uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Supriya Mayank Goyal
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Epidemiology)
GRADE / DEGREE

Department of Epidemiology and Community Medicine
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

The Effects of Hormone Replacement Therapy on Coronary Heart Disease in Postmenopausal Women: Systematic Review and Assessment of Methods

TITRE DE LA THÈSE / TITLE OF THESIS

George Wells
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

Heather Clark
CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Julian Little

Peter Tugwell

Gary W. Slater
LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES /
DEAN OF THE FACULTY OF GRADUATE AND POSTDOCORAL STUDIES

THE EFFECTS OF HORMONE REPLACEMENT THERAPY ON CORONARY HEART DISEASE IN POSTMENOPAUSAL WOMEN: SYSTEMATIC REVIEW AND ASSESSMENT OF METHODS

by

Supriya Mayank Goyal

Thesis submitted to the School of Graduate Studies and Research in partial fulfillment of the requirements for the M.Sc. degree in Epidemiology

University of Ottawa

Thesis Supervisors:

George A. Wells, MSc, PhD

and

Heather Clark, MSc, MD

Your file Votre référence
ISBN: 0-494-14912-4
Our file Notre référence
ISBN: 0-494-14912-4

ABSTRACT

Hormone replacement therapy (HRT) was commonly prescribed to postmenopausal women to decrease the risk of coronary artery disease. Non-randomized studies have demonstrated upto 40% reduction in women using HRT. The first randomized trial to study this association failed to find the anticipated cardiac benefit. Since then, several trials addressing primary or secondary prevention of heart disease with HRT have failed to find any association. This thesis analyzed all available data on HRT and heart disease to identify factors contributing to the divergent results. We found that in non-randomized studies, current users of HRT had a decreased risk of heart disease while the summary statistic for RCTs failed to show any association. Measurement of exposure and socioeconomic status were important confounding variables. The strengths and limitations of study designs were discussed. A second component of the thesis was the modification of Newcastle Ottawa scale, a quality assessment tool for non-randomized studies.

Executive Summary

Objectives: 1) To conduct a systematic review and a meta-analysis of the pertinent literature to evaluate the association between the risk of coronary artery disease (CAD) and hormone replacement therapy (HRT) among postmenopausal women. 2) To develop and evaluate a modified Newcastle Ottawa scale (NOS) and compare it with the original scale.

Methods: For the systematic review, a comprehensive search strategy was developed to search electronic databases (Medline, Embase, Current Contents, Premedline, Cochrane Controlled Trials Register, Cochrane Reviews, Biological Abstracts, Dissertation Abstracts, Pubmed and Biocentral) and the key journals and conference proceedings were hand searched. Randomized controlled trials (RCTs) and non-randomized studies (study designs) examining the association between exposure to HRT (intervention) and CAD (outcome) among postmenopausal women (population) were eligible for inclusion in the review. Data abstraction was done using standardized forms. Data on publication status, population characteristics, study intervention, outcome and on adjusted risk factors was collected. Two independent reviewers assessed the quality of studies using Jadad's scale and allocation concealment for the RCTs and the Newcastle Ottawa scale (NOS) for the non-randomized studies. Data synthesis was performed when appropriate by using the fixed effects model to calculate the summary estimate and its standard error for RCTs, case control and cohort studies. Statistical tests for association and homogeneity were performed. Pertinent subgroup and sensitivity analysis was done.

For the modification of NOS, the original scale was critically appraised with input from other users of the scale. Qualitative analysis was done to identify factors that could improve the performance of the comparability component of the original NOS. The new items generated were then tested for validity (face, content and criterion) and reliability. The results of the original and modified NOS were compared.

Results: Seven RCTs, 23 case control and 30 cohort studies were included in the systematic review. Of the seven trials, four were on primary prevention of CAD and HRT use. The summary relative risk for RCTs on secondary prevention was 1.01 (95% CI: 0.86, 1.19). None of the trials showed a decrease in CAD with HRT use. The summary odds ratio for case control studies for current users of HRT and overall CAD was 0.75 (95% CI: 0.66, 0.84) and for fatal CAD was 0.47 (95% CI: 0.32, 0.69). The summary relative risk for cohort studies for current users of HRT and overall CAD was 0.60 (0.52, 0.70) and for fatal CAD was 0.65 (0.57, 0.75). Among ever users of HRT the incidence of fatal CAD was decreased in both non-randomized study designs. Further analysis did not show any association between CAD and duration of HRT use or with use of estrogen alone versus estrogen and progestin. The decreased risk was most consistently seen with 0.625 mg of estrogen dose and in women who initiated HRT early after menopause. On combining the summary estimates from case control studies that adjusted for socioeconomic status as a confounder the decrease in the risk of CAD with HRT use was

not as marked and the upper CI was 1.00. Similar analysis among cohort studies continued to show a cardio-protective effect with HRT use. There was no association between study quality and study results. Addition of unpublished literature accentuated the treatment results for both the non-randomized study designs. Among cohort studies on women with pre-existing heart disease, initiating HRT after a cardiac event resulted in increased incidence of recurrent CAD (RR: 1.25, 95% CI: 0.97, 1.62) while in women who had been on HRT at the time of their initial cardiac event, the risk of a recurrent CAD event was decreased (RR: 0.58, 95% CI: 0.42, 0.81).

The comparability component of the NOS was modified and expanded to include three instead of two items in the original scale. The face and content validity of these items was good. The items were clearly worded and addressed the question of comparability between the two groups. The scoring of the modified NOS compared favorably with the scoring of Black and Down's scale, another validated scale for quality assessment. The Spearman's correlation coefficient was 0.91 ($p<0.01$) and 0.77 ($p<0.01$) for case control and cohort studies respectively demonstrating a high degree of correlation between the two scales. The inter-rater reliability was high with a Cohen's kappa of 0.94 and it was easier and faster to apply than Black and Down's scale. On comparison with the original NOS, the new item evaluating studies adjusting for a group of pre-determined risk factors improved the performance of the instrument.

Conclusions: This meta-analysis differs from previous meta-analyses as it includes the results of the RCTs which are discordant from the results of non-randomized studies. Evaluation of explanatory variables demonstrated the difficulties in accurate measurement of an exposure such as HRT in non-randomized studies as well as the potential confounding by socioeconomic status. These explain some of the discrepancy between the results of the two study designs.

The modified NOS is a good instrument for assessment of quality of non-randomized studies but needs to be tested in other areas of research.

Acknowledgements

The completion of this thesis has been a challenge and the task would have been abandoned had it not been for the continued support, encouragement and guidance from my mentor and supervisor, Dr George Wells. Thanks Dr Wells, for believing that I could accomplish this goal. Despite being faced by his own challenges, he made time for innumerable meetings going in detail about meta-analyses of non-randomized studies and providing lucid explanations for my doubts and questions. I would also like to extend my thanks to Dr Heather Clark, my co-supervisor, who has been readily available and provided thorough reviews of my various thesis drafts. I would like to especially thank her for being the independent reviewer for quality assessment of several articles, which was a laborious task. A special thanks to Laura Wells, for doing quality assessment of all studies using the modified Newcastle Ottawa scale.

A heart felt thanks to Anne Gray for always being there, coordinating meetings with George, providing valuable advice on formatting and giving finishing touches to the thesis. I would like to thank Joan Peterson, Dr Ann Cranny, Dr Elaine Jolly and Dr Ian McDonald for their advice on the review. I would also like to thank Fay Draper and Dr Rama Nair for their support during this period.

I would like to recognize the technical support provided by Margaret Sampson and Jesse McGowan.

A very special and loving thanks to my sons, Dhruv and Arnuv, for patiently waiting for all those hours that I should have played with them but had to work on my thesis. Thanks to my mother, mother-in-law and Marie, my children's caregiver and friends for their constant support through trying times. I have no words to thank Mayank, my husband, who has stood like a pillar of strength behind me through good days and bad days and without whose support this endeavor would not achieved fruition.

TABLE OF CONTENTS

List of Tables

List of Figures

List of Abbreviations

HRT	Hormone replacement therapy
CAD	Coronary artery disease
PEPI	Post menopausal estrogen progestin interventions
HERS	Heart and estrogen/progestin replacement study
WHI	Women's health initiative
RCT	Randomised controlled trial
NOS	Newcastle Ottawa Scale
CVD	Cerebrovascular disease
PVD	Peripheral vascular disease
IHD	Ischemic heart disease
CHD	Coronary heart disease
MI	Myocardial infarction
UA	Unstable angina
CHF	Congestive heart failure
CRP	C reactive protein
LDL-C	Low density lipoprotein cholesterol
HT	Hypertension
SBP	Systolic blood pressure
DBP	Diastolic blood pressure
HDL-C	High density lipoprotein cholesterol
TGs	Triglycerides
NCEP	National cholesterol education program
BMI	Body mass index
DM	Diabetes mellitus
HbA1c	Hemoglobin A1c
OR	Odds ratio
RR	Relative risk
IL-6	Interleukin 6

hs CRP	High sensitivity C-reactive protein
VLDL-C	Very low density lipoprotein cholesterol
NETA	Norethindrone acetate
MPA	Medroxyprogesterone acetate
MP	Micronized progesterone
NGM	Norgestimate
QCA	Quantitative coronary angiography
CAC	Coronary artery calcium
CEE	Conjugated equine estrogen
ERA	Estrogen – replacement and atherosclerosis study
PHOREA	Post menopausal hormone replacement against atherosclerosis
WELL-HART	Women's estrogen – progestin lipid lowering hormone atherosclerosis regression trial
IMT	Intima-media thickness
WAVE	Women's angiographic vitamin and estrogen trial
EPAT	Estrogen in the prevention of atherosclerosis trial
EE	Ethinyl estradiol
MOOSE	Meta-analysis of observational studies
NRSMG	Non-randomised studies methods group
SERMS	Selected estrogen receptor modulators
DVT	Deep vein thrombosis
CHA	Canadian Heart Association
PHASE	Papworth HRT atherosclerosis study
ESPRIT	Estrogen in the prevention of reinfarction trial
WISDOM	Women's international study of long duration oestrogen after menopause
GHC	Group health cooperative
HTA	Health technology assessement
USPTFC	US Preventive Task Force
CAST	Cardiac arrhythmia suppression trial

Chapter 1.0

INTRODUCTION

1.1 Menopause and hormone replacement therapy (HRT)

Menopause ends the reproductive period of a woman's life with a decrease in the levels of estrogen and progesterone and cessation of menstruation. Many women experience a number of symptoms around the perimenopausal period. These include vasomotor symptoms (hot flashes, night sweats, palpitations), genitourinary complaints (vaginal dryness, dyspareunia, stress incontinence, urgency, recurrent urinary tract infections), sleep disturbances (insomnia, fatigue, early wakefulness), mood disturbances (anxiety and irritability), and accelerated decline in bone mass leading to osteoporosis and increased fractures. The diminished levels of circulating estrogen are responsible for these symptoms and this was the original rationale for recommendation of hormone replacement therapy (HRT). HRT has been prescribed for short periods of time for control of vasomotor symptoms and for longer periods for prevention of osteoporosis and heart disease.

1.2 Historical timeline for HRT and coronary heart disease (CAD)

HRT was first utilized as a medication in the early 20th century. In the 1930s early formulations of HRT were comprised of estrogen alone, either as a patch or as an oral preparation. These preparations were effective in relieving vasomotor symptoms and led to widespread acceptance of this new therapy. In the late 1960s, the first clinical trial to study the association between heart disease and estrogen replacement was actually performed in men with pre-existing heart disease. This trial was known as the Coronary Drug Project[1] and was stopped earlier than planned due to an increased risk of thromboembolic and cardiovascular events in the treatment group. The second clinical trial to study this association was not undertaken for another 23 years. This time it was performed in postmenopausal women.

In the 1970s, it was first recognized that unopposed estrogen exposure in women with an intact uterus led to an increase in endometrial cancer and therefore progestins

were added to counter this effect. Epidemiological evidence suggested a steep increase in heart disease in women around the time of menopause. At the time, this increase in heart disease was attributed to the diminished levels of circulating estrogen. In the 1980s, a deluge of studies were published looking at the association between HRT and heart disease. In the early 1990s, two meta-analyses [2;3]were published that unequivocally concluded HRT decreased the incidence of heart disease and this benefit would outweigh any risks of HRT. These meta-analyses pooled results from several non-randomized studies. Following the publication of these results HRT was widely endorsed for prevention of heart disease in postmenopausal women.

HRT was available in several preparations and premarin (a conjugated estrogen preparation) was one of the most widely dispensed prescription drug in the United States between 1990 to1995. The first combination HRT pill was introduced in the late 1990s and was called prempro (premarin with medroxyprogesterone acetate (MPA)).

Three pertinent clinical trials addressing the association of HRT and CAD were started in late 1990s: The Postmenopausal Estrogen/Progestin Interventions (PEPI) [4]trial of HRT and CAD risk factors, the Heart and Estrogen/progestin Replacement Study (HERS)[5] of HRT and the effect on secondary prevention of CAD, and the National Institute of Health sponsored Women's Health Initiative (WHI)[6] of HRT and the effect on primary prevention of CAD.

The results of the PEPI trial demonstrated an improvement in the lipid profiles of participants. However, results of the HERS trial failed to show any decrease in clinical coronary events. The WHI had two parallel arms, one studying the effects of estrogen and progestin in women with an intact uterus and one studying the effects of estrogen alone in women who had undergone hysterectomy. The WHI arm studying the effect of estrogen with progestin on CAD was stopped early in 2002 due to an increase in the incidence of invasive breast cancer in the participants that exceeded the designated boundary of the planned stopping rules which were a part of the design of the study. The effects on CAD were equivocal. The second arm of WHI studying the effect of unopposed estrogen on CAD continued for an additional 20 months after termination of the first arm. This was also stopped earlier than scheduled by the data safety and monitoring board due to an increased risk of stroke in participants and the likelihood that neither cardioprotection nor

breast cancer risk would be demonstrated in the remaining intervention period. The study reported a lack of association between the use of HRT and incidence of heart disease. In view of this new data, HRT and its association with heart disease is being subjected to tremendous scrutiny. There is dichotomy between results from non-randomized studies and the more recent clinical trials.

1.3 CAD and HRT: unresolved issue

The association between HRT and CAD has been addressed in basic science studies, animal studies, non-randomized studies, trials using surrogate markers and more recently in randomized controlled trials (RCTs). The results from these studies have been divergent. The data on effect of HRT on coronary risk factors and results of non-randomized studies suggested a decreased risk of CAD among HRT users. None of the randomized trials found this protective effect. These results need to be analyzed further to understand where the truth lies. For example, the studies need to be compared on the population of women included, outcomes of interest and exposure characteristics (different preparations, routes, and doses of HRT). There are well-established risk factors for development of CAD and evaluation of these studies should assess whether these factors have been taken into consideration in an appropriate fashion. Non-randomized studies have certain inherent biases that need to be addressed and similarly, the limitations of the trials need to be reviewed. A detailed analysis of all these factors in the available literature may provide some definitive answers, as well as identify the unanswered questions that need further research and would help in our clinical decision making. This translated into the primary objective of this thesis which was to conduct a systematic review.

1.4 Assessing quality of studies

One important component in the conduction of a systematic review is assessing the quality of individual studies. Inclusion of poor quality studies can exaggerate the summary statistic[7]. The effect of study quality on results is often evaluated by a sensitivity analysis. There are several quality assessment tools available in the literature but majority of these are designed for RCTs. The Newcastle Ottawa scale (NOS) has

been identified as one of the six quality assessment tools which could be used for the quality assessment of non-randomized (case-control and cohort) studies[8]. This scale has three components addressing the selection, comparability and exposure or outcome ascertainment of the two groups being studied. The comparability component of the NOS has two items addressing the control of confounding variables. However, experience has demonstrated that most studies score a star on both items and differentiation between studies achieving more comparable groups from those that fail to achieve comparable groups is not distinct. An attempt was made to modify this component of the scale to improve the performance of this instrument. This was the secondary objective of this thesis.

1.5 Objectives of the Study

1) The first objective of this thesis was to systematically examine the association between the risk of CAD and HRT among postmenopausal women. This was achieved through a systematic review of the literature and meta-analyses of all appropriate studies within each study design. The results of non-randomized studies were compared and contrasted with those of randomized controlled trials with emphasis on the strengths and the weaknesses of the study designs.

