heller D, Nieves J, Lindsay R. Short-term endometrial effects of raloxifene, tamoxifen, and Premarin [abstract]. *Fertil Steril* 1998;70(3 Suppl 1):S186-7.
102. Parsons A, Walsh B, Walls E, Siddhanti S, Basson R, Heath H, et al. Effect of concomitant raloxifene and vaginal Premarin on endometrial thickness and histology in a 6-month randomized controlled trial [abstract]. *Menopause* 1999;6(4):341.
103. Parsons A, Nachtigall L, Merritt D, Basson R, Heath H, Plouffe L. Vaginal Premarin vs. Replens in women with pre-existing vaginal atrophy receiving oral placebo or raloxifene: effects on objective endpoints [abstract]. *Menopause* 1999;6(4):341.
104. Parsons A, Nachtigall L, Sulak P, Basson R, Heath H, Plouffe L. Vaginal Premarin vs. Replens in women with pre-existing vaginal atrophy receiving oral placebo or raloxifene: effects on subjective endpoints [abstract]. *Menopause* 1999;6(4):340.
105. Walsh BW, Kuller LH, Wild RA, Paul S, Farmer M, Lawrence JB, et al. Effects of raloxifene on serum lipids and coagulation factors in healthy postmenopausal women. *JAMA* 1998;279(18):1445-51.
106. Walsh BW, Kuller LH, Wild RA, Paul S, Shah A, Anderson P. The effect of raloxifene on markers of cardiovascular risk in healthy, postmenopausal women [abstract]. *Atherosclerosis* 1997;134(1-2):192.
107. Walsh BW, Paul S, Wild RA, Dean RA, Tracy RP, Cox DA, et al. The effects of hormone replacement therapy and raloxifene on C-reactive protein and homocysteine in healthy postmenopausal women: a randomized, controlled trial. *J Clin Endocrinol Metab* 2000;85(1):214-8.

108. Mijatovic V, van der Mooren MJ, Kenemans P, de Valk-de Roo GW, Netelenbos C. Raloxifene lowers serum lipoprotein(A) in healthy postmenopausal women: a randomized, double-blind, placebo-controlled comparison with conjugated equine estrogens. *Menopause* 1999;6(2):134-7.
109. Mijatovic V, Netelenbos C, van der Mooren MJ, de Valk-de Roo GW, Jakobs C, Kenemans P. Randomized, double-blind, placebo-controlled study of the effects of raloxifene and conjugated equine estrogen on plasma homocysteine levels in healthy postmenopausal women. *Fertil Steril* 1998;70(6):1085-9.
110. de Valk-de Roo GW, Stehouwer CD, Meijer P, Mijatovic V, Kluft C, Kenemans P, et al. Both raloxifene and estrogen reduce major cardiovascular risk factors in healthy postmenopausal women: A 2-year, placebo-controlled study. *Arterioscler Thromb Vasc Biol* 1999;19(12):2993-3000.
111. Ross PD, Davis JW, Epstein RS, Wasnich RD. Pre-existing fractures and bone mass predict vertebral fracture incidence in women. *Ann Intern Med* 1991;114(11):919-23.
112. Ross PD, Davis JW, Wasnich RD. Bone mass and beyond: risk factors for fractures. *Calcif Tissue Int* 1993;53 Suppl 1:S134-7.
113. Cummings SR. Treatable and untreatable risk factors for hip fracture. *Bone* 1996;18(3 Suppl):165S-7S.
114. Ross PD. Prediction of fracture risk. II: other risk factors. *Am J Med Sci* 1996;312(6):260-9.
115. Center for Drug Evaluation and Research, Food and Drug Administration. *Medical review: supplemental NDA for raloxifene hydrochloride: Evista*. Rockville (MD): The Center; 1999 Sep 30. Available: http://www.fda.gov/cder/foi/nda/99/20815S3_Evista.htm (accessed 2003 Jan 30).
116. Barrett-Connor E, Grady D, Sashegyi A, Anderson PW, Cox DA, Hozzowski K, 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.
117. Garnero P, Hausherr E, Chapuy MC, Marcelli C, Grandjean H, Muller C, et al. Markers of bone resorption predict hip fracture in elderly women: the EPIDOS Prospective Study. *J Bone Miner Res* 1996;11(10):1531-8.
118. Melton LJ, Khosla S, Atkinson EJ, O'Fallon WM, Riggs BL. Relationship of bone turnover to bone density and fractures. *J Bone Miner Res* 1997;12(7):1083-91.
119. Riggs BL, Melton LJ, III. Bone turnover matters: the raloxifene treatment paradox of dramatic decreases in vertebral fractures without commensurate increases in bone density. *J Bone Miner Res* 2002;17(1):11-4.
120. Sarkar S, Mitlak BH, Wong M, Stock JL, Black DM, Harper KD. Relationships between bone mineral density and incident vertebral fracture risk with raloxifene therapy. *J Bone Miner Res* 2002;17(1):1-10.
121. Cummings SR. The paradox of small changes in bone density and reductions in risk of fracture with raloxifene. *Ann N Y Acad Sci* 2001;949:198-201.
122. Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM, 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.
123. Riggs BL, Hartmann LC. Selective estrogen-receptor modulators -- mechanisms of action and application to clinical practice. *N Engl J Med* 2003;348(7):618-29.