2) The second objective of the thesis was to develop and evaluate the use of a modified Newcastle Ottawa Scale (NOS) to assess the quality of non-randomized studies and compare it with the original NOS.

Chapter 2.0

BACKGROUND

This chapter consists of two main sections. The first section describes menopause and the physiologic changes in the cardiovascular system around menopause. Coronary artery disease is discussed with emphasis on epidemiology, clinical presentation, risk factors for CAD and the unique presentation in women. The effects of exogenous estrogen and progesterone on the risk factors for CAD are explored. Studies on HRT and surrogate markers of CAD are discussed and the currently available preparations of HRT are presented. The second section focuses on systematic reviews with emphasis on limitations of this tool for non-randomized studies. The role of quality assessment of individual studies in a systematic review is explored and the instrument used for this review is discussed.

2.1 Natural history of menopause

The transition period in a woman's life from a reproductive to a non-reproductive phase is defined as the climacteric and is a natural process of aging. Menopause is the cessation of periods and occurs during the climacteric. The interplay of several hormones is responsible for the cyclical bleeding that marks the reproductive life of a woman. Hormonal changes begin to occur around the age of thirty and ovarian function diminishes gradually ending some years after menopause[9]. The perimenopausal period is a biological phenomenon which cannot be predicted by chronological age only and extends from just before menopause to several years afterwards[10]. In North America, the average age of menopause is 51.4 years[11].

Estradiol is the predominant endogenous hormone produced by the ovaries in pre-menopausal women. As a woman ages, estradiol levels diminish with a sharp decline during the perimenopause. This decline is partially offset by the synthesis of a less active estrogen called estrone. Estrone is synthesized by the enzyme aromatase by converting androstenedione in adipose tissue and this enzyme is stimulated by alcohol. Among obese women, especially who partake moderate amounts of alcohol have higher levels of

circulating estrogen in the peri-menopausal period than the women who are thinner and abstain from alcohol[12].

Surgical menopause is the term used when both ovaries are removed surgically resulting in an abrupt decline in endogenous estradiol levels.

2.2 Coronary Artery Disease: Definition

Cardiovascular disease (CVD) refers to a broad category of circulatory conditions including coronary artery disease, cerebrovascular disease and peripheral vascular disease (PVD). The 10th revision of the International Classification of Diseases (ICD) from the World Health Organization (WHO) has classified the morbidity and mortality due to CAD under the codes 390 to 458[13]. The term CAD is used interchangeably with terms such as ischemic heart disease (IHD) or coronary heart disease (CHD). CAD is defined as an impaired function of the heart muscle usually due to ischemia or reduced blood flow. The most common cause of reduced blood flow is atherosclerosis of the coronary vessels. CAD may present symptomatically as angina (stable or unstable (UA), myocardial infarction (MI), congestive heart failure (CHF) or sudden death.

2.3 Cardiovascular disease in menopausal women

CVD is the leading cause of death in both men and women in Canada and most of industrialized world accounting for 36% of overall mortality of which 21% is due to CAD alone[14]. Yet, it has been perceived as primarily afflicting middle aged men. There is a relative paucity of studies on prevalence, presentation, diagnosis and prognosis of cardiovascular diseases in women. A study performed in 1997 found 60% of women in all age groups felt that cancer should be their primary health concern and only 20% reported that a physician had ever advised them about heart disease [15]. More recently, Mosca et al[16] found that the proportion of women who felt heart disease was the primary killer in women increased from 30% to 60% in three telephone surveys conducted in the years 1997, 2000 and 2003. It is well established that the rate of cardiovascular events begins to rise in women around menopause and sharply increases with increasing age.

Early publications from the Framingham cohort[17] demonstrated that initial manifestations of coronary events are delayed on average by 10 years in women as

compared to men. The incidence of the more serious outcomes, for example, MI and sudden death, lags by 20 years in women. This gap in coronary events starts to narrow around 65 years of age with the incidence of CAD doubling in men and tripling in women with respect to that in the 35-65 years age group[18]. Among women of the same age, the incidence of CAD in postmenopausal women is three times higher than those who remain premenopausal [19]. This rise has been attributed to the diminishing presence of estrogen and is the principal rationale for evaluating exogenous estrogens as a form of therapy to decrease the prevalence of CAD.

2.3.1 Clinical presentation of CAD in women

The most common presenting symptom of CAD in women is angina which accounts for more than 50% of the initial clinical events[19]. In contrast, 43% of men presented with acute MI as their chief clinical presentation without any forewarning angina. Angina in women often presents atypically, being reported as a burning sensation or abdominal fullness rather than the 'typical' central chest pain reported by men. Thus, women have not been targeted for risk factor modifications of CAD or for diagnostic procedures to evaluate CAD. Women who present with MI as the initial clinical event are more likely to die of it than men (31.8% versus 23.1% respectively). This has been attributed partly to the older age of women at presentation and partly due to higher prevalence of comorbidities (such as unstable angina, hypertension and diabetes) [20].

CAD may present as congestive heart failure (CHF) or sudden death. CHF due to myocardial ischemia is increasing in both sexes but the increase may be more marked among women particularly those with diabetes[21]. Presentation of CAD as sudden death is less common in women than men and this difference has been partly attributed to tobacco use and partly to gender[22]. Women admitted to the hospital for MI do worse than their male counterparts. The average duration of hospitalization for CVD among Canadian women tends to be considerably longer than in men (26.1 days versus 17.1 days respectively) [17;23]and mortality is higher as well. Following discharge from hospital, re-infarction within a year is higher in women, 40% as compared to 14% in men[24]; recurrent angina and congestive heart failure are also more frequent. Studies have shown

that women are less likely to undergo cardiac bypass or angioplasty than men but when they do, they tend to have poorer outcomes[25;26].

2.3.2 Risk Factors for CAD in Women

The risk factors for CAD in women are not all unique to gender, but there exists a difference in primacy and potency of the traditional risk factors from those in men. The correlation of the three major risk factors (smoking, hypertension, hypercholesteremia) with CAD are similar for both sexes. However, diabetes, circulating levels of lipoproteins, body fat distribution and obesity, gonadal hormone status, and family history of CAD appear to have a different connotation for women when compared to men[27]. Recently identified risk factors include levels of C-reactive protein (CRP), homocysteine and lipoprotein(a). These factors have not been studied extensively and it is unclear if the risks are different for men and women.

Some of these risk factors are non-modifiable, for example age, gender, and family history while the remaining can be modified or treated. The INTERHEART study [28] evaluated the effects of nine modifiable risk factors on acute MI in a standardized case control design with enrollment from 52 countries. The nine modifiable risk factors included smoking, hypertension, diabetes, abnormal lipids including apolipoprotein levels, waist/hip ratio, dietary patterns, physical activity, alcohol consumption, and psychosocial factors. The study found that collectively, these nine factors accounted for 90% of the population attributable risk (PAR) for acute MI in men and 94% in women. These modifiable risk factors and contribution of menopause as a risk factor are discussed in the following section.

Smoking: Cigarette smoking is the most important known preventable risk factor for CAD in women and is responsible for 21% of the mortality due to CVD[29]. Smoking is associated with an increase in low-density lipo-protein cholesterol (LDL-C), free fatty acids, platelet aggregation, and insulin resistance. Although the prevalence of women smokers has always been lower than men this gender gap has now narrowed from 17% in 1965 to 5% in 1998[30]. Another important difference between the genders is the percentage of women who quit smoking has always been lower than men (46.1% vs

50.9% in 1998) [30] The adverse effect of smoking is dose-related[31]and works synergistically with other risk factors[32;33]. Although the cardiovascular risk increases in both sexes, the incidence of MI increases six folds in women and only three folds in men who smoked over 20 cigarettes a day compared to those who have never smoked. Additionally, women who smoke and use oral contraceptive pills have a higher risk of venous thrombosis and cerebro-vascular complications[34]. Smoking appears to eliminate the protective effect of estrogen on the cardiovascular system, and is associated with an early menopause. Cessation of smoking for one year decreases the risk of MI by 50% and after 3 years the risk becomes comparable to non-smokers[35].

Hypertension (HT): The risk of acute coronary events and CAD mortality increases with progressively higher levels of systolic blood pressure (SBP) and diastolic blood pressure (DBP). This has been repeatedly documented in prospective studies in both women and men[36;37]. Control of blood pressure using life style modifications and/or drug therapy decreases the rates of CAD in both genders. The prevalence of hypertension increases with age from 21% in women aged 35-44 years to 50% in women between 65-74 years[38]. Due to a high prevalence of this condition the attributable risk of hypertension for CAD is high. Hypertension is as frequent in women as men till the age of 65, after which it is higher in women[39]. There are certain causes of HT (eclampsia, hypertension due to oral contraceptives and primary pulmonary hypertension) which are unique to women. The impact of HT on CAD is stronger in women than in men[39]. It has been reported that lowering diastolic blood pressure is associated with a reduction in acute coronary events in women even when the diastolic blood pressure was in the normal range[39].

Lipid Profile: A lipid profile with elevated levels of total cholesterol and low-density lipoproteins (LDLs) and decreased levels of high-density lipoproteins (HDLs) are established risk factors for CAD in women and men. An increase in total cholesterol or LDL-C of 1% or a decrease in HDL-C of 1% is associated with an increased risk of CAD of 2 to 4.7%[19;40]. Hypertriglyceredemia is an independent risk factor for CAD and is positively correlated with diabetes, oral contraceptive use, obesity and older age. Menopause is associated with an adverse change in the lipid profile. Within 6 months of

menopause, there is a 5 to10% increase in each of LDL-C, total cholesterol and triglycerides (TGs) and a gradual reduction in HDL-C levels over two years[41;42]. Sixty five percent of Canadian women between the ages of 45-64 years have serum total cholesterol values above the desired levels of 5.2 mmol/L[43]. Secondary prevention trials of lipid-lowering agents have consistently shown a significant benefit of treatment in both genders resulting in a 35% to 46% reduction in the relative risk of CAD[44-46]. Based on these observations, the recent National Cholesterol Education Program (NCEP) Guidelines[47]emphasized stringent control of lipid levels.

Obesity: Excess body weight, both overweight (Body mass index (BMI) between 25 and 29.9 kg/m^2) and obesity (BMI $\geq$ 30 kg/m^2), is considered a major independent risk factor for CAD in both men and women[48]. An increased waist circumference (>102 cm in men and >88 cm in women), is an even better prognostic marker for CAD than weight or BMI alone[48]. Obesity also contributes to an increase in the prevalence of hypertension, insulin resistance, dyslipidemia, metabolic syndrome and type 2 diabetes mellitus, all of which increase the risk of CAD. The age-adjusted prevalence of obesity in North American women was 33% in the year 2000[49]. The increase in coronary risk with weight gain is more dramatic in women than men. After the age of 21 years, a weight gain of 15 kilograms was associated with 83% increased coronary risk in women as compared to 46% in men[50]. Besides body weight, body fat distribution is an important determinant of metabolic and cardiovascular disease. Patients with abdominal or 'apple' type obesity are at a greater risk of developing CAD than those who have a 'pear' type obesity with greater fat around the hips[51].

Diabetes Mellitus (DM): Diabetes is an independent risk factor for CAD. Approximately 60% of patients who have diabetes will die from CVD and women appear to be at a higher risk than men[52;53]. The relative risk of developing CAD in people with diabetes varies from 3 to 7 in women versus 2 to 3 in men. Glycosylated hemoglobin A1c (HbA1c), a reflection of blood sugar levels over the previous 90 days, has been studied in relation to risk of cardiovascular diseases in a prospective cohort study. The authors reported that a 1% rise in HbA1c, increased the relative risk of cardiovascular diseases by

24% to 28% in both men and women[54]. This was independent of all other predictors of cardiovascular disease. The prevalence of DM is increasing and was 7.9% in the year 2001[55]in North America. Aggressive management of dyslipidemia in diabetics has been associated with significant reduction in CAD risk[56].

Metabolic syndrome is a constellation of conditions that increase the risk of CAD. There is no consensus on the precise definition of this condition. Patients with insulin resistance, abdominal obesity, dyslipidemia, and hypertension are considered to have metabolic syndrome[57]. This occurs in both genders affecting 44% of people over 50 years of age. Whether patients with metabolic syndrome are at a higher risk than what is accounted for by the individual risk factors is a moot question[58] as aggressive management of each of the above mentioned risk factors in such patients will decrease the risk of CAD. Polycystic ovary disease, pregnancy, and pre-eclampsia are conditions unique to women and all these conditions predispose to the development of metabolic syndrome and subsequently increased risk of CAD.

Sedentary life style and dietary patterns: Physical activity attenuates the impact of several risk factors of CAD including obesity, hypertension and dyslipidemia. Some cross-sectional studies have documented lower weight, lower blood pressure and improved lipid profile in women who had higher levels of physical activity[59]. In a recent prospective study[60] women who underwent cardiac catheterization for cardiac ischemia and reported better physical fitness scores had fewer cardiovascular risk factors and a decreased risk of cardiovascular events. The INTERHEART study[28] found decreased odds of acute MI in people who exercised and this decrease was more marked for women than men (odds ratio (OR): 0.48 versus 0.77 respectively). This study also reported that consumption of moderate levels of alcohol had significantly greater reduction in odds of developing acute MI in women in comparison to men. Consumption of vegetables and fruits was associated with decreased odds of acute MI in both genders in this study.

Socioeconomic status: Lower socioeconomic status has consistently been associated with increased cardiovascular mortality and morbidity, particularly in women[27]. Women who take HRT tend to belong to a higher socioeconomic class and are at a lower risk of

developing CAD. This has been reported in several non-randomized studies evaluating the association between HRT and risk of CAD. Also women with higher incomes and belonging to higher social class are more likely to exercise and have an overall healthier life style decreasing their risk of CAD. Such women are more likely to participate in research studies as well. These have been described as the 'healthy user bias' and 'compliance bias', both of which may explain some of the cardio-protective effect attributed to HRT in non-randomized studies. These biases will be explored in greater detail in this review.

Menopause: The etiological basis of the association between menopause and an increased risk of CAD remains elusive. Endogenous estrogen levels do not correlate with the severity of atherosclerosis or coronary events[61]. A study from California evaluated the age at menopause and mortality from CAD and documented a U shaped relationship with a higher mortality in women who attained early (35-40 years) or late (55 years) menopause[62]. The physiology of this association has not been explained.

Following menopause women have higher hematocrits, blood viscosity and fibrinogen levels. Several inflammatory markers, including CRP, homocysteine, interleukin-6 (IL-6), lipoprotein(a) and their association with endogenous and exogenous hormones is being explored to understand the complex relation between menopause and CAD. Parity, age at menarche and age at first childbirth do not seem to have a significant correlation with developing CAD[63].

Emerging risk factors: Recent breakthroughs in the pathophysiology of CAD have identified additional risk factors which may predict the risk of developing CAD. Levels of CRP have been shown to predict cardiovascular events in healthy population as well as those who have pre-existing heart disease [64]. A specific type of CRP called the high sensitivity CRP (hs-CRP) seems to add predictive value above that of currently well established risk factors [65]. Further research on its role in accurate prediction of CAD needs to be done. The American Heart Association suggests it is optional to investigate for hs-CRP levels in individuals with moderate risk of CAD. Two other factors being studied are homocysteine and lipoprotein(a). There is inconsistent evidence that higher

levels of homocysteine are associated with increased risk of CAD. Similarly, epidemiologic evidence has shown increased levels of lipoprotein (a) are independently correlated with risk of CAD. Further research is needed to clarify the roles of these new risk factors.

2.4 Hormones and their effect on cardiovascular system

2.4.1 Effects of oral estrogen on the cardiovascular system.

Lipid profile: The overall effect of exogenous estrogen is to reduce cholesterol accumulation in peripheral tissues and increase its biliary secretion. Specifically, low dose exogenous estrogen decreases LDL-C, lipoprotein (a) and increases the levels of HDL-C[4].Exogenous estrogens elevate the levels of triglycerides and it is mainly the high levels of very low density lipoprotein cholesterol (VLDL-C) which have an impact on progression of atherosclerosis. These effects of estrogen are mostly mediated though hepatic expression of specific genes. Estrogen reduces the oxidation of LDL-C and may act as an anti-oxidant[66].

Vascular Effects: The vascular function is a complex balance of the vasodilatory and vasoconstrictory factors in the body. Estrogen has direct and indirect effects on the vasculature[66]. The direct effects are of two types: rapid nongenomic effect and slower, long-term effect mediated through genetic expression. The rapid effect includes arterial vasodilatation mediated through nitric oxide synthetase in the endothelium and by stimulating the calcium-activated potassium channels in the cell membrane of vascular smooth muscle cells. The long-term effects of estrogen are an increase in the gene expression of vasodilatory enzymes such as prostacyclin synthase and nitric oxide synthase and enhancement of endothelial cell growth in response to vascular injury. The vasodilator effects of estrogen is seen in the large proximal epicardial vessels[67] as well as the arterioles of the coronary vasculature[68]. Estrogens inhibit endothelial and smooth muscle cell hyperplasia, collagen and elastin deposition and these in turn delay the process of atherosclerosis[69]. Continuous estrogen treatment also results in a decrease in fibrinogen levels, antithrombin and protein S but increased levels of CRP.

Autonomic factors: Estrogens down regulate the activity of mono-amine oxidase and potentiate the effect of catecholamines[70]. This results in a lowering of systolic and mean blood pressure by decreasing peripheral vascular resistance. About 5% of women develop an idiosyncratic reaction to exogenous estrogens and develop high blood pressure which reverts when the hormone is stopped[71]. Estrogen therapy also reduces plasma renin, angiotensin converting enzyme and endothelin1, all of which result in vasodilatation[72]. It improves overall pancreatic function, decreasing insulin levels and improves insulin resistance[4].

Although the evidence is insufficient to be certain, most cardio-protective effects of estrogen have been attributed to the direct effect of estrogen[73]. The net effect of estrogen is decreased vascular resistance and increased perfusion. It also reverses the hypertrophy of heart wall and septum which occurs after menopause. Several studies suggest the beneficial effects of estrogen in salvaging myocardium from cardiac ischemia and consequences of re-perfusion[35]. Experiments in female cats and rabbits have shown a 50% reduction in cardiac necrosis with a bolus administration of estradiol an hour after occluding the left anterior ascending coronary artery[74;75]. In another experiment, female monkeys were fed an atherogenic diet to induce the formation of plaques. Those given oral estrogen showed a remarkable decrease in the plaque size [76]. Primate data also demonstrated that continuous estrogen therapy keeps the coronary blood vessels healthy but if estrogen replacement was done after a hormone-free interval, it cannot reverse the damage to coronary vessels[77]. Thus, it appears estrogens affect the function of cardiac myocytes, lowering their respiratory and energy demands. This anti-ischemic effect has also been demonstrated in women receiving sub-lingual estradiol by a prolongation of time to ST-depression and total exercise time[78]. A hypo-estrogenic state may be a risk factor for CAD and estrogen replacement could have a cardio-protective role.

2.4.2 Metabolic effects of non-oral therapy

Oral HRT formulations are rapidly absorbed after ingestion into the portal venous system and reach the liver. In the liver, estrogen is converted to estrone, estrone sulphate and other estrogen conjugates, decreasing the bio-availability of estrogen to 2% to10% of

the original dose[79]. Thus, higher oral doses are required to ensure therapeutic levels after the 'first-pass effect of the liver'. HRT formulations administered through alternate routes avoid this hepatic first pass and thus protein synthesis by the liver remains untouched. Most of the research on non-oral therapy refers to transdermal estradiol patches. A recent trial using transdermal 17 beta estradiol and NETA showed beneficial effects on vascular function and CAD risk factors[80].

Lipids: Several studies have confirmed that utilization of transdermal estradiol fails to increase the levels of HDL-C in postmenopausal women[81;82]. The levels of lipoprotein (a), the highly atherogenic variant of LDL, decrease with non-oral therapies[83]. Transdermal preparations reduce the levels of TGs, although the mechanism is not clearly known, whereas oral preparations raise the levels of fasting TGs. A recent meta-analysis of 17 population based cohort studies concluded that increase in TGs is associated with an increase in the risk of CAD. This is more marked in women than men and may be an independent predictor of CAD[84].

Vascular effects: The effects of oral and transdermal estrogen on markers of coagulation showed transdermal estradiol had no effect on CRP levels in contrast with oral therapy [85]. Physiologic studies on primates and humans have shown transdermal estradiol reduces vascular resistance in uterine and carotid arteries[86] but this benefit has not been apparent in studies on treadmill performance in CAD patients[87]. Fibrinogen levels decrease with use of both oral and non-oral estrogen therapy.

Autonomic effects: Transdermal estradiol tends to reduce blood pressure but the clinical significance of this association remains to be seen[88]. The effects of non-oral HRT on carbohydrate metabolism have not been studied extensively but a few small studies of transdermal estradiol show a reduction in fasting plasma insulin levels and an improved insulin sensitivity among users[89]. More recent studies on insulin kinetics and transdermal HRT have failed to confirm this effect[85] emphasizing the inconsistency in available literature.

2.4.3 Effect of Progesterone on the cardiovascular system

Progesterone is added to estrogen replacement regimens to protect against endometrial hyperplasia and cancer[90]. Progestins tend to be antagonistic in their actions to estrogen leading to a concern that the beneficial effects of estrogen would be attenuated by progesterone. The effects tend to vary between different progestin preparations: medroxyprogesterone acetate (MPA) or natural micronized progesterone (MP) are primarily progestational and norethindrone acetate (NETA) and norgestimate (NGM)) are primarily androgenic.

Several trials have evaluated the effect of various regimens of HRT in healthy post-menopausal women[4;91;92] and have shown differing effects on lipid profiles depending on the dose and type of progestin used. The combination of 17 beta estradiol and norgestimate provided the maximum beneficial effect by lowering LDL-C, raising HDL-C and minimizing the rise in TGs [93]. Regimens with MPA showed a small rise in HDL-C in comparison with estrogen used alone and had no benefit on decreasing TGs; this effect was more marked with 10 mg doses rather than 5 mg doses. Regimens with NETA demonstrated a decrease in HDL-C levels from the baseline. Natural progesterone had no effect on lipid profile [4].

Progesterones do not antagonize the estrogenic effects on the structure of vascular walls[94]. However in postmenopausal women, they can induce vasospasm of the proximal coronary vessels in postmenopausal women and in animal studies, they reverse the vasodilatory effect of estrogens [95]. Future research may shed light on the complex interplay between progestins and estrogen combinations and how they complement or negate each others actions.

2.5 Role of HRT in preventing atherosclerosis

Progression of atherosclerosis is the hallmark of CAD and the reversal of this process should reduce the risk of clinical coronary events. Although this process is similar in men and women, the age at which women pass through these stages is later than in men. Around menopause, women appear to be in a relatively early stage of atherosclerosis, providing an opportunity to intervene that is not available in men. The measurement of atherosclerosis has been used as a surrogate marker for clinical coronary

events. Serial coronary angiography is the current gold standard for measuring progression of the atherosclerotic process. Studies have shown a change in luminal dimensions measured by quantitative coronary angiography (QCA) correlates with clinical events[96]. However, the test is invasive, and used generally in symptomatic population detecting the effect of interventions at a later stage of disease. Increased coronary artery calcium (CAC) score is another surrogate marker shown to correlate with subclinical atherosclerosis [97].

In order to assess early progression of atherosclerosis, a non-invasive technique that measures intima-media thickness (IMT) of carotid arteries has emerged as a surrogate measure of clinical coronary artery disease[98;99]. The degree of atherosclerosis parallels that of coronary arteries and the abdominal aorta[100] and is positively associated with CAD in women and men[101;102]. This effect of estrogen on IMT has been looked for in women with established coronary disease (i.e. for secondary prevention) as well in those without any CAD (i.e. for primary prevention). Some of the relevant studies are discussed.

Secondary prevention studies on HRT and progression of atherosclerosis

In order to measure the effects of HRT on CAD, secondary prevention studies target a population at a high risk of recurrent CAD.

A sub-study[103] of the HERS trial with 362 participants (mean age 66.7 yrs) evaluated the effect of HRT on delaying progression of atherosclerosis. Postmenopausal women with established CAD either received oral conjugated equine estrogen (CEE) combined with MPA or a placebo and were followed for an average of 4.1 years. The progression of atherosclerosis was greater in the placebo group but not significantly different from the treatment group.

The Estrogen-Replacement and Atherosclerosis Study (ERA) enrolled 309 postmenopausal women (mean age of 65.8 years) with angiographically documented CAD [104]. Women were randomized to three arms: oral HRT either 0.625 mg of CEE alone, a combination of 0.625 mg CEE with 2.5 mg of MPA or a placebo. The study duration was 3.5 years. Coronary angiograms were performed on all patients at the end of

treatment and there was no difference in the progression of coronary plaques in any group.

The Postmenopausal Hormone Replacement Against Atherosclerosis (PHOREA) study[105] evaluated combination HRT (1 mg of 17 beta estradiol (E2) and 0.025 mg gestodene, a new progesterone) versus placebo in clinically healthy postmenopausal women (40 to 70 yrs of age) with documented intima-media thickening greater than one mm. These women were followed on average for 48 weeks and no difference was found in the IMT in either group. Of note, all women had advanced atherosclerosis with a carotid intima thickening greater than one mm. As well, the duration of trial was not long enough to demonstrate slowing of progression of atherosclerosis. Several previous studies on antiatherosclerosis therapies have shown that at least two years are required to demonstrate a slowing of progression of atherosclerosis [44].

The Women's Estrogen-progestin Lipid-Lowering Hormone Atherosclerosis Regression Trial (WELL-HART)[106] enrolled 226 women with established CAD with an average age of 63.5 years. These women were randomized to one of three groups to receive either daily oral micronized 17 beta estradiol and MPA placebo, daily micronized oral 17 beta estradiol and MPA for 12 sequential days each month, or placebo for both hormones. The primary outcome in this study was the mean change in stenosis of carotid vessels from baseline to follow-up as measured by quantitative coronary (QTC) angiography. Carotid IMT was also measured. Again, no difference on delaying the progression of atherosclerosis in any of the three groups was found after 3.3 years of follow-up.

The Women's Angiographic Vitamin and Estrogen (WAVE)[107] trial was conducted to determine the effect of HRT (CEE with MPA or CEE alone in women who had a hysterectomy) and vitamin E and C supplements. Postmenopausal women whose angiogram demonstrated a 15% to 75% stenosis in one of the coronary arteries were eligible. A total of 423 women with an average age of 65 years were randomized in a 2 by 2 factorial design, alone or in combination for HRT and the vitamins. The progression of CAD was measured by QTC angiography. There was no difference in coronary disease progression in any of the four groups receiving HRT or vitamins after an average follow up of 2.8 years.

In 2004, Husak et al [108] conducted a cross-sectional study of 843 postmenopausal women who had undergone their first cardiac catheterization. The presence and severity of CAD (defined as ≥ one diseased coronary vessel with ≥ 50% stenosis) was studied in association with HRT use. A third of all women were on HRT and 78% of these were using estrogen alone. Multivariate modelling was done to adjust for cardiac risk factors and co-morbidities. This study reported a significant cardioprotective effect among the women who had been using estrogen therapy.

Primary prevention studies on HRT and progression of atherosclerosis

Studies on primary prevention of atherosclerosis on women with no pre-existing heart disease are described in this section.

One randomized trial showed that use of HRT was associated with decrease in subclinical atherosclerosis. The Estrogen in the Prevention of Atherosclerosis Trial (EPAT) [109] assessed the progression of subclinical atherosclerosis in postmenopausal women randomized to either oral 17 beta-estradiol or a placebo. In total, 222 women with an average age of 61 years, no known coronary disease and nonsmokers, were enrolled. There was significantly less progression of carotid IMT in the estrogen group after two years as compared to those on placebo. The results of this trial were different from those of the previously discussed secondary prevention trials and one other primary prevention trial. An important difference was women in this trial were at an average 4 to 6 years younger than all other trials. On comparison with HERS trial, the interval from menopause to randomization was an average of 13 years as compared to 23 years in HERS trial.

DeKleijn et al [110] enrolled 121 peri-menopausal Dutch women and randomized them to three arms: 17 beta estradiol in combination with desogestrel, CEE with desogestrel or placebo. The outcome measures were carotid IMT and end-diastolic lumen diameter. No difference in outcomes was reported after two years of follow up and progression of atherosclerosis was similar in all three groups.

Increased coronary artery calcium (CAC) score has been shown to correlate with subclinical atherosclerosis [97]. A recent meta-analysis of four cohort studies reported a CAC score between 1 and 100 consistently predicted a doubling of the risk of coronary

events, a score between 100 to 400 increased the relative risk to four times and scores higher than 400 had a seven times increased risk of CAD[97]. The mean duration of follow-up was 3.6 years. Another prospective study conducted in 2003 evaluated the association between HRT and coronary artery calcium score[111]. The average age of women in the study was 59 years and the average duration of use was 9 years. Current users of HRT were significantly more likely to have this score below 100, and less likely to have scores over 400 than non-users. The study concluded that current use of HRT resulted in a significant reduction in CAC and was cardioprotective.

Results from studies using surrogate markers of clinical coronary disease are discordant, with some studies reporting a decrease in progression of subclinical atherosclerosis while others reporting a lack of association. The number of studies on secondary prevention are greater than those on primary prevention. A major point emerging from these discordant results is that the approach of using hormone replacement in postmenopausal women to maintain hormone levels as in premenopausal women in an attempt to prevent coronary disease is an over-simplistic one. The pathophysiology of how estrogen and progesterone are affecting the process of atherosclerosis needs further delineation. The improvement in imaging techniques, newer bio-markers are making it easier to track the changes of atherosclerosis more accurately. This coupled with further research should give us more answers to the complex association between menopause and CAD.

2.6 HRT: What's available

Hormone replacement therapy is available in various formulations and delivery modes. Oral administration of HRT has been the most popular route with transdermal a second choice. The commonly available formulations and their delivery modes are discussed in the following section.

2.6.1 Oral preparations

Oral preparations are predominantly available as a combined formulation of estrogen and progesterone. The most common estrogens used in these formulations include: conjugated equine estrogen (CEE), micronized estradiol (17 beta estradiol) or

ethinyl estradiol (EE). One study suggests equipotent doses of these three estrogens would be 1 mg of 17 beta estradiol being equivalent to 0.625 mg of CEE or 5 micrograms of EE per day[112]. Worldwide, 17 beta-estradiol is the most frequently prescribed estrogen except in North America where CEE is prescribed most often[113].

The most common progestins used in these formulations are: medoroxyprogesterone acetate (MPA), norethindrone acetate (NETA) or norgestimate (NGM). More recently HRT regimens have used newer progestins such as gestodene and desogestrel. MPA is structurally related to progesterone while the latter two are 19-nortestosterone derivatives. Micronized progesterone is a natural progestin that is also available as an oral preparation.

These commonly available oral preparations, the trade names and their compositions are summarized in Table 2.1.

Table 2.1 Oral HRT formulations and commonly prescribed doses.

Trade name	Estrogen	Progestin
PREMPRO/Premplus	CEE 0.625 mg/0.45mg/0.3mg	MPA 2.5 mg/5 mg/1.5 mg
Activella	17 beta estradiol 1 mg	NETA, 0.5 mg
Prefest	17 beta estradiol 1 mg	NGM 0.09 mg
Premphase	CEE 0.625 mg daily	MPA 5mg from day 15-28
FemHRT	EE 5 micrograms	NETA 1 mg
Estratest /Syntest DS	Esterified estrogens 1.25 mg	Methyltestosterone 2.5 mg
Estratest HS/Syntest HS	Esterified estrogens 0.625 mg	Methyltestosterone 1.25 mg

CEE: conjugated equine estrogen; MPA: medroxyprogesterone acetate; NETA: norethindrone acetate; EE: ethinyl estradiol

2.6.2 Non-oral therapies

Alternatives routes of delivery of HRT include transdermal, subcutaneous implant, vaginal gel or cream, hormone impregnated pessary or vaginal ring, and intra-uterine device. Most often estrogen is delivered by an alternative route and the progestins are co-administered orally. Recently, there has been development of combined transdermal patches, long-lasting progestin implants and intra-uterine devices for delivery

of progesterone. Transdermal gels and patches have been a popular mode of administration among the alternative routes, especially in Europe. Implant use has been limited as a minor surgical procedure is required. Vaginal pessaries and creams are not popular as they are considered messy to use.