124. Ioannidis JP, Contopoulos-Ioannidis DG. Reporting of safety data from randomised trials. *Lancet* 1998;352(9142):1752-3.
125. Ioannidis JP, Lau J. Completeness of safety reporting in randomized trials: an evaluation of 7 medical areas. *JAMA* 2001;285(4):437-43.
126. Haas S. Recommendations for prophylaxis of venous thromboembolism: International Consensus and the American College of Chest Physicians Fifth Consensus Conference on antithrombotic therapy. *Curr Opin Pulm Med* 2000;6(4):314-20.
127. Clagett GP, Anderson FA, Jr., Geerts W, Heit JA, Knudson M, Lieberman JR, et al. Prevention of venous thromboembolism. *Chest* 1998;114(5 Suppl):531S-60S.
128. Delmas PD, Genant HK, Crans GG, Stock JL, Wong M, Siris E, 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.
129. Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. *JAMA* 2002;288(3):321-33.

Appendix 1: Literature Search Strategy

Dialog® OneSearch® and Alerts®

| | |
|----|--|
| de | = descriptor, e.g. in MEDLINE®, HealthSTAR a Medical Subject Heading or MeSH, a controlled, thesaurus term; in EMBASE, etc., a thesaurus term as well. |
| ti | = title (i.e. word has to occur in title field of the bibliographic record) |
| ab | = abstract (i.e. word has to occur in abstract field of bibliographic record) |
| ! | = explode; picks up narrower terms as well, i.e. terms which are conceptually subsets of a broader term |
| () | = words must be adjacent |
| ? | = truncation symbol |
| rd | = reduce duplicates, i.e. duplicate references are removed |

| DATABASES | SEARCH TERMS |
|--|--|
| Dialog® OneSearch® Adis LMS Drug Alerts, BioBusiness, Biosis Previews®, Current Contents Search®, EMBASE, HealthSTAR, Inside Conferences, International Pharmaceutical Abstracts, MEDLINE®, PASCAL, and Toxline®. | <ol style="list-style-type: none"> 1. raloxifene/ti,ab 2. tn=raloxifene 3. keoxifene/ti,ab 4. tn=keoxifene 5. Set 1: Set 4 6. osteoporosis/de 7. osteoporosis! 8. fractures/de 9. fracture/de 10. fracture reduction/de 11. bone demineralization, pathologic/de 12. bone demineralization, pathologic!/de 13. bone atrophy/de 14. bone density/de 15. bone regeneration/de 16. bone regeneration!/de 17. bone demineralization!/de 18. calcification, physiologic/de 19. bone mineralization/de 20. osteoporosis/ti, ab 21. (facture OR fractures)/ti,ab 22. fracture()reduction?/ti,ab 23. bone()demineralization/ti,ab 24. bone()atrophy/ti,ab 25. bone()density/ti,ab 26. bone()regeneration/ti,ab 27. bone()demineralization/ti,ab 28. physiologic()calcification/ti,ab 29. bone()mineralization/ti,ab 30. Set 7 : Set 29 31. dt=clinical trial 32. clinical trial!/de 33. clinical trials/de 34. clinical trials!/de |

| | |
|--|--|
| | 35. dt=randomized controlled trial 36. dt=controlled clinical trial 37. random allocation/de 38. double-blind method/de 39. random?/ti,ab 40. placebo?/ti,ab 41. controlled ()trial?/ti,ab 42. double()blind?/ti,ab 43. dt=meta-analysis 44. meta-analysis 45. meta()analy?/ti,ab 46. metaanaly?/ti,ab 47. quantitative?()review?/ti,ab 48. quantitative?()overview?/ti,ab 49. quantitative?()review?/ti,ab 50. evidence based medicine/de 51. multicenter study/de 52. randomized controlled trial/de 53. drug comparison/de 54. comparative study/de 55. toxicity/de 56. adverse effect/de 57. contraindications/de 58. side effects/de 59. adverse reactions/de 60. adverse reaction/de 61. drug toxicity/de 62. controlled()clinical()trial?/ti,ab 63. multicenter()stud?/ti,ab 64. multicentre()stud?/ti,ab 65. toxic?/ti,ab 66. adverse/ti,ab 67. effect?/ti,ab 68. adverse()effect?/ti,ab 69. adverse()effect?/ti,ab 70. contraindicat?/ti,ab 71. side()effect?/ti,ab 72. adverse()reaction?/ti,ab 73. Set 31: Set 72 74. Set 5 AND Set 30 AND Set 73 = References 75. estrogen replacement therapy/de 76. hormone replacement therapy/de 77. hormone substitution/de 78. hormonal therapy/de 79. estrogen therapy/de 80. Set 75 : Set 79 81. Set 74 AND Set 80 = References (subset of Set 74) |
|--|--|

Appendix 2: Additional Details Concerning Methods

Selection Process: Citations were entered into a Reference Manager database, with duplicates removed manually. Two individuals (HMS, AC) independently “broad screened” the title, abstract and key words for each citation by liberally applying the eligibility criteria to determine whether to retain it. It was retained if it contained pertinent information. If one broad screener was at best uncertain as to its potential relevance, it was entered into the next phase of the review and deemed to be “potentially relevant.”

Hard copies of the full reports were retrieved and assessed for relevance independently by two reviewers (HMS, TC) who applied the eligibility criteria strictly. All disagreements were resolved by forced consensus. The reasons for excluding studies were noted. Before completing the relevance assessment, the reviewers undertook a reliability study, which yielded a 95% agreement with 20 randomly selected, full reports.³¹ The reason for the one disagreement was discussed to prevent its recurrence.

Assessment of Methodological Quality of Trials from Reports: The quality of each included trial was assessed independently by two reviewers (HMS, TC) who were familiar with the validated, three-item Jadad *et al.* scale,³³ which assesses randomization, double-blinding and the inclusion of data for dropouts and withdrawals. Total scores range from zero to five, with scores <3 indicating poor quality. The concealment of each trial’s allocation to treatment³⁴ was also evaluated by the reviewers (grade A=adequate; B=unclear; C=inadequate) (Appendix 3). Scoring differences were settled by forced consensus.

Data Synthesis and Analysis of Efficacy and Safety Data: If appropriate, binary data from trials were pooled using the fixed effects Mantel-Haenszel estimate of the relative risk.³⁵ Statistical heterogeneity of the relative risks across studies was assessed using the chi-square Q statistic, but the power of this test may be low and clinical insight may be more relevant in understanding heterogeneity.³⁶ The DerSimonian-Laird³⁷ random effects estimate of the relative risk was used to pool trials if there was between-trial statistical heterogeneity. When there was little statistical heterogeneity, the DerSimonian-Laird estimate was identical to the fixed effects estimate. As an index of precision,³⁸ a 95% confidence interval was calculated for each estimate.

When percentages of patients responding to treatment were reported, the numbers of patients were computed by multiplication with rounding. When the goal of a MA was to compare all doses of raloxifene to a comparator, different doses were pooled in trials by summing the number of patients. Where appropriate, it was planned to derive the numbers needed to treat (NNT) and harm (NNH) for primary outcome and adverse event outcome data respectively.

Continuous (i.e., BMD, one adverse event) data were used to derive the weighted mean difference. While “percent change from baseline” (i.e., BMD) was a poor outcome measure,³⁹ it could be synthesized in a fixed effects MA by computing the difference in mean percent change from baseline between treatment and control; and weighting by the reciprocal of the squared standard error. The chi-square Q statistic and the DerSimonian-Laird approach were used as required.