2.6.2.1 Transdermal therapy

The available estrogen and progestin transdermal patches are summarized in Table 2.2[114]. Most estrogen transdermal patches contain 17-beta estradiol in a reservoir of ethanolic gel. An adhesive protective back adheres the patch onto the skin allowing the estrogen to diffuse across a semi-permeable membrane on the lower surface of the patch. The rate of delivery of 17-beta estradiol varies from 25 to 200 micrograms/day.

Table 2.2 Estrogen and Progestin Transdermal Patches

Trade name	Active ingredients	Available strengths (mg/day)	Patch application site	Application schedule
Combipatch	Estradiol and NETA	0.05/0.14 0.05/0.25	Lower abdomen	Twice weekly
Estraderm	Estradiol	0.05/0.1	Trunk	Twice weekly
Alora	Estradiol	0.05/0.075/0.1	Abdomen, hip, buttocks	Twice weekly
Climara	Estradiol	0.025/0.05/0.075/ 0.1	Trunk	Twice weekly
Vivelle	Estradiol	0.037/0.05/ 0.075/ 0.1	Trunk	Twice weekly
Vivelle-dot	Estradiol	0.037/0.05/0.075/ 0.1	Abdomen	Twice weekly
Esclim	Estradiol	0.025/0.037/0.05/ 0.075/0.1	Buttocks, femoral triangle, upper arm	Twice weekly

Weekly replacement of the patches is required. An alternative delivery system is a matrix patch where the hormone is contained in adhesive layer rather than a reservoir. This formulation is thinner and often better tolerated. The dose and the rate of release of estradiol is similar to regular patches. A common side-effect of transdermal patches has been the local irritation at the site of application of the patch.

2.7 CAD and HRT: Systematic reviews and meta-analysis

There have been six previously published meta-analyses conducted in an attempt to determine if use of HRT is harmful, protective or unrelated to the risk of CAD[2;3;115-118].

Five of the six meta-analyses were conducted by pooling the summary estimates from individual studies while Hemminki et al evaluated the impact of HRT on cardiovascular events by combining adverse event data from randomized trials that had compared HRT with either placebo, no therapy or vitamins. In this review, the term pooling has been used to describe quantitative synthesis of the results (i.e. the summary estimates of individual studies).

The results from these meta-analyses on the association between the risk of CAD and use of HRT have been inconsistent. The earliest meta-analysis was published in 1991[3] and included 13 case control studies, three cross-sectional studies, 15 cohort studies and one randomized trial. The quantitative overview of all studies taken together yielded a relative risk (RR) of 0.56 (95% CI: 0.50, 0.61) demonstrating a significant decrease in risk of CAD for women who used HRT. The analysis was also performed separately for each different study design. The hospital based case control studies showed a non-significant trend toward an adverse effect of HRT with a summary risk of 1.33 (95% CI: 0.93, 1.91) while all other study designs showed a statistically significant decrease in risk of CAD with HRT use. Grady et al[2] published another meta-analysis in1992 utilizing the same set of studies except one Polish study was excluded[119] and a study done in 1991 was included[120]. This yielded a pooled RR of 0.65 (95% CI: 0.59, 0.71) demonstrating a decrease in risk of CAD in women who had used estrogen compared to women who had never used estrogen. The pooled RR for fatal CAD was similar (0.63, 95% CI: 0.55, 0.72). The study sought to look for differences between estrogen users versus estrogen plus progestin users but concluded that there were too few studies to draw any meaningful inference. Grodstein et al[116] published a meta-analysis with the same set of studies as included in the previous two meta-analyses. The RRs of individual studies ranged from 0.17 to 4.2 and yielded a summary estimate of 0.64 (95% CI: 0.59, 0.68). All of these meta-analyses included studies which had assessed any cardiovascular disease rather than only coronary artery disease. Also, some of the cohort studies included did not have an internal control group. Grady et al discussed the presence of bias in non-randomized studies but concluded that such a large and consistent reduction in the risk of CAD is unlikely to be all due to bias.

In an innovative approach, Hemminki et al[117] examined the incidence of cardiovascular disease (not CAD alone) from published clinical trials reporting outcomes other than CVD in postmenopausal women on HRT. Their meta-analysis included a total of 22 trials in which the number of women with cardiovascular events in each study were summed and divided by the number of women originally allocated to the groups. Data on cardiovascular events were either provided incidentally or as an adverse event. The odds ratio (OR) for women taking hormones versus not taking hormones was 1.39 (95% CI: 0.48, 4.18) demonstrating an insignificant increase in the odds of cardiovascular events. The results of this meta-analysis were in complete discordance with the results of the non-randomized studies as well as the previous three meta-analyses.

A meta-analysis published in 1998 included 12 case control, 3 cross-sectional and 10 cohort studies[115]. This meta-analysis excluded studies with outcomes other than CAD and cohort studies without an internal control group. On pooling the studies, the authors reported a RR of 0.70 (95% CI: 0.65, 0.75) demonstrating a significantly decreased risk of CAD among users of estrogen in comparison with never users. Subgroup analysis on the risk of CAD in women who were on combination of estrogen and progestin versus never users yielded a RR of 0.66 (95% CI: 0.53, 0.84). The authors discussed the potential sources of bias in non-randomized studies and stressed the need for randomized trials to confirm these benefits.

The most recent meta-analysis by Humphrey et al[118] was published in 2002 to evaluate the role of HRT in the primary prevention of CVD and CAD. This meta-analysis included only studies of good or fair quality based on their own criteria[121] which are not validated. The analysis included 11 case control studies, 9 cohort studies, and one randomized trial. They analyzed the association between HRT use (current, past, ever and never) and the incidence and mortality of CAD alone and that of CVD. The pooled estimate for the RR of CAD mortality and incidence in current users was 0.62 (95% CI : 0.4, 0.9) and 0.80 (95% CI: 0.68, 0.95) demonstrating a 40% decrease in risk of fatal CAD and a 20% decrease in risk of overall CAD. Relative risk for past users and ever users showed an insignificant decrease in risk of CAD. Sensitivity analysis was conducted on several confounding variables. The authors reported that among studies adjusting for socioeconomic status or education, there was no association between current

HRT exposure and CAD with relative risks varying between 0.97 to 1.11. The authors mention that similar results were found when analysis was adjusted for alcohol consumption, exercise or both but the RRs are not stated.

2.7.1 Rationale for the present systematic review

Since the previous meta-analyses were conducted, other non-randomized studies have been published. Also, none of the previously published meta-analyses have included the recent randomized trials addressing the association between HRT and CAD. In our present systematic review, these data have been included. Studies have been grouped according to the exposure classification (current, past and ever use) as well as the outcome measure (fatal, nonfatal or overall CAD) to provide groupings of similar studies.

2.8 Systematic Reviews and Meta-analyses

Readers of medical literature are inundated with new scientific information and interpreting this information is a challenge to health care professionals. The systematic review is one tool which has emerged to help summarize studies in a meaningful way. The definition of a systematic review is "the application of strategies that limit bias in the assembly, critical appraisal, and synthesis of all relevant studies on a specific topic"[122]. If quantitative pooling of the data is also performed, then the review is defined as a meta-analysis. In 1992, an international organization of multi-disciplinary team of health professionals initiated the 'Cochrane Collaboration', with a goal to prepare, maintain and ensure the accessibility of systematic reviews on the effects of health care interventions (www.cochrane.org).

The methodology of a systematic review and meta-analysis have been well described and is comprised of several steps: a comprehensive literature search pertinent to the question of interest; a critical appraisal and synthesis of abstracted data from the individual studies; and finally a statistical pooling of data, if appropriate, to provide a consensus of results. The randomised controlled trial is the 'gold standard' research design in the evaluation of medical interventions and the primary design used in Cochrane reviews. A large, well-designed RCT would yield similar comparison groups except for the intervention. Thus, any differences in outcomes between groups are

attributable to the differences in effect of the intervention. The effect estimates from randomized designs are consistent, that is, they converge on population parameters as sample size increases. If more than one RCT has studied similar populations, interventions, and outcomes, then pooling data together provides a summary estimate of the effect size with narrower confidence intervals.

There are limitations of RCTs; these studies are expensive and often measure clinical efficacy of an intervention. They tend to be limited in size and duration of follow-up. Also, less common adverse effects or those occurring later than the follow-up period of most trials cannot be studied well in an RCT. The generalizability of RCTs is restricted as patients enrolled in trials are often different from the average patient seen in clinical practice. For example, most RCTs exclude the very old, the very young, the very frail or those on multiple therapies. As well, the specialized centres where RCTs are often conducted have different populations from primary care practices where most patients are treated. RCTs may not be a suitable design in certain clinical situations due to ethical considerations. In these contexts, non-randomized studies may provide pertinent information. An initiative called Meta-analysis of Observational Studies (MOOSE) has prepared guidelines for systematic reviews of observational studies[123]. Similarly, the Cochrane Non-Randomized Studies Methods Group (NRSMG) is establishing guidelines for including non-randomized studies in systematic reviews (www.cochrane.dk/nrsmg) .

2.8.1 Systematic reviews for non-randomized studies

A systematic review provides an overview of the existing literature about a specific health question. Therefore, the studies included usually encompass a substantial variety of populations, treatment regimens and outcomes. A meta-analysis can produce a summary estimate with improved power and precision and provide insights on why there are differences across studies. Before statistically combining studies, the differences between studies, or the 'heterogeneity', should be adequately investigated and addressed[124].

A meta-analysis is based on the assumption that each study is providing an unbiased estimate of the effect of intervention and the variability is only a reflection of random variation. The final summary estimate would also be free from bias and would

have greater precision. When meta-analysis is extrapolated to non-randomized studies a concern is that the absence of randomization in these studies will fail to eliminate unknown confounders or effect modifiers. In non-randomized studies the participants are assigned to a group based on their previous experiences or on the outcome of interest. For example, a cohort study follows two groups, one exposed and other non-exposed to the intervention and the outcome in each group is measured. In a case-control study, groups are determined by the outcome that has already occurred in the cases, and controls are selected out of those who are free of that outcome. The exposure is measured in both groups. In both these non-randomized designs the comparative groups may differ on aspects other than the one being measured and the measure of association may be true or may be due to confounding factors. A meta-analysis may then produce a precise but misleading overall estimate of the association being studied. Opinions on performing meta-analysis on non-randomized studies vary from firm opposition[125] to a belief that if well-done these may provide useful information[126;127].

2.8.2 Assessing the quality of studies in a systematic review/meta-analyses

The effect of quality assessment on the results of a meta-analyses has been well studied and inclusion of poor-quality studies may exaggerate the summary estimate by 30 to 50 percent[7;128]. If the results of included studies are flawed, then the conclusions of pooled analysis would also be compromised. Assessing the quality of a study may provide an indication of the validity of the results of that study[127]. Similarly, the results of meta-analyses are more likely to be a valid estimate if they take into account the quality of individual studies included in them.

2.8.2.1 Quality Assessment Tools

Several check-lists and scales have been formulated to assess the internal validity, and sometimes the external validity of studies. Internal validity is the extent to which the results of a study are accurate for the circumstances being studied. External validity is the generalizability or the extent to which the results of a study can be applied to other circumstances. Quality of studies is usually assessed as part of a sensitivity analysis to evaluate the effect of exclusion or inclusion of poor quality studies on the

summary estimate[129]. A scale allows a quantification of quality that can then be used for sensitivity analyses. A check-list provides an objective standard for comparing different studies. A recent comprehensive text on systematic reviews in health care mentions over 25 scales and 9 checklists for quality assessment of randomized trials[130]. One such validated scale is the Jadad's scale and this was used for the quality assessment of RCTs in this review. However, the hazards of quality scoring of clinical trials for meta-analysis have been described by Juni and Egger who concluded that relevant methodological aspects should be assesssed individually rather than using summary scores [131].

There are several scales that have been used for quality assessment of non-randomized studies. The majority of these are scales modified from ones used to assess the quality of randomized trials. Deeks et al[8] published a comprehensive review of quality assessment tools that are available for non-randomized studies and identified 194 potential tools. Based on five internal validity domains and 3 or more core items, they found six of these instruments suitable for use in systematic reviews [132-137]. The Newcastle Ottawa Scale (NOS) was one of these six tools which could be used for assessment of quality of case-control and cohort studies. This was pertinent to the current systematic review, as the majority of studies evaluating the risk of CAD and HRT were non-randomized studies.

2.8.2.2 The Newcastle Ottawa Scale

The NOS was developed specifically for the quality assessment of non-randomized studies (both case-control and cohort). The NOS judges a study on three broad perspectives: the selection of study groups, the comparability of groups (exposed and non-exposed for cohort design; cases and controls for case control design) and the ascertainment of either outcome (in cohort design) or the exposure (in case control design). For each of these components, there are two to four items assessing how a study addresses the inherent biases pertinent to that component. If an item has been appropriately addressed a 'star' is given. For the selection of study groups, the NOS evaluates the definitions of the groups and assesses if the groups are representative of the community. A maximum of four stars can be awarded in this category. The comparability of groups is evaluated on the control of the confounding variables in both study designs.

One star is awarded for the most pertinent factor matched or controlled for between the cases and controls or the exposed/unexposed groups. Another star is awarded for any additional relevant factor matched or controlled for in the study. The final component assesses a study for the ascertainment of exposure in a case control design and outcome in a cohort design, each using three items. For case control studies the reliability of measuring exposure in both groups and the non-response rate are evaluated. For cohort studies an unbiased outcome measurement and the adequacy of follow-up are evaluated.

Using the NOS a study can be awarded a maximum of nine stars. The greater the number of stars, the better the study quality. The NOS is currently undergoing a formal evaluation of its validity and reliability. Results from studies where it has been applied suggest a high inter-rater reliability (intra-class correlation of 0.92 for cohort and 0.82 for case control)[138]. In the same study, the results of the NOS were compared to scores on the Black and Down's scale for assessing criterion validity and a Spearman's correlation coefficient of 0.88 and 0.61 was obtained for cohort and case control studies respectively.

2.8.2.3 Rationale for modifying the original NOS

The NOS scale has been used to evaluate study quality in the fields of HRT and breast cancer[138] and users of this scale have reported on its simplicity and ease of administration. However, concern has been raised regarding the comparability component of this scale. The comparability component may not be sensitive enough to differentiate between studies achieving better comparability than those studies which fail to achieve similar comparison groups. Ensuring the two groups are comparable on all factors other than the one under study is one of the important steps in reducing bias in non-randomized study designs. Studies achieving greater comparability should achieve a higher score on their quality assessment. The original NOS has only two items evaluating the comparability among the two groups. The first question addresses the primary confounding factor while the second gives credit to any other relevant factor adjusted for in the study. Most studies control for more than two factors and this results in many studies scoring the maximum of two stars on this component. One of the objectives of this thesis was an attempt to modify the comparability component of the original NOS to increase its sensitivity.

Chapter 3.0

METHODS

This chapter consists of two main sections. In the first section, all the steps undertaken to conduct the systematic review have been described. These included an extensive literature search, screening for relevant articles, data abstraction, and finally data synthesis and analysis. The second section of methods describes the process used to modify the original Newcastle Ottawa Scale for assessing the quality of non-randomized studies.

3.1 Systematic Review

3.1.1 Literature Search Strategy

We used the search strategy developed by the Cochrane Collaboration[139] and there were no restrictions on language or year of publication. Relevant studies were identified through: a) electronic databases using the Ovid interface, b) trial registries such as Current Controlled Trials and Cochrane Controlled Trials Registry, c) cross-referencing the bibliographies of retrieved articles, d) personal communication with primary authors of published non-randomized or clinical trials and, e) manufacturers of HRT. Unpublished material was eligible for inclusion in this review.

Electronic Search Strategy

Electronic databases that were searched included: Medline (from 1966 to April 2004), Embase (1960 to June 2003), Current Contents (from 1960 to April 2004), Premedline, Cochrane Controlled Trials Register, Cochrane Reviews, Biological Abstracts and Dissertation Abstracts, Pubmed and Biocentral. Searches were conducted under the supervision of two librarians who have a masters degree in library sciences (JM and KC). Three broad concepts were searched as subject terms: 1) hormone replacement therapy, which included estrogen (unconjugated/conjugated) and progestins (natural and synthetic), 2) cardiovascular disease including coronary heart disease, myocardial infarction, angina, hyperlipidemia and arrythmias, and 3) study designs such as

randomized and controlled trials, observational studies, cohorts, and case-controls. These terms (text and Medical Subjects Headings (MeSH)) were combined with Boolean operators and functions such as explode, focus and truncation were used. The search strategies for Medline and the Embase databases strategies are provided in Appendix 1a and 1b respectively. Subject headings were tailored to each bibliographic database and so were the free text terms in the titles and abstracts. An additional search was conducted in Medline using the subject terms and a filter designed to detect previously published systematic reviews in this area.