In any trial where multiple doses were compared to a comparator, the mean percent change from baseline for treatment was computed by taking a weighted average of the mean percent change from baseline for the different doses, with weights proportional to the inverse of the squared standard error of each mean percent change from baseline. Not all variances were available, so variance imputation was required.⁴⁰ For studies in which the standard error (SE) was available, the standard deviation (SD) was computed as:

$$SD = \sqrt{n} * SE, \text{ where } n = \text{number of participants.}$$

The median standard deviation was computed for outcomes (Tables 6 and 8). SEs were computed for the studies where they were not reported as:

$$SE = \text{median } SD / \sqrt{n}.$$

The difference in mean percent difference from baseline, between treatment and control, was computed. The SE of the difference was computed as:

$$SE = \sqrt{SE^2(\text{treatment}) + SE^2(\text{control})}.$$

Sensitivity and Subgroup Analyses: Where baseline population data were unavailable for a given study, the inclusion criteria were used in Appendix 4 and to derive summary statistics (Tables 1 to 3). If the inclusion criterion information was expressed as a range (e.g., two to eight years post-menopause) or as a minimum requirement (e.g., ≥ 2 years post-menopause), the midpoint and the lower bounds respectively were entered into calculations.

Publication Bias: Funnel plots (e.g., effect size versus precision) involving primary outcomes were to be inspected, while tests of publication bias were to be used where appropriate. The latter yielded quantitative views of a funnel plot's degree of asymmetry (rank correlation test, graphical test⁴³), the number of "missing" or unobserved trials given the dispersion of trial estimates (trim and fill method⁴⁴) and the pooled estimate of efficacy or safety when adjusted for the impact of publication bias (trim and fill method).

Appendix 3: Tools to Assess Quality of Randomized Controlled Trials

1. Randomization:

Is the study described as randomized (i.e., including words such as randomly, random, randomization)?

Yes = 1 No = 0 = ___

A trial reporting that it is “randomized” receives one point. Trials describing an appropriate method of randomization (table of random numbers, computer generated) receive an additional point.

Appropriate = 1 Not appropriate = 0 = ___

If the report describes the trial as randomized and uses an inappropriate method of randomization (e.g., date of birth, hospital numbers), a point is deducted.

Total points: 0 1 2 Score = ___

2. Double-blinding: Is the study described as double-blind?

Yes = 1 No = 0 = ___

A trial reporting that it is double-blind receives one point. Trials that describe an appropriate method of double-blinding (identical placebo: colour, shape, taste) receive an additional point.

Yes = 1 No = 0 = ___

If the report describes the trial as double-blind and uses an inappropriate method (e.g., comparison of tablets versus injection with no dummy), a point is deducted.

Total points: 0 1 2 Score = ___

3. Withdrawals and dropouts: Is there a description of withdrawals and dropouts?

Yes = 1 No = 0 Score = ___

A trial reporting the number of and reasons for withdrawals or dropouts receives one point. If there is no description, no point is given.

Overall score: ___

Low = 0 to 2 points

Moderate = 3 to 4 points

High = 5 points (maximum)

4. Adequacy of Allocation Concealment: (circle one)

- Central randomization; numbered or coded bottles or containers; drugs prepared by a pharmacy; serially numbered, opaque, sealed envelopes = Adequate
- Alternation; reference to case record number or date of birth = Inadequate
- Allocation concealment is not reported or fits neither category = Unclear

Appendix 4: Raloxifene for Primary and Secondary Prevention of Osteoporosis in Postmenopausal Women

| Author (Lilly Trial Id) | Yr | N | Mean Age (SD); Range | Years Post- menopause (Mean; SD) (Range) | Trialist's Stated Goal of Osteoporosis Intervention | Intervention Length | Raloxifene Intervention (mg/day) (Supplements) | Comparator Intervention Dose | Publication Status | Number of Centres | Trial Quality: Total Jadad Score/Allocation Concealment |
|--|------------------------------|-------|---|--|--|---|--|------------------------------------|--|------------------------------|--|
| Secondary prevention trials:^a efficacy and safety data | | | | | | | | | | | |
| Delmas Eastell | 2000 2000 | 7,705 | 66.5 (6.8) yrs; 31 to 80 yrs | 19.3 (8.0) yrs (>2 yrs*) | Tx: substudy 1: <2.5 SD below peak femoral neck or lumbar BMD;* | 4 yrs (mean=39.4 mos; median=47.4 mos) | RLX 60, RLX 120 (Ca 500 mg/day; vitamin D 400 to 600 IU/day) | Placebo | Abstract Abstract | 180 (25 countries) | 4/adequate |
| Ettinger | 1999 | | | | substudy 2: low BMD (undefined) and ≥1 moderate or severe VFx or ≥2 mild VFXs or ≥2 moderate VFxs* | 2 yrs (median= 28.9 mos) | in year 4, option to use bone- active agents other than oral estrogens | | Journal | | |
| Ensrud | 1998 | | | | | 2 yrs | | | Abstract | | |
| Ettinger Ettinger | 1998 1998 | | | | | 2 yrs 2 yrs | | | Abstract Abstract | | |
| Cauley Cummings Cummings Maricic | 2000 1999 1998 2000 | | | | | 4 yrs 3 yrs 2 yrs 52 wks | | | Abstract Journal Abstract Abstract | | |
| MORE trial (GGGK) | | | | | | | | | | | |
| Lufkin Lufkin Lufkin Lufkin Nickelsen | 1998 1998 1997 1999 | 143 | 68.4 (5.0) yrs; 51 to 76 yrs; 45 to 75 yrs* | 22.6 (1.1) yrs (NR) ≥5 yrs*, or estradiol and FSH levels | Tx: BMD <10 th percentile and ≥1 non-traumatic vertebral Fx* | 52 wks | RLX 60, RLX 120 (Ca 750 mg/day; vitamin D 800 IU/day) | Placebo | Journal Abstract Abstract Abstract Journal | 2 (US) | 3/unclear |
| (GGGN) | | | | | | | | | | | |

Id=identification; yr=year; N or n=sample size; RLX=raloxifene; SD=standard deviation; SE=standard error; wks=weeks; mos=months; BMD=bone mineral density; Fx=fracture; VFX=vertebral Fx; NR=not reported; Lilly=Eli Lilly Inc.; Px=prevention; Tx=treatment; CEE=conjugated equine estrogen; *inclusion criteria; Ca=calcium; FSH=follicle stimulating hormone; ^aaccording to World Health Organization's (WHO) BMD definition of osteoporosis, osteopenia and normal; ^btwo identical trials [Delmas/GGGF (n= 601) and GGGG (n=544)]; ^ca crossover at 3 mos for topical vaginal interventions; likely a Lilly trial=authors' associations suggest Lilly sponsorship yet unconfirmed via trial code information provided by Canadian representative of Eli Lilly.

| Author (Lilly Trial Id) | Yr | N | Mean Age (SD); Range | Years Post- menopause (Mean; SD) (Range) | Trialist's Stated Goal of Osteoporosis Intervention | Intervention Length | Raloxifene Intervention (mg/day) (Supplements) | Comparator Intervention Dose | Publication Status | Number of Centres | Trial Quality: Total Jadad Score/Allocation Concealment |
|---|------------------------------|-----|---|--|--|---------------------------------------|--|---|---|---|--|
| Primary and secondary prevention trials (i.e., osteoporotic and osteopenic women): ^aefficacy and safety data | | | | | | | | | | | |
| Meunier Meunier Meunier Sarkar (GGGP) | 1999 1998 1998 1998 | 129 | 60.2 (6.7) yrs; 50 to 75 yrs* | 12.5 (9.0) yrs (NR) ≥1 yr* or estradiol and FSH levels | Tx: ≤2.8 below BMD peak with ≤3 previous Fx; 60% were <2.5 SD below | 104 wks 52 wks 52 wks 52 wks | RLX 60, RLX 150 (Ca 1,000 mg/day; vitamin D 300 IU/day) | Placebo | Journal Abstract Abstract Abstract | 8 (France) | 3/unclear |
| Johnell Johnell Stock (GGHV) | 2000 1999 1999 | 330 | NR (NR); NR (≤75 yrs*) | NR (NR) (≥2 yrs*) | Tx and Px: <2 SD below peak femoral neck BMD* | 52 wks | RLX 60; RLX 60 + ALN 10 mg/day (Ca 0.5 g/day; vitamin D 400 IU/day to 600 IU/day) | Placebo; ALN 10 mg/day | Abstract Abstract Abstract | NR (Sweden? Australia? Belgium?) | 2/unclear |
| Morii Morii (likely a Lilly trial) | 1997 1996 | 93 | 56 (NR) yrs; 38 to 65 yrs | Median= 7 (NR) yrs (NR) (serum estradiol and FSH levels confirmed) | Tx and Px: <2 SD below peak BMD* | 12 wks of 24 wk trial | RLX 30, RLX 90 (Ca 800 mg/day) | Placebo | Abstract Abstract | NR (Japan) | 1/unclear |
| Primary and secondary prevention trials (i.e., osteoporotic, osteopenic and “healthy” women): ^aefficacy and safety data | | | | | | | | | | | |
| Prestwood Gunness Prestwood (GGGM) | 2000 1997 1997 | 51 | 64.4 (5.8) yrs; 55 to 85 yrs* | 18.2 (8.5) yrs (≥5 yrs*) | Tx and Px: 3.0 SD below peak BMD to 1 SD above* | 26 wks | RLX 60 (NR) | CEE 0.625 mg/day (women with uterus given Provera in last 2 wks) | Journal Abstract Abstract | NR? | 3/unclear |