Updating was done throughout the data collection stage of the project until April 2004 and consisted of periodic re-running of the Medline and Current Contents search strategies with restriction of the results to material added since the previous update.

Hand search Strategy

Journals on women's health, cardiology and general medicine (Table 3.1) were searched from June 2002 to April 2004 and the pertinent abstracts were reviewed manually. Reference lists from recent textbooks, review articles, and reports from primary studies were examined to identify additional studies. Proceedings from key conferences (e.g. American Journal of Epidemiology) were also reviewed.

Table 3.1 A list of hand-searched journals (from June 2002 to April 2004)

American Journal of Cardiology	Controlled Clinical Trials
American Journal of Epidemiology	Epidemiology
American Journal of Obstetrics & Gynecology	International Journal of Epidemiology
Annals of Internal Medicine	International Journal of Fertility
British Medical Journal	Journal of American Medical Association
Canadian Journal of Cardiology	Lancet
Cardiovascular Research	Maturitas
Circulation	New England Journal of Medicine
Climacteric	Obstetrics and Gynecology

3.1.2 Study Eligibility Criteria

A study was eligible for inclusion in this review if the following criteria were met:

1. **Population**: Postmenopausal women with either natural or surgical menopause and with or without CAD.
2. **Intervention or Exposure**: Administration of HRT (any preparation of estrogen alone or as a combination of estrogen/progesterone administered through oral, transdermal or parenteral route).
3. **Outcomes**: Coronary artery disease including fatal and non-fatal myocardial infarction and angina. Studies were excluded if they: a) did not differentiate cardiovascular events between coronary, stroke or peripheral vascular disease; b) used survival as an outcome without specifying the cause of death; or c) used surrogate markers (angiographic, coronary risk factors, laboratory markers) instead of clinical coronary events.
4. **Study Design**: Non-randomized studies (prospective or retrospective cohorts, case-controls and cross-sectional) and randomized controlled trials (RCTs) comparing HRT with placebo or *any* standard care were included. Case-series, ecological studies or expert opinions were excluded.

3.1.3 Selection Process

The selection of articles was conducted in two steps. Initially, the inclusion criteria were applied liberally in order to maximize the number of identified studies and prevent an important study from being excluded. A document was retained if the title, abstract or key words suggested it may contain information on the efficacy and/or safety of HRT on CAD in postmenopausal women. After this initial screen all retained citations comprised the 'potentially relevant' documents. The author reviewed all citations and consulted with an expert in meta-analysis (GW) for any study for which the author had any doubts on eligibility.

In the next stage of screening each article was reviewed for eligibility based on the inclusion and exclusion criteria defined above. An affirmative answer to each of the criteria implied that the article had passed the 'relevance filter' and would be included in the

systematic review. For each excluded study, the reason for rejection was recorded. Multiple publications of the same data or updated articles were assessed and the most recent publication from the same dataset was used, except in cases, where the previous article had published a different outcome of interest. When a study had missing data on any essential parameter, three attempts were made to contact the author. If unsuccessful, the study was excluded from the quantitative pooled analysis.

Once the screening was complete, a final list was prepared of all the articles included in the systematic review.

3.1.4 Data Abstraction

The author abstracted and documented the content of each included study using a standard data abstraction form (Appendix 2 a, b and c) to obtain information on the descriptive and quantitative characteristics of individual studies. Descriptive characteristics included: publication status (year and language of publication, whether published, sources of funding); study design (type of study design, number of centres, sample size); population characteristics (demographics, excluded populations, risk factors); study intervention or exposure (dosage, frequency of administration, route of delivery, treatment duration, compliance, and, any co-interventions); and data on outcome (fatal or non-fatal or overall CAD).

A random selection of 10% of studies was reviewed by an independent reviewer (MG) to verify data abstraction. A calibration exercise was conducted prior to data abstraction to decrease abstraction bias. Consensus was used to resolve any disagreements. The original reports were not masked as there is conflicting evidence regarding the benefit of this practice[140].

3.1.5 Assessment of the Methodological Quality of Included studies

Non-randomized studies were assessed for quality by two independent reviewers (HC, SG) applying the original Newcastle Ottawa Scale (NOS) versions for case-control and cohort studies (Appendix 3 a and b). A detailed description of NOS is provided in Section 2.8.3.3.

RCTs were assessed for quality by the same two reviewers (HC, SG) using the 3-item Jadad scale[141] (Appendix 4). Items assessed on the Jadad scale included randomization, double blinding, and accounting for the dropouts and withdrawals. The range of scores is from zero to five, with a score less than three indicating a poor quality study. Schulz et al have shown a lack of concealment of allocation is associated with an exaggerated treatment effect [142]. Thus, in addition to Jadad, the concealment of each trial's treatment allocation [143] was also evaluated by the same two reviewers (Grade A= adequate; B= unclear; C= inadequate) (Appendix 4).

An overall Cohen's kappa statistic was calculated for each item. Regarding the three components of NOS, it was decided a priori that studies receiving two or more stars in the selection component, two or more stars in the exposure/outcome component, and one or more star in comparability component were better quality studies and a kappa statistic was calculated for the three components separately. An intraclass coefficient was calculated for the overall score to assess the inter-observer agreement for the quality assessment of included studies for each of the scales. Scoring differences were settled by consensus and a final score was obtained.

3.1.6 Data synthesis and analysis

Non-randomized studies: Currently available soft-ware packages for quantitative synthesis of studies (Meta-analyst, Revman, Biostats) are designed for the analysis of randomized trials. These packages do not allow the input of a summary estimate directly but require individual frequencies in each group to be entered and the program calculates the summary estimate. These software packages are unsuitable for non-randomized studies because such studies provide summary estimates adjusted for a number of potential confounders. There is no option in these standard software packages for entering odds ratios or relative risks directly. Quantitative synthesis for non-randomized studies was performed using Microsoft Excel (Appendix 5).

In a meta-analysis, the simplest formulation assumes that the observed treatment effects in each study are normally distributed and the only source of uncertainty results from the sampling of study participants. The weight applied to each study is equal to the inverse of the variance of the estimate for that study. This type of variance can be

characterized as 'within-study' variation, and is a function of the sample size of the study and the variability in the patient responses within the study. This is the premise of a 'fixed effects model' where the between study variation is assumed to be negligible.

On the contrary, if the between study variation is not negligible then a 'random effects model' should be applied for statistical inferences. This takes into account between-study variations. The population in this model is the one in which there are infinitely many possible populations. In our present review, the population under consideration is all postmenopausal women, the exposure is HRT and the results are being pooled separately for the three study designs (RCTs, case control and cohorts). The 'fixed effects model' was used for those studies in which heterogeneity was minimal. For those studies that appeared to have substantial differences in patient characteristics or any other parameter, quantitative synthesis was not performed and the variables contributing to differences between studies were explored.

Summary estimate for non-randomized studies: Standard formulae were applied for the fixed effects model[144]. For each study, the adjusted odds ratioor relative risk and the upper and lower 95% confidence intervals (CI) were entered in a spreadsheet. A log-transformation of the summary estimate was performed (Y_i). The standard error (SE) was calculated based on the 95% confidence interval using the formula:

$$SE = \frac{(UCI - LCI)}{2(z_{\alpha/2})} \qquad (1)$$

where $z_{\alpha/2} = 1.96$ corresponding to a 95% CI.

The weight assigned to the estimate (W_i) was calculated by taking the inverse of the square of the SE:

$$W_i = \frac{1}{(SE)^2} \qquad (2)$$

The overall summary estimate ($\hat{\theta}$) was calculated by dividing the product of the log transformed measure (Y_i) and it's weight (W_i) by the sum of all weights ($\sum W_i$) :

$$\hat{\theta} = \frac{\sum W_i Y_i}{\sum W_i} \qquad (3)$$

The results of the quantitative analysis were displayed in a Forrest Plot using Harvard Graphics 3.0 (Regd). The log transformed values of the summary estimates and their CIs were plotted. A line representing 'no effect' i.e. log (OR) = 1 or log (RR) = 1 was also plotted. Orientation of the plot has values to the left of the 'No effect' line or less than one, demonstrating a favorable effect of HRT on CAD while those on the right or greater than one demonstrate a harmful effect. Studies crossing the 'No effect' line indicate no significant association between HRT and CAD.

Summary estimate for RCTs: The 'Revman' software package[139] was used for pooling data from the RCTs.

For all study designs, the appropriate chi-square (χ^2) values were calculated. For non-randomized studies, these were calculated using the same spreadsheet:

$$\chi^2(overall) = \sum W_i Y_i^2 \qquad (4)$$

$$\chi^2(association) = \sum (W_i Y_i)^2 \approx \chi_1^2 \qquad (5)$$

$$\chi^2(\text{hom}ogeneity) = \sum W_i (Y_i - \hat{\theta})^2 \approx \chi_{k-1}^2 \qquad (6)$$

$$I^2 = 100\%(Q - df)/Q \qquad (7)$$

where Q is the Cochran's heterogeneity statistic

df is the degrees of freedom

An association between HRT and CAD would result in χ^2 (*association*) value

being significant. The χ^2 (*homogeneity*) explores the homogeneity among the studies pooled together. Significance indicates heterogeneity among the studies. Three indicators of statistical heterogeneity were utilized: first, a visual inspection of Forrest plots demonstrating the relative risks or the odds ratios with 95% CIs provided an indication of heterogeneity, second, if the p-value obtained by the chi-square homogeneity was greater than 0.1, it indicated that the studies may not be heterogeneous, and third, an I^2 statistic was calculated as described by Higgins et al [145]. The I^2 statistic describes the percentage of total variation across studies that is due to heterogeneity rather than chance. Negative values of I^2 are equivalent to zero and it's value lies between 0% to 100%. A value of 0% indicates no observed heterogeneity and larger values show increasing heterogeneity. If there was heterogeneity present across the variables (exposure, outcome, or control of confounding variables) the studies were not combined quantitatively but explored further to identify factors contributing to it.

3.1.7 Sensitivity and Subgroup analyses

Sensitivity analysis is defined as an assessment of the robustness of results of the statistical synthesis by estimating and comparing the effects of intervention in different groups of studies[139]. For this review, a sensitivity analysis was performed using the following variables: unpublished versus published literature, study quality, and including only those studies which adjusted for a measure of socioeconomic status. Stable summary estimates using all studies and the various sensitivity analyses would reflect robust results[139] and greater credibility.

In subgroup analysis, statistical analysis is done for particular subgroups of patients across the studies that are of direct interest and not just to assess the robustness of overall results as in a sensitivity analysis. This type of analysis is helpful in delineating differences across specific variables and is considered hypothesis generating. We conducted a subgroup analysis assessing the risk of CAD with the duration of HRT use (less than 3 years or more), estrogen alone versus estrogen and progestin and for dose of estrogen in the HRT preparation. The studies evaluating younger postmenopausal women were pooled in a subgroup analysis to assess the effect of early initiation of HRT on the risk of CAD in cohort studies.

3.1.8 Publication Bias

Publication bias is defined as the tendency of journals to publish only those studies which demonstrate a statistically significant result[146]. There is evidence to suggest that positive studies are more likely to be published [147], are more often published in English[148], are cited more often[149]and are more likely to appear as multiple publications[150]. These studies are more likely to be located easily and included in systematic reviews increasing the likelihood of this bias.

Visual inspection of funnel plots was used to detect this bias. The measures of precision that can be used on the Y-axis include inverse of standard error (SE), variance, or log sample size[151] and number of pateinets. We plotted the treatment effect size (RR or OR) on the X-axis and the number of patients as a measure of precision of a study on the Y-axis. A 'funnel plot' is based on the premise that the precision in the estimation of the treatment effects will increase as the sample size of component studies increases. Thus, effect sizes from smaller studies will scatter widely at the bottom and become narrower as the study size increases. If there were no publication bias the shape would resemble an inverted funnel.

An increase in sample size improves the precision of the treatment effect for RCTs. But in non-randomized studies without the advantage of randomization, there may be several biases which may systematically affect the treatment effect. A large sample size and a small SE may provide precise estimates but these may not be accurate. The biases operating in such studies have to be evaluated before concluding a treatment effect is real. Hence, evaluating publication bias in non-randomized studies using funnel plots is controversial and was not done for this review.

For the RCTs, funnel plots were constructed using Microsoft excel with number of patients plotted on the Y-axis and the RR or OR on the X-axis for all included trials. The shape of the graph was evaluated visually to determine if it resembled an inverted funnel.

3.2 Modifying the Newcastle Ottawa Scale

In order to formally modify the existing Newcastle Ottawa Scale, several steps were taken. The initial steps included a critical appraisal of the original NOS to identify any deficiencies, interviewing individuals who had used the original NOS for their experience with NOS and generating new items. A decision was made a priori that changes will be made to the comparability component of NOS alone. Final steps included establishing the validity and reliability of the modified component of NOS.

3.2.1 Survey of users of NOS

A purposive sample of individuals who had applied the NOS to their areas of research were selected. These authors were identified from the original developers of the scale and by a literature search on Medline and Embase search engines. These groups were contacted and a telephone interview with a set of open-ended questions (Figure 3.1) was administered. The responses were analyzed qualitatively to identify the reasons for poor performance of items in the comparability component.

Figure 3.1 Questionnaire used for gathering information on original NOS.

- Name of the Reviewer
- Type of Study Design/s to which NOS was applied
- Area of health research being addressed
- Their opinion on NOS scale in terms of
  - Clarity
  - Ambiguity in any item
  - Time taken for application
  - Addressing most of the important biases in non-randomized study designs
  - Item redundancy
- Poor performance of items in comparability component
- Any suggestions on improving the NOS

3.2.2 Face and Content Validity

The validation process began by assessing content validity of the new items for the comparability component of the NOS. Content validity refers to the comprehensiveness, or how adequately the sampling of questions reflects the aim of the items. In this context, our aim was to formulate items which would measure the comparability of the two groups in non-randomized studies on all variables other than the one being studied. We felt it was important that the scale be kept concise and simple to apply. Thus, by consensus between the two groups who had formulated the Newcastle Ottawa Scale (one in Newcastle, Australia and one in Ottawa, Canada) it was decided to restrict the number of items to about three for the 'comparability component'. All the new items were to provide some measure of comparability between groups under study. For example, items could address the choice of relevant confounding variables, the methods used to adjust these variables or if the analysis had been conducted appropriately. Each of the items, comments or suggestions would then be debated and final items would be selected to replace the existing two items on the comparability component of the NOS.

3.2.3 Inter-rater reliability

The modified scale was to be applied to all the non-randomized studies selected for the systematic review by two independent reviewers (SG and LW). Inter-rater reliability was evaluated by calculating a kappa statistic for each item on the scale.

3.2.4 Criterion validity

Criterion validity is a formal method of testing validity of a scale. In its simplest interpretation, criterion validity measures how well an instrument correlates with an existing 'gold standard' measuring the same outcome. There is no accepted 'gold standard' for assessing the quality of non-randomized studies. One of the existing and validated tools used to assess the quality of case control and cohort studies is the Black and Down's scale[132]. One reviewer (SG) administered this scale to all the non-randomized publications. In order to determine if the modified NOS produces results which correspond to a well established scale as well or better than the original NOS, the

scores obtained on each of these scales were compared to those achieved on Black and Down's scale. To reduce bias, there was a difference of four weeks between using NOS, and Black and Down's scale. The total stars/points from each scale were summed. The level of association was evaluated by using the Spearman's rank correlation coefficient. Separate coefficients were obtained for case control and cohort studies.

Chapter 4.0

RESULTS

The results are presented in two main sections. The first section is comprised of the results of the systematic review on the association between HRT and risk of CAD in postmenopausal women. The second section presents the results of validity and reliability testing of the modified NOS.

4.1 Results of the systematic review

4.1.1 Summary of search results

A total of 4367 articles were identified with the comprehensive search strategy. Of these citations, only 2049 (46.9%) were unique with the remainder being repeat articles. Figure 4.1.1 provides a summary of the search results and the screening process involved in the selection of the articles for the systematic review. There was a 42% overlap of identified studies among Medline and Embase databases.