Id=identification; yr=year; N or n=sample size; RLX=raloxifene; SD=standard deviation; SE=standard error; wks=weeks; mos=months; BMD=bone mineral density; Fx=fracture; VFx=vertebral Fx; NR=not reported; Lilly=Eli Lilly Inc.; Px=prevention; Tx=treatment; CEE=conjugated equine estrogen; *inclusion criteria; Ca=calcium; FSH=follicle stimulating hormone; ALN=alendronate sodium; ^aaccording to World Health Organization's (WHO) BMD definition of osteoporosis, osteopenia and normal; ^btwo identical trials [Delmas/GGGF (n= 601) and GGGG (n=544)]; ^ca crossover at 3 mos for topical vaginal interventions; likely a Lilly trial=authors' associations suggest Lilly sponsorship yet unconfirmed via trial code information provided by Canadian representative of Eli Lilly.

| Author (Lilly Trial Id) | Yr | N | Mean Age (SD); Range | Years Post- menopause (Mean; SD) (Range) | Trialist's Stated Goal of Osteoporosis Intervention | Intervention Length | Raloxifene Intervention (mg/day) (Supplements) | Comparator Intervention Dose | Publication Status | Number of Centres | Trial Quality: Total Jadad Score/Allocation Concealment |
|---|--|-------|--|---|---|--|--|--|--|---|--|
| Primary prevention trial (i.e., osteopenic and "healthy" women): ^aefficacy and safety data | | | | | | | | | | | |
| Johnston ^b Shah ^b Bjarnason Delmas Delmas Bjarnason Delmas (GGGF) | 2000 1998 1998 1999 1997 1997 1997 | 1,145 | 54.6 (SE=0.01) yrs; ^b 45 to 60 yrs* | NR (NR) ^b 4.8 (2.0) yrs= Delmas trial only (2 to 8 yrs*) | Px: 2.5 SD below peak BMD to 2.0 SD above peak* =55% osteopenic and 45% "normal" participants (for Delmas data only) | 36 mos 24 mos 36 mos 24 mos 24 mos 24 mos | RLX 30, RLX 60, RLX 150 (Ca 400 to 600 mg/day) | Placebo | Journal Abstract Abstract Abstract Journal Abstract Abstract | 11: GGGF + North America (GGGG) (GGGF=8 European countries) | 4/unclear |
| Tsai (a Lilly trial) | 2001 | 116 | 56.9 (4.2) yrs; 47 to 66 yrs | NR (NR) | Px: 2.5 SD below, to 1 SD above, peak bone mass* | 12 mos (no hysterectomy*) | RLX 60 (Ca 500 mg/day) | CEE 0.625 mg/day + 5 mg/day medroxy- progesterone acetate | Journal | 2 (Taiwan) | 3/unclear |
| Primary prevention trial (i.e., no additional population definition): ^aefficacy and safety data | | | | | | | | | | | |
| Pavo Stock (Lilly trial but no ID provided) | 1999 1999 | 128 | 59 (NR) yrs; NR | NR (NR) (≥2 yrs*) | Px: no selection criteria or baseline data reported | 12 mos | RLX 60 (no supplements or vitamins allowed) | Placebo | Abstract Abstract | ≥2 (i.e. multi- centre) (Russia) | 2/unclear |

Id=identification; yr=year; N or n=sample size; SD=standard deviation; RLX=raloxifene; SE=standard error; wks=weeks; mos=months; BMD=bone mineral density; Fx=fracture; VFx=vertebral Fx; NR=not reported; Lilly=Eli Lilly Inc.; Px=prevention; Tx=treatment; CEE=conjugated equine estrogen; *inclusion criteria; Ca=calcium; FSH=follicle stimulating hormone; ^aaccording to World Health Organization's (WHO) BMD definition of osteoporosis, osteopenia and normal; ^btwo identical trials [Delmas/GGGF (n= 601) and GGGG (n=544)]; ^ca crossover at 3 mos for topical vaginal interventions; likely a Lilly trial=authors' associations suggest Lilly sponsorship yet unconfirmed via trial code information provided by Canadian representative of Eli Lilly.