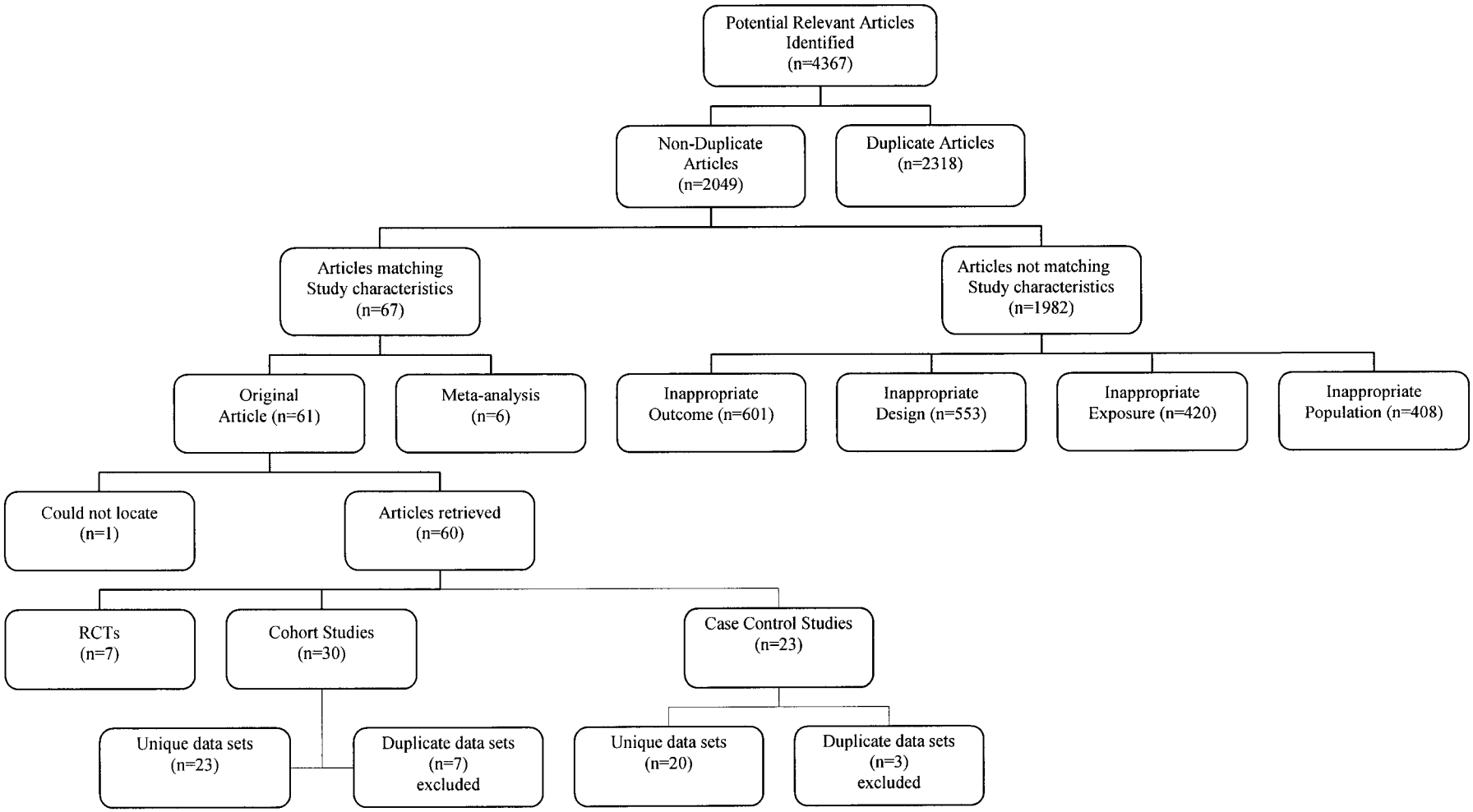
4.1.1.1 Excluded Articles

Of the unique articles, 1982 (96.7%) of the articles did not meet the inclusion criteria. Thirty percent of the articles (n=601) were excluded as the outcome of interest studied was not clinical CAD. Of the 601 excluded studies, 210 articles measured cardiac risk factors as their outcome while 193 measured a surrogate marker for heart disease, including CRP, endothelin and interleukine levels. Another 144 articles assessed heart disease other than CAD as their outcome and 54 articles evaluated all-cause mortality including heart disease but did not analyze results for CAD separately.

Twenty eight percent (n=553) of the articles were excluded as they were either review articles (n=471), case series or expert opinions (n=56) or ecological studies (n=26) with entire geographical areas as their unit of measurement rather than individual patients.

Twenty one percent (n=420) articles were excluded as the exposure measured was not HRT. Of the 420 articles, 206 evaluated the use of either oral contraceptives,

Figure 4.1.1 Flow chart illustrating the selection process of articles included in the systematic review on the association between CAD and HRT in postmenopausal women



selective estrogen receptor modulators (SERMs), or phytoestrogens. Of the remainder, 115 articles studied endogenous hormone levels, 67 studied menopause and its various effects on women's health while 32 articles studied various components of nutrition as the exposure.

Another 21% (n=408) of the articles were excluded because the population studied was not appropriate. Of these, 184 studies had been conducted on animals, 92 were on men and in the rest (n=132), only selected groups of postmenopausal women were included (for example, end stage kidney disease, diabetics, etc).

4.1.1.2 Articles matching the inclusion/exclusion criteria

A total of 67 articles matched the study eligibility characteristics for this review. Of these, 61 were original articles and six were meta-analyses. The six meta-analyses have been described in the background in section 2.7.

Among the remaining 61 articles only one article was in a language other than English and did not have an English abstract. It was identified from the bibliography of one of the previous meta-analysis. This article could not be retrieved despite several attempts. Two unpublished studies (one case-control and one cohort) were identified from dissertation abstracts and one (RCT) from a review article. Among the 57 published studies, there were six randomized trials, 22 case control studies and 29 cohort studies. These studies are discussed in detail in the following sections.

4.1.2 Randomized Controlled trials on HRT and CAD

The search strategy identified a total of seven randomized trials meeting the inclusion/exclusion criteria for the systematic review. Four of these trials examined secondary prevention and three examined primary prevention of CAD. This section summarizes the salient features of the six published and one unpublished trial.

4.1.2.1 Characteristics of RCTs

Table 4.1.1 summarizes the details of RCTs evaluating the effects of HRT on CAD in postmenopausal women. Of note, the WHI trial had two parallel arms using oral

estrogen, one studied the effects of oral estrogen with progestin while the other studied the effects of oral estrogen alone on the risk of CAD.

Secondary Prevention trials

1) Heart Estrogen/Progestin Replacement study (HERS) (1998) [5] : This was the first published randomized trial undertaken with the objective of testing whether HRT in the form of oral estrogen and progesterone reduced the risk of coronary events in women with established heart disease. HERS was a multi-centre, randomized, double-blind, placebo controlled trial. Postmenopausal women with known CAD and an intact uterus were eligible. Women who had a coronary event in the last six months, had severe heart disease (New York Heart Association class IV or Congestive Heart Failure class III), uncontrolled HT or diabetes or a history of deep vein thrombosis (DVT) were excluded. The primary outcome measure was a combined outcome of nonfatal and fatal CAD. The sample size was calculated to have a power of 90% at a two-tailed alpha of 0.05 to detect an intention to treat effect size of 24%. A total of 2763 postmenopausal women with a mean age of 67 years were randomized to either a single tablet containing CEE (0.625mg) and MPA (2.5mg) or a placebo tablet. The duration of the study was 4.1 years and follow up was available for all of the patients.

The intervention group had 172 events and the placebo group had 176 events (relative hazards ratio: 0.99 (95% CI: 0.8, 1.22) with a non-statistical difference between the two groups. Time trend analysis showed a statistically significant increase in coronary events in the first year in the intervention group with an excess risk of 52%. This risk decreased over the next 4 years with a significant trend for benefit (p=0.009). Results from HERS trial were in contrast with previous epidemiological evidence which had suggested a cardio-protective effect in HRT users. There was a 11% decrease in LDLs and a 10% rise in HDL levels in the intervention group (p<0.001) but no decrease in clinical coronary events. Several sub-group analyses were performed to explain these unexpected results.

The proportion of women still on study medications dropped to 75% in the intervention group and 81% in the placebo group by the end of third year. In the 'per protocol' analysis restricted to women who had at least 80% compliance with treatment,

Table 4.1.1 Characteristics of trials done on HRT and CAD in postmenopausal women

Study year	Location	Type of prevention trial	Total number of patients	Average age of women (yrs)	Outcome measured	HRT preparation	Relative risk (95% CI)	Follow up in years
HERS 1998	USA	Secondary	2763	67	CAD events (fatal and nonfatal)	CEE* 0.625 mg and MPA#, 2.5 mg or placebo	0.99 (0.80-1.22)	4.1
Hall et al 1998	Sweden	Secondary	60	59	Symptoms of angina	17β estradiol patch with MPA, placebo or CEE+MPA	Not available	1
PHASE 2002	United Kingdom	Secondary	255	66.3	UA, proven MI or cardiac death	17β estradiol 2.5 mg patch alone or with NETA**	1.29 (0.84-1.95)	4
ESPRIT 2002	England and Wales	Secondary	1017	62.6	Reinfarction or cardiac death	Estradiol valerate 2 mg or placebo	0.99 (0.70-1.41)	2
Natchigall 1979	USA	Primary	84 pairs of women	54	MI	CEE 1.25 mg + MPA	Incidence rate in users: 1.4% incidence rate in nonusers: 4.3%	10
WHI 2002	USA	Primary	16,608	63.3	CAD (fatal and nonfatal)	CEE 0.625 mg and MPA, 2.5 mg or placebo	1.29 (1.02-1.63)	5.2
WHI 2004	USA	Primary	10,739	63.6	CAD (fatal and nonfatal)	CEE 0.625 mg	0.91 (0.75-1.12)	6.8
WISDOM	UK, Australia, New Zealand	Primary	Was to enroll 16,000. Stopped early after 5,000	Not available	CAD (fatal and nonfatal)	CEE alone CEE + MPA	Not available	Not available

*CEE: Conjugated equine estrogen; # MPA: Medroxyprogesterone acetate; ** NETA: Norethi-testosterone acetate

the relative hazards ratio was (0.87, 95% CI: 0.67, 1.11) lower than with the intention to treat analysis but it was not statistically significant. The study concluded the use of oral CEE and MPA did not reduce the overall rate of CAD events in postmenopausal women with established CAD.

The HERS trial was conducted rigorously, the treatment and placebo groups were similar and assessed equally, and outcome was ascertained in all participants. As a 'null trial', having adequate power to detect a difference was an important requirement. The sample size calculations were done to obtain a 90% power for a two-tailed test to detect a 24% difference between treatment and placebo. The event rate and compliance were lower than expected reducing the power of the study. This was partially offset by an 18% over-recruitment of women. Following are a few limitations: The population of postmenopausal women included in the study was relatively old with an average of 18 years after menopause occurred. More than 50% of women were over-weight. The study only tested a single preparation of oral HRT but the one that was used most commonly in North America.

The study concluded that 65 year old women with heart disease should not be started on HRT, specifically, a combination pill containing CEE + MPA, for the prevention of recurrent CAD. The trial did not answer the question of whether a 50 years old menopausal woman with no heart disease should be given HRT or not.

2) Hall et al, 1998 [152] **:** In this randomized study, the authors assessed the effects of HRT on symptoms of angina pectoris, quality of life and investigated factors determining compliance in postmenopausal women with CAD. Sixty postmenopausal women aged 44 to 75 years (mean age 59 years) with known CAD participated in the study. This study did not report other inclusion/exclusion criteria, the method of recruitment and randomization. The women were assigned to one of three treatment groups: 1) transdermal 17 beta estradiol at a dose of 50 μg/24 h for 18 days followed by 10 days of combined treatment with oral MPA (5 mg); 2) transdermal placebo; or 3) CEE (0.625 mg) orally for 18 days followed by 10 days of combined treatment with MPA (5 mg). Each of the treatment cycles were administered for a period of one year. Clinical evaluations of the women including a cardiac history and physical examinations were

performed at baseline and after 3, 6 and 12 months of treatment as well as 6 weeks after completion of treatment. The symptoms of angina were evaluated using the Canadian Heart Association (CHA) protocol.

Of the 60 women who participated in the study only 46 women were followed to completion. Baseline characteristics between the three groups were similar and of note, all women were overweight. The patients were divided into two groups: one with mild (CHA I to II) and one with moderate or severe (CHA III or IV) angina pectoris and the change from baseline was recorded. No significant difference existed between the three groups on the symptoms of angina compared at baseline or after 6 and 12 months of treatment although the study did not present the results as relative risk and 95% confidence intervals. From this study, the authors concluded there were no adverse effects of HRT on women with CAD on angina symptoms in any of the three groups.

This trial assessed two routes of HRT, oral and transdermal, for the symptoms of angina. Limitations of the study included: the trial was unblinded and the method of randomization was unclear; the study had a small sample size and of the 60 women who participated, 25% did not complete the study; no summary estimates or confidence intervals were provided; and the population of included women was an average of 11 years post menopause and all were over-weight.

3) Papworth HRT Atherosclerosis Study (PHASE) 2002[153] : Clarke et al 2002 conducted an unblinded, randomized trial to evaluate the effect of transdermal HRT in postmenopausal women who had angiographically proven CAD. Women were excluded if they had carcinoma of the breast or genital tract, family history of breast cancer, undiagnosed postmenopausal bleeding, cholelithiasis, untreated hypertension, or previous abnormal cervical smear. The primary outcome measure was a combined outcome of admission to hospital with unstable angina, proven MI or cardiac death. Outcome assessment was blinded. Women with a hysterectomy were randomly allocated to an estrogen patch or no therapy in a 2:1 ratio. Women with a uterus were randomly allocated to a combined estrogen/progestin patch or no therapy in a 1:1 ratio. The patch was administered every 4 days to women who had undergone hysterectomy, and for women with intact uterus, it was given until the 14th day and followed by 4 patches of a

combination of 17-beta estradiol (3 mg) and norethisterone (4 mg). A total of 255 (mean age 66.3 years) women were randomly assigned to either intervention group or no therapy. After a follow up of 3 years, an a priori planned interim analysis was performed. Using an intention to treat analysis, the primary end point rate was 15.4 events per 100 patient years for the HRT group compared to 11.9 for the control group (relative risk 1.29 (95% CI: 0.84, 1.95)). The most frequently observed outcome was admission for unstable angina and this occurred most often in the first two years after randomization.

A total of 53 (40%) women withdrew from the intervention group and 8 (7%) from the placebo group. The most common reason for stopping HRT was vaginal bleeding. Using a 'per protocol' analysis, the relative risk increased to 1.49 (95% CI: 0.93, 2.36). Based on these results, the monitoring committee suggested an early closure of the trial. Interestingly, this trial showed a higher event rate in the HRT group in the first 2 years. The study concluded that in postmenopausal women with angiographically proven CAD, treatment with transdermal HRT does not reduce the incidence of acute coronary events.

This was the only large trial to assess the effect of transdermal route of HRT on CAD. The study was designed for 90% power to detect a 18% reduction in intervention group but was stopped early on the basis of interim results. There were several limitations: first, the early stopping of the trial coupled with a high drop out rate in the intervention group decreased the power of the study, second, the trial was unblinded but both groups were followed equally and outcome assessment was blinded. Third, the population of women in this study was older and had been menopausal for over 17 years.

4) Estrogen in the Prevention of Re-Infarction Trial (ESPRIT), 2002[154] : This was a randomized, double-blinded, placebo-controlled trial evaluating the effect of HRT in postmenopausal women with known CAD. All women between the ages of 50 to 69 years admitted to hospital with a first time diagnosis of MI between 1996 and 2000 and had survived were eligible for inclusion. Exclusion criteria included: the use of HRT 12 months prior to the index MI, vaginal bleeding, history of breast or genital cancer, liver disease, severe renal disease or venous thromboembolism. A total of 1017 women (mean age 62.6 years) were randomized to receive oral estradiol valerate (2 mg) or placebo for 2

years. The primary outcomes were cardiac death, re-infarction and all-cause mortality and these were determined by a blinded assessor. Follow up was complete for all women. The study had a power of only 56% with a two-sided test at 5% significance to detect a difference in the rate of nonfatal reinfarction or cardiac death in the two groups. A total of 123 women had an ischemic cardiac event, 62 in the intervention group and 61 in the placebo group. In the intention to treat analysis, the frequency of re-infarction or cardiac death did not differ significantly between the two groups at 2 years (relative risk 0.99; 95% CI 0.7, 1.41). The analysis by time showed the lowest cardiac mortality in the intervention group at 3 months post-recruitment and no evidence of early harm.