| Author (Lilly Trial Id) | Yr | N | Mean Age (SD); Range | Years Post- menopause (Mean; SD) (Range) | Trialist's Stated Goal of Osteoporosis Intervention | Intervention Length | Raloxifene Intervention (mg/day) (Supplements) | Comparator Intervention Dose | Publication Status | Number of Centres | Trial Quality: Total Jadad Score/Allocation Concealment |
|--|--|-----|--|---|--|---|---|---|---|----------------------------|--|
| Primary prevention trials (i.e., osteopenic or 'healthy' women): ^asafety data only | | | | | | | | | | | |
| Fugere Fugere Shah Jolly Glant Scheele Scheele Prestwood (GGGZ) | 2000 1999 1998 1997 1997 1997 1997 1997 | 136 | 56 (5.1) yrs; 45 to 66 yrs | 5.9 (4.5) yrs (1 to 17 yrs) | Px Healthy* | 24 mos 24 mos 12 mos 12 mos 12 mos 12 mos 12 mos (no hysterectomy*) | RLX 150 (NR) | CEE 0.625 mg/day + medroxy- progesterone acetate 2.5 mg/day | Journal Abstract Abstract Abstract Abstract Abstract Abstract | 11 (Canada) | 3/unclear |
| Goldstein Thiebaud Goldstein (GGHG) | 2000 1999 1998 | 415 | 54.7 (3.5) yrs; NR 47 to 60 yrs* | 4.6 (1.9) yrs (2 to 8 yrs*) | Px: healthy* 2.5 SD below to 2 SD above peak BMD* (i.e. osteopenic and "normal") | 12 mos (no hysterectomy*) | RLX 60, RLX 150 (Ca 520 mg/day) | Placebo: CEE 0.625 mg/day | Journal Abstract Abstract | 33 (US=32, UK=1) | 5/adequate |
| Boss Draper Huster (GGGB/US) | 1997 1996 1996 | 251 | 53.1 (3.4) yrs; 45.2 to 60.9 yrs | Median=35.6 (NR) mos; (6 to 87 mos) | Px: healthy* | 8 weeks (no hysterectomy*) | RLX 200, RLX 600 (Ca 520 mg/day) | Placebo; CEE 0.625 mg/day | Journal Journal Abstract | 11 (US) | 4/unclear |

Id=identification; yr=year; N or n=sample size; SD=standard deviation; RLX=raloxifene; SE=standard error; wks=weeks; mos=months; BMD=bone mineral density; Fx=fracture; VFx=vertebral Fx; NR=not reported; Lilly=Eli Lilly Inc.; Px=prevention; Tx=treatment; CEE=conjugated equine estrogen; *inclusion criteria; Ca=calcium; FSH=follicle stimulating hormone; ^aaccording to World Health Organization's (WHO) BMD definition of osteoporosis, osteopenia and normal; ^btwo identical trials [Delmas/GGGF (n= 601) and GGGG (n=544)]; ^ca crossover at 3 mos for topical vaginal interventions; likely a Lilly trial=authors' associations suggest Lilly sponsorship yet unconfirmed via trial code information provided by Canadian representative of Eli Lilly.