The study had a high rate of non-compliance especially in the treatment group (57%) as ascertained from the women directly or by their family physicians. The main reason for non-compliance was vaginal bleeding. Repeat analysis was done excluding those women who did not adhere to the protocol and there was no difference in the results. The study concluded that oral estradiol valerate does not reduce the overall risk of recurrence of cardiac events in women who had survived a myocardial infarction.

This was the only trial to evaluate estradiol valerate, a more common preparation used in most of the world other than North America and it was conducted rigorously. There were a few limitations. First, the study was quite under-powered from inception and this was made worse by poor compliance. Second, the follow up period was short and no inferences could be made for the long-term effects of estradiol valerate on CAD. Third, the women were an average of 16 years after menopause.

Primary prevention

1) Nachtigall et al 1979[155]: Nachitall et al conducted a double-blinded, controlled prospective study on the long-term effects of estrogen replacement therapy (ERT) on postmenopausal women. Women who had been menopausal for at least two years by history and had never been exposed to HRT were eligible for the study. Women were excluded if they had acute heart disease, hypertension, any malignancy or had undergone a hysterectomy. A total of 329 eligible women consented to participate. Of these, 84 matched pairs were selected on the basis of age and diagnosis (diabetes, custodial care and arteriosclerosis) and 130 women were excluded as there was no suitable match

available. Within each pair, a research nurse randomly assigned one member to treatment and the other to placebo. The treatment group received oral premarin, 2.5 mg daily and MPA, 10 mg daily for 7 days each month and the control group received a placebo. The duration of the study was 10 years and all patients were hospitalized for the entire duration of the study. Several outcomes were measured, including overall mortality, all malignancies, and myocardial infarction. The physician's diagnosis was considered reliable and details on each outcome ascertainment were not provided in the study. The study states that a serious attempt was made to prevent the research physician from knowing whether a woman belonged to treatment or placebo group. Complete follow up was available for all women. The women underwent a yearly examination, careful history and a chart review to identify any incidents or complications since the last exam.

In 30 pairs out of the 80 pairs, the blinding code was broken due to the occurrence of a complication, with vaginal bleeding being the most common. One woman among the treated group and three among the placebo group had a MI. The results were presented as the mean incidence in the treated group (1.2%) and the control group (3.7%). The difference between the two groups was not statistically significant. The incidence of MI was well below the national norms for women in this age group in both treatment and control group. The authors concluded that "no protective influence can be ascribed to this dosage of estrogen, but statistically, no increased risk of heart disease was found".

There were several limitations of this study. The sample size was small and the power of the test differences was low. The dose of estrogen was very high and cannot be applied to clinical practice as this dose is not currently used for HRT.

2) The Women's Health Initiative (WHI) : The WHI was started in 1992 as a primary prevention study to assess the major health benefits and risks of the most commonly used combined HRT in United States. Two parallel randomized, double-blinded placebo controlled clinical trials were conducted to evaluate the preventive effects of HRT (combined estrogen plus progestin in women with an intact uterus and estrogen alone in hysterectomized women) on cardiovascular events in mostly healthy postmenopausal women. Postmenopausal women between the ages of 50 to 79 years and likely to reside in the area for 3 years were eligible. Exclusion criteria were based on competing risks

(any medical condition with a predicted survival < 3 years), safety (prior breast cancer, other cancer in last 10 years except nonmelanoma skin cancer, anemia, low platelets), and adherence issues (dementia, alcoholism). Random permuted blocks, stratified by age of participant and centres were used for to achieve randomization. The primary outcome measure was combined fatal and nonfatal heart disease and the principal adverse outcome was development or diagnosis of invasive breast cancer. The effect of HRT on overall health was an important consideration and to measure it objectively, a global index was created. This was defined as the earliest occurrence of CAD, invasive breast cancer, stroke, pulmonary embolism, endometrial cancer, colorectal cancer, hip fracture and death due to other causes.

WHI Estrogen plus Progestin arm 2002[6]: The combination HRT arm of the trial recruited 16,608 women with an intact uterus who were randomly assigned to either receive oral CEE (0.625mg) and MPA (2.5 mg) or placebo. The study was scheduled to continue for 8.5 years. However, after 5.2 years of follow-up the data monitoring and safety committee recommended an early closure of this arm. The test statistic for invasive breast cancer had exceeded the stopping boundary and the global index supported that risks exceeded benefits while the effect on CAD was equivocal. Follow up was complete in 96.5% of the participants and all primary analyses used time to event and intention to treat analysis. The hazard ratio for CAD was 1.29 (95% CI: 1.02, 1.63) with 164 events in the intervention group and 122 in the placebo. Using group sequential methods to correct for multiple analyses over time the adjusted 95% CI for CAD was 0.85 to 1.97 making the increased risk statistically non-significant. Time trend analyses showed the increased risk was evident from the beginning and was mainly due to an increase in nonfatal MI. A subgroup of women (n=400) with pre-existing CAD at baseline had a hazards ratio of 1.28 (95% CI: 0.64, 2.56) for recurrent CAD. Analyses for interactions between age, BMI, prior hormone use, smoking, diabetes, aspirin, or statin use failed to show any benefit or harm by HRT with regards to heart disease.

A substantial number of women had stopped the study drugs at some point during follow up (42% in intervention arm and 38% in placebo). Of note, these participants were not lost to follow up. A sensitivity analysis that excluded non-adherent women increased the hazards ratio for CAD to 1.51. Also, 6% of women in the intervention and 11% of

women in the placebo group initiated hormone treatment from their own clinicians. There was a large disparity in blinding and 41% of women in the intervention group and 7% in the placebo group became aware of their treatment, largely due to vaginal bleeding in the intervention group.

WHI was a rigorously conducted trial and was designed to answer the role of HRT in primary prevention of CAD among postmenopausal women. However, among the included women 36% had hypertension, 49% were current or ex-smokers, and 34% were obese, the proportions being similar in both groups. Of all women randomized, 400 women had a prior coronary event. The average age of women was more than a decade after menopause. Only 10% of women were between the ages of 50 to 54 years, and 25% were over 70 years of age. Power was reduced due to markedly low compliance. WHI reported that overall CAD was 29% greater in the intervention group compared with placebo (37 vs 30/10000 person-years), although the absolute increase in events was small. The nominal statistical significance disappeared when the results were adjusted for multiple analyses.

WHI Estrogen alone arm 2004[156] : Another 10,739 women who had undergone hysterectomy were randomized to only oral CEE (0.625 mg) or placebo. The eligibility and exclusion criteria and the outcome variables were the same as those in the combined estrogen and progestin arm except all included women had undergone hysterectomy. This trial was also stopped early in February 2004 at the recommendation of the data safety monitoring committee due to the increase in risk of stroke among estrogen users. The study duration was 6.8 years and only 5.2% women were lost to follow-up. The relative risk for CAD (fatal and nonfatal) was found to be 0.91 (95% CI: 0.75, 1.12) and the confidence intervals increased to 0.72 to 1.15 on adjusting for multiple analyses. There were 441 women with known history of MI or revascularization procedures. In this subgroup, the effect of CEE relative to placebo was insignificant with a hazard's ratio of 1.04 (95% CI: 0.63, 1.71). The study concluded estrogen therapy had no significant effect on the incidence of coronary heart disease. Exploratory analyses by age at enrollment and treatment effect showed that younger postmenopausal women between the ages of 50 to 59 years on CEE had a decreased risk of CAD but the decrease was not statistically significant.

At the end of the study, 53% of women had stopped the study medication. A sensitivity analysis of only compliant women did not change the estimates for CAD. Also, 5.7% of women in CEE group and 9.1% in the placebo group initiated hormone therapy through their own clinician. Unblinding was minimal being less than 2% in either group.

The WHI was a well-designed trial with appropriate methodology and a follow up of 6.8 years. The study was designed to have a power of 81% to detect the difference in the rate of CAD between the two groups but due to early termination, this was reduced to 72%. The study had limited generalizability as it included older postmenopausal women, tested a single oral preparation of HRT, and more than 50% of women were overweight. There was a high discontinuation rate of study interventions but it was similar in both groups. There was also a high rate of cross over from placebo to treatment use.

3) Womens's International Study of Long Duration Oestrogen after Menopause (WISDOM) trial: This randomized, double-blinded, placebo controlled trial was initiated in 1999 to investigate the long term effects of estrogen combined with progesterone and of estrogen alone on the incidence of cardiovascular disease, breast cancer, dementia and osteoporosis. The trial aimed to include 16,000 postmenopausal women from the United Kingdom and 6,000 women from Australia and New Zealand. The intervention was oral CEE with MPA and oral CEE alone. The trial was halted abruptly in November 2002 after WHI published its results from the combined estrogen and progestin arm. The main reason given for halting the trial was the results from this trial would not be available for the next decade and would not change clinical practice in view of the WHI results. No formal results from the 5,000 women already recruited for the study have yet been published.

4.1.2.2 Quality assessment of RCTs

Table 4.1.2 summarizes the quality assessment of RCTs using the Jadad's scale and assessing allocation concealment. On the overall Jadad score, three trials achieved a high score of four or more indicating good quality studies. The three trials with lower

scores did not have appropriate randomization and were not double blinded. The description of withdrawals was provided by most trials. Allocation concealment was adequate in the three trials with high scores and unclear in the remaining three trials. The scoring for the six RCTs was identical by the two independent reviewers.

4.1.2.3 Quantitative Analysis of RCTs

The results of the individual trials are shown as a Forrest plot in Figure 4.1.2. There were four trials on secondary prevention in whom all included women had pre-existing coronary artery disease, and three on primary prevention in whom the majority of included women were free of heart disease.

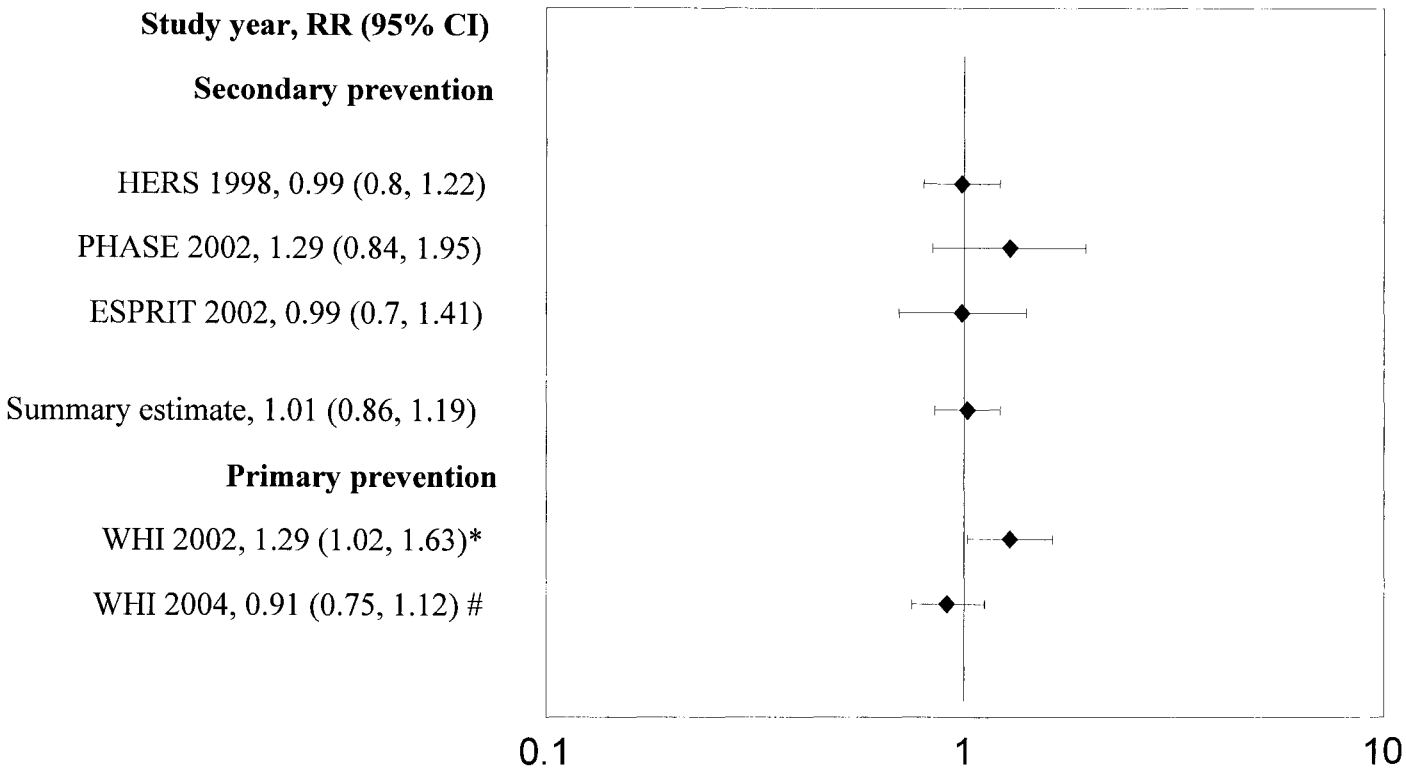
Among the trials on secondary prevention, one trial did not present any summary estimate or confidence intervals for its results. In the remaining three trials, two assessed the effect of oral HRT and one assessed transdermal HRT. Among the two trials using oral HRT, the estrogen preparation used was CEE in one and ethinyl estradiol in the other. The CEE was given in combination with progestin while the other gave it unopposed although the women had an intact uterus. None of the trials found a significant association between HRT and CAD among women with heart disease. Figure 4.1.2 illustrates the results of quantitative pooling of these three trials. A fixed effects model was used for data synthesis and yielded a summary relative risk of 1.01 (95% CI: 0.86, 1.19). The chi-square statistic for association was 0.17 (p=0.90) indicating a lack of association between HRT use and risk of CAD. The p-value for the Cochran Q test for homogeneity was greater than 0.10 demonstrating the studies were not very different and using a fixed effects model was appropriate.

Among the three trials on primary prevention, the study by Nachtigall et al[155] presented the incidence of MI in treated and control groups as percentages but no relative risk was provided. This study used a high dose of unopposed estrogens. The WHI used oral CEE either alone or in combination with MPA and the results are demonstrated in

Table 4.1.2 Quality assessment of RCTs evaluating the association between HRT and CAD in postmenopausal women

Study, year	Randomization	Double-blinding	Withdrawals/dropouts	Total score	Allocation concealment
HERS 1998	2	2	1	5	Adequate
Hall 1998	0	0	1	1	Unclear
PHASE 2002	1	0	1	2	Unclear
ESPRIT 2002	2	1	1	4	Adequate
Natchigall 1979	0	2	1	3	Unclear
WHI 2002 and WHI 2004	2	2	1	5	Adequate

Figure 4.1.2 Risk of CAD in users of HRT versus non-users of HRT among postmenopausal women in RCTs: relative risks and 95% confidence intervals



* After adjusting for multiple analysis, the 95% CI were 0.85, 1.97
\# After adjusting for multiple analysis, the 95% CI were 0.72, 1.15

Figure 4.1.2. The WISDOM trial was not completed and no results have yet been published from it. Thus a quantitative analysis could not be performed. Of note, both the published trials failed to show any association between oral HRT and CAD in relatively healthy women.

Pooling studies on primary and secondary prevention is not warranted as the populations of postmenopausal women under study are different. All six trials failed to show any association between hormone replacement and coronary heart disease.