| Author (Lilly Trial Id) | Yr | N | Mean Age (SD); Range | Years Post- menopause (Mean; SD) (Range) | Trialist's Stated Goal of Osteoporosis Intervention | Intervention Length | Raloxifene Intervention (mg/day) (Supplements) | Comparator Intervention Dose | Publication Status | Number of Centres | Trial Quality: Total Jadad Score/Allocation Concealment |
|--|------------------------------|-----|--|---|--|--|---|---|---|--|--|
| Freedman San Martin Cohen Laksh-manan (GGGH) | 2001 2000 1998 1999 | 619 | 52.9 (4.7) yrs (based on n= 168); 45 to 60 yrs* | 6.1 (4.2) yrs based on only n=168; (NR) | Px healthy* no breast cancer | 24 months (breast data; n= 168) 36 mos (vaginal data; n= 619) (5 yr trial) (hysterectomy <15 yrs before study*) | RLX 60, RLX 150 (NR) low dose vaginal estrogens (except estradiol) up to 3 times per wk during trial | Placebo; CEE 0.625 mg/day | Journal Abstract Abstract Abstract | 38 (10 countries in North America, Europe, South Africa, New Zealand) | 3/unclear |
| Vardy (not a Lilly trial?) | 1998 | 33 | NR (NR) | NR (NR) | Px: healthy* | 18 wks | RLX 60 (NR) | Placebo; tamoxifen 20 mg/day; CEE 0.625 mg/day | Abstract | NR | 2/unclear |
| Parsons Parsons Parsons (GGIK) | 1999 1999 1999 | 187 | 59 (NR) yrs; NR | 9.5 (NR) yrs (NR) | Px: naturally post- menopausal;* genito-urinary atrophy* | 3 mos ^c | RLX 60 (NR) | Placebo; open label vaginal Premarin 0.5 g/day or Replens ^c | Abstract Abstract Abstract | NR | 2/unclear |
| Walsh Walsh Walsh (GGGY) | 1998 1997 2000 | 390 | 59.3 (6.3) yrs; 45 to 72 yrs* | 11.0 (7.5) yrs; (≥1 yr*) | Px: Healthy* | 6 mos (if hysterectomy, then 50 to 72 yrs*) | RLX 60, RLX 120 (NR) | Placebo: CEE 0.625 mg/day + medroxy- progesterone acetate 2.5 mg/day | Journal Abstract Journal | 8 (US) | 5/unclear |
| Mijatovic Mijatovic Valk deRoo (likely a Lilly trial) | 1999 1998 1999 | 60 | 54.8 (3.5) yrs; NR | NR (NR) (NR) | Px: healthy* | 24 months (5 yr trial) (all hysterecto- mized*) | RLX 60, RLX 150 (Ca 500 mg/day) | Placebo: CEE 0.625 mg/day | Journal Journal Journal | 1 (Nether- lands) | 3/unclear |

Id=identification; yr=year; N or n=sample size; SD=standard deviation; RLX=raloxifene; SE=standard error; wks=weeks; mos=months; BMD=bone mineral density; Fx=fracture; VFx=vertebral Fx; NR=not reported; Lilly=Eli Lilly Inc.; Px=prevention; Tx=treatment; CEE=conjugated equine estrogen; *inclusion criteria; Ca=calcium; FSH=follicle stimulating hormone; ^aaccording to World Health Organization's (WHO) BMD definition of osteoporosis, osteopenia and normal; ^btwo identical trials [Delmas/GGGF (n= 601) and GGGG (n=544)]; ^ca crossover at 3 mos for topical vaginal interventions; likely a Lilly trial=authors' associations suggest Lilly sponsorship yet unconfirmed via trial code information provided by Canadian representative of Eli Lilly.

Appendix 5: Classification of Trials by Bone Mineral Density of Participants at Baseline

In each qualitative or quantitative synthesis, the goal was to use trial populations' respective baseline BMD data to categorize the studies (e.g., low versus high BMD). Some studies did not select populations according to this scheme. Thus, several trials contained a heterogeneous collection of BMD levels among the enrolled population (i.e., category 5 below) (Appendix 4). To avoid combining data from homogeneously defined populations with very low BMD and homogeneously defined studies with BMD levels higher than the 2.5 SD threshold, it was acknowledged a priori that there were five possible trial types, each of whose data could be synthesized separately:

1. homogeneous (i.e., homog) very low BMD: trials where all participants had BMD <2.5 SD below the young adult peak bone mass
2. heterogeneous (i.e., heterog) very low BMD: category 1 very low BMD trials and category 5 (i.e. mixed very low and higher BMD level) trials
3. homogeneous (i.e., homog) higher BMD: trials where all participants were diagnosed as having BMD higher than the 2.5 SD threshold (2.5 SD below the young adult peak bone mass)
4. heterogeneous (i.e., heterog) higher BMD: category 3 trials and category 5 trials
5. mixed low and high BMD: in each trial, there were participants who had a full range of BMD levels and who would all be receiving raloxifene (this is not a mix of treatment and prevention).

Adding category 5 trials to each of category 1 and 3 trial collections of pooled participants yielded two heterogeneous groups of women wherein fewer than all (i.e., <100%) of the participants had very low BMD and some had BMD levels greater than the 2.5 SD threshold. Qualitative summaries (section 3.2.1, Tables 1 to 3) and a meta-analysis (section 3.2.2, Tables 4, 6 to 8, Figures 2 to 5) were planned to evaluate the possible impact of varying BMD definitions of "population" on the empirical picture of raloxifene's efficacy and safety.

These categories were not mutually exclusive; one trial may be included in several categories. For example, all trials in category 1 were also listed in category 2. In the MORE trial, the two subgroups fitted into different categories: subgroup 1 was low BMD; subgroup 2 was mixed, as women with prevalent fractures were enrolled regardless of their BMD.