4.1.3 Case Control Studies on HRT and CAD

4.1.3.1 Characteristics of case control studies

A total of 22 published case control studies and one unpublished study satisfied the inclusion/exclusion criteria for this review. Some of the studies conducted on the same population had been published more than once. Pettiti et al[157] published results in 2000 from the same population as reported by Sidney et al in 1997[158]. The study by Sidney et al was retained for this review as the details of the methods are only presented in this paper and the control of confounders was more complete. Similarly, a report published by Psaty et al[159] was updated after three years by Heckbert et al [160]. In this case the recent report by Heckbert et al was included in this review as the number of patients had been updated. A subgroup analysis on the regimen of HRT used was presented only in Psaty's paper and was included for the pooled subgroup analysis. Jick et al[161;162] published two studies from the same data set, one with any estrogen use (including oral contraceptive pills), while the second study was restricted to estrogen use for noncontraceptive purposes. One study on estrogen use failed to differentiate between HRT use and oral contraceptive use and was excluded. The latter study was retained for this review. The unpublished study was only included in a sensitivity analysis and is described in section 4.1.3.5.

The data extracted from the 19 unique published case control studies[158;160;161;163-178] is summarized in Table 4.1.3. The articles were published between 1976 to 2003 with six studies being conducted in Europe and the remainder in North America. Studies published after 1992 generally had a larger number of cases. In most studies there was

more than one control per case, ranging from two to ten controls per case. In six of the 18 studies, there was a single control for each case. In 13 studies the controls were selected from the general population while in the remaining studies, they were selected from hospitals. In hospital based studies, the controls were women admitted to hospitals for illnesses other than CAD or any other condition contraindicating the use of estrogen or progesterone. In most studies (n=13), it was clearly stated that women with previous heart disease were excluded and hence primary prevention of CAD was being evaluated.

A total of five studies out of the 19 published studies were excluded from the quantitative review. Four were excluded as their study population was comprised of women who were not all menopausal and a separate analysis focused on postmenopausal women was not provided. One study was excluded as it failed to provide the 95% confidence intervals or the standard error for its odds ratio. A brief description of these excluded studies follows. First, Croft et al[168] published a study in 1989 to assess risk factors for acute MI in women. Women between the ages of 20 and older were included and 158 cases of acute MI were identified from the Royal College of General Practitioners oral contraceptive study cohort. Of these, 108 women were over the age of 45 years but there was no information on the menstrual status. An adjusted odds ratio of 0.8 (95% CI: 0.3, 1.8) was reported demonstrating no significant association between use of HRT and risk of acute MI. The results were not reported separately for postmenopausal women. Second, La Vecchia et al[170] reported on a study of Italian women who were all under the age of 55 years. This study was conducted to evaluate the risk factors for MI in younger women. The details of HRT use were not provided. The authors reported a statistically insignificant increase in risk of MI in HRT users. Again, no separate analysis was shown for postmenopausal women. Third, the study reported by Jick et al[161] was a hospital based study restricted to women between the ages of 39 to 45 years. The study had 17 cases and 34 controls of which 13 cases and 21 controls were postmenopausal. They found an increased risk of CAD in users of MI with a RR of 7.5 (95% CI: 2.4, 24) and did not present separate results for postmenopausal women. Of the 14 cases, 13 were smokers. The study stated the presence of risk factors for CAD in cases and controls but did not adjust for any of them. Fourth, Rosenberg et al[174] published a study in 1980 to assess the relationship between noncontraceptive estrogens and nonfatal

Table 4.1.3 Characteristics of the population based case control studies on the association between HRT and CAD in postmenopausal women

Author year	Study Location	No. of cases		No. of controls		Exposure to estrogen	Risk adjusted Odds Ratio (95% CI)	Outcome measured	Variables controlled (by design or analysis)
		HRT	No HRT	HRT	No HRT				
Chilvers # 2003	East midlands, UK	188	371	437	681	Ever Current	0.74 (0.55, 0.99)* 0.53 (0.32, 0.88)*	Nonfatal CAD & fatal CAD separately	Age, DM, HT, smoking, FH, social class, alcohol use, HCBS
Varas-Lorenzo # 2000	England and Wales	133	839	855	3879	Current	0.72 (0.59, 0.89)	Overall CAD	Age, DM, HT, smoking, FH, obesity, HC, type of mp,
Sidney # 1997	California, USA	124	314	150	288	Current	0.96 (0.66, 1.4)	Nonfatal CAD	Age, DM, HT, smoking, FH, BMI, SES, race
Grodstein # 1997	Boston, USA	43	289			Current	0.47 (0.32, 0.69)	Fatal CAD	Age, DM, HT, smoking, FH, BMI, HC, type of and age at mp
Heckbert 1997	Washington, USA	126	621	411	1274	Current	0.74 (0.57, 0.96)	Overall CAD	Age, DM, smoking, previous angina
Mann # 1994	UK	117	1404	562	5522	Current	0.83 (0.66, 1.03)	Overall CAD	Age, DM, smoking, HC, HT, type of mp
Rosenberg # 1993	Boston USA	176	647	176	635	Current	0.90 (0.3, 1.3)	Nonfatal CAD	Age, DM, smoking, HT, BMI, HC, education, FH, alcohol use, exercise

DM: diabetes mellitus; HT: hypertension; FH: family history; HCBS: health conscious behaviour score; HC: high cholesterol; mp: menopause; SES: socioeconomic status; BMI: body mass index

* OR for nonfatal CAD

Studies which excluded women with heart disease

Table 4.1.3 Contd. Characteristics of the population based case control studies on the association between HRT and CAD in postmenopausal women

Author year	Study Location	No. of cases		No. of controls		Exposure to estrogen	Risk adjusted Odds Ratio (95% CI)	Outcome	Variables controlled (by design or analysis)
		HRT	No HRT	HRT	No HRT				
Avila # 1990	Boston, USA	18	85	157	564	Current	0.7 (0.4, 1.4)	Overall CAD	Age, women on antihypertensives, diuretics, andtidiabetics and antiarrythymics
Croft # 1989	UK	9	149	32	126	Ever	0.8 (0.3, 1.8)	Overall CAD	Age, HT, smoking, , SES, OCP use
Beard # 1989	Rochester, USA	19	67	45	105	Ever	0.55 (0.24, 1.3) 0.82 (0.46, 1.47)	Overall CAD & Angina alone	Age, DM, HT, smoking, type of mp
Bain 1981	Eleven states of USA	32 64	56	825 1390	1048	Current Ever	0.7 (0.4, 1.1) 0.8 (0.6, 1.3)	Nonfatal CAD	Age, DM, HT, smoking, FH, Obeisty, HC, type of mp, angina
Ross # 1981	LA, USA	133 cases, no numbers for HRT use provided		133 living controls, 124 deceased controls		Ever	0.43 (0.24, 0.75) for living 0.57 (0.33, 0.99) for deceased	Fatal CAD	Age, DM, smoking, HC, HT, stroke, FH, alcohol use, cardiovascular drugs
Adam # 1981	England and Wales	9	67	26	125	Ever	0.65 (estimated)	Fatal CAD	Age
Pfeffer # 1978	California, USA	17 59	137	61 172	348	Current Ever	0.78 (0.38, 1.63) 0.87 (0.57, 1.38)	Overall CAD & Nonfatal CAD*	Age, DM, HT

DM: diabetes mellitus; HT: hypertension; FH: family history; HC: high cholesterol; mp: menopause
Pfeffer et al also provided adjusted ORs for nonfatal CAD separately: (Current users: 0.68, 95% CI: 0.32, 1.42 and ever users: 0.86, 95% CI: 0.54, 1.37)
Studies which excluded women with heart disease

Table 4.1.3 Contd. Characteristics of hospital based case control studies on the association between HRT and CAD in postmenopausal women

Author year	Study Location	No. of cases		No. of controls		Exposure to estrogen	Risk adjusted Odds Ratio (95% CI)	Outcome	Variables controlled (by design or analysis)
		HRT	No HRT	HRT	No HRT				
La Vecchia 1989	Italy	9	154	7	237	Current	2.95 (0.8, 10.8)	Nonfatal CAD	Age, DM, HT, smoking, FH, social class, HC, alcohol use, age at mp
Szklo # 1984	Baltimore, USA	11	28	14	31	Ever	0.61 (0.2, 1.87)	Nonfatal CAD	Age, DM, HT, smoking, education, type of mp, stroke
Rosenberg # 1980	Boston, USA	27	477	120	1832	Current	1.0 (0.6, 1.7)	Nonfatal CAD	Age, DM, HT, smoking, HC, age and type of mp
Jick 1978 #	Boston, USA	9	8	4	30	Ever	7.5 (2.4, 24)	Nonfatal CAD	type of mp
Rosenberg 1976	Boston, USA	8	328	330	6400	Ever	0.97 (0.48, 1.95)	Nonfatal CAD	Age, DM, smoking, past h/o CAD, type of mp

DM: diabetes mellitus; HT: hypertension; FH: family history; HC: high cholesterol; mp: menopause
Studies which excluded women with heart disease

MI in women younger than 50 years. Half of the included population was pre-menopausal. They found no association between estrogen use and MI but did not provide separate analysis for postmenopausal women.

The fifth study that was excluded was by Adam et al[163] and it was a pilot case control study designed to evaluate the association between cardiovascular disease and HRT. The authors obtained the death certificates of 248 women between the ages 50 to 59 yrs who either had a MI or subarachnoid hemorrhage as the stated diagnosis. Of these, only 97 women could be included due to reasons of incorrect diagnosis (n=33), not traceable (n=25) and no response from the general practitioner (n= 93). No attempt was made to improve the rate of response. The overall use of HRT was only 3% among cases and control group. No details on type, preparation, dose or route of HRT were provided. The unadjusted effect size was not calculated but could be estimated (OR 0.69). The authors mentioned a lack of association between the risk factors of acute MI and use of HRT.

Thus, a total of 14 studies were includ ed for the quantitative review. The details of exposure ascertainment in these 14 studies are presented in Table 4.1.4. The following section describes the method and frequency of ascertainment, the classification of HRT use, and the preparations and routes of HRT in the case control studies. In most publications, ascertainment of exposure to HRT was a one-time assessment. The only exception was the study by Grodstein et al[169] who updated the exposure status every two years by a repeat questionnaire. This was a nested case control study and had the advantage that the exposure had been ascertained before the outcome occurred. For the remainder of the studies, five used pharmacy or medical records to identify prescriptions for HRT, two depended on self-reporting of HRT use and the rest conducted personal or telephone interviews with the subjects. Sidney et al[158] used pictures of the pills to improve recall of HRT use among the study subjects. Compliance with medication cannot be measured without regular updating. Pharmacy records confirm a prescription of HRT but do not provide information on whether the medication was taken and for what duration.

The exposure status was classified as current by ten studies and as ever by the remaining. Four studies presented separate analyses for current and ever use of HRT. All

Table 4.1.4 Ascertainment of hormone replacement therapy in case control studies

Author year	Ascertainment of HRT use	Data on regimen, dose, duration or route of HRT	Definitions of HRT use
Chilvers 2003	Interview using a structured questionnaire	Combined E + P , oral, patches and implants; analysis by duration and regimen	Current use, past use and ever use not defined.
Varas Lorenzo 2000	National Formulary	Oral CEE or CEE + MPA (76%); transdermal (9%), analysis by duration and regimen	Recent/Current: within 6 mo of index date, past use= use stopped 6 mo before index date, nonuse= no prescription of HRT
Sidney 1997	Personal interview	Oral CEE or CEE + MPA (84%); analysis by duration and regimen	Current and past not specifically defined. Lifetime HRT use assessed
Grodstein 1997	Questionnaires every 2 years from 1976 to 1992	Oral HRT, Analysis by duration of HRT for all cause mortality not CAD alone	Current and past use not defined. Patient classified according to the last questionnaire before death
Heckbert 1997	GHC computerized pharmacy database	Oral E or E + P; analysis by duration; analysis by regimen*	Current: prescription to last till index date, ever=2 prescriptions
Mann 1994	Computer records	E or E + P: no details provided	Current use : within 6 months of index date, past use : > 6 months before the index date.
Rosenberg 1993	Personal or telephone interviews	Unopposed CEE or CEE + MPA; analysis by duration and regimen	User ≥ 1mo use, Nonuse < 1 mo use. 2/3rds had taken it for over 1 year.
Avila 1990	Computerized Pharmacy records	Analysis by duration and dose of HRT	Current: using for 12 mo after a prescription, recent user: used 12 to 23 mo after last prescrition, nonusers: never users or stopped >23 mo ago.
Beard 1989	Medical records	Steroidal estrogen preparations listed, premarin most commonly used	No definitions of users/nonusers

* Available from previous publication by Psaty et al

Table 4.1.4 Contd. **Ascertainment of HRT in case control studies**

Author year	Ascertainment of HRT use	Data on regimen, dose, duration or route of HRT	Definitions of HRT use
Szklo 1984	Personal interviews by trained nurses	Estrogen therapy. No data on other parameters	Ever and never use based on interviews
Bain 1981	Self reported on a questionnaire	No data	Current: when the interval between menopause and hospitalization was ≤ duration of use
Ross 1981	Medical records	Oral ERT (CEE), analysis by dose of ERT but no CIs provided.	Use/nonuse not defined
Pfeffer 1978	File of estrogen prescriptions	Oral estrogens (CEE)	User/ever/current not defined. Continuous use ≥ 30 days of ERT
Rosenberg 1976	Personal interviews by trained nurses	CEE most commonly used ERT, duration of use analyzed but no CIs	Ever: regular use defined as used on a scheduled basis. Irregular users were excluded

studies defined current use of HRT as being on hormones at the time of the index coronary event. Some defined current use as a time period a woman had been on HRT before the index event which ranged from a month of use to a year. Of the seven studies that classified their HRT users as ever users, three stated the HRT use had been for over one year and the remaining four studies did not quantify the duration of use. The most common estrogen preparation used was oral conjugated equine estrogen (CEE), followed by ethinyl estradiol and estrone. Most studies do not provide any details on the type of progesterone used and a few mentioned the use of oral MPA.

In the majority of studies the route of HRT was oral. Two recent studies, Chilvers et al[167] and Varas Lorenzo et al[178], analyzed the risk of CAD by two common routes of HRT, oral and transdermal. Both the studies were population based studies and conducted in Europe. Ever use of HRT was reported as 34% and 22% respectively in the two studies. Oral CEE was the most common HRT used. Transdermal patches were used by less than 10% of cases in both studies. Nonfatal MI was the primary outcome in the study by Chilvers et al while Varas Lorenzo analyzed overall CAD (fatal and nonfatal). The adjusted ORs in both studies showed that oral estrogens decreased the risk of CAD. The adjusted ORs for the association between transdermal HRT and risk of CAD were 1.7 (95% CI: 0.58, 4.98) in the study by Chilvers et al and 0.75 (95% CI: 0.47, 1.21) in the study by Varas Lorenzo. The direction of association was opposite in the two studies although neither achieved statistical significance. The number of women on transdermal estrogen was small in both trials. The details on duration, dose and regimen of HRT are analyzed as part of the subgroup analysis.

The outcome of interest in this review was clinical CAD. A total of six studies analyzed the risk of overall CAD (fatal and nonfatal), two analyzed only fatal CAD, four studied only nonfatal CAD, one study presented the results for overall CAD and nonfatal CAD separately while one presented the risks for nonfatal CAD and fatal CAD separately. For quantitative analysis, we grouped the studies according to the outcomes studied.

In order that the cases and controls are comparable on all factors other than HRT use, studies need to consider adjusting for known risk factors for CAD. Most of the case-control studies have listed the potential confounders and have adjusted them either in the

design or the analysis (Table 4.1.3). These potential confounders include age, diabetes, smoking, high blood pressure, family history of CAD, BMI or obesity, high cholesterol, socio-economic status, type of menopause and age at menopause.

4.1.3.2 Quality of case control studies

The NOS was applied to the 19 published case control studies for assessing the quality of these studies. The quality scores of these studies ranged from a low score of three to a maximum score of nine. Six of the 19 studies scored seven or higher suggesting a good quality study. Most of the studies with a low score of five or less were published prior to 1981 with two exceptions, Sidney et al[158] was published in 1997 and had a low score of five while Pfeffer et al[172] published in 1978 achieved a high score of nine. The results of NOS are summarized in Table 4.1.5.

For the first component of selection of cases and controls, 11 studies had an adequate definition of cases with independent validation or record linkage. Most cases were comprised of postmenopausal women with a fatal or nonfatal CAD based on either WHO criteria, ICD 9 or American Heart Association guidelines. Almost half of the studies had cases representative of the general population but in the rest, there were potential for selection biases; for example, women in whom exposure could not be ascertained were excluded making the recruitment non-consecutive and few studies selected hospital controls. These two items contributed to low scores in this category. The last item in selection component assesses if the controls were free of the outcome of interest. The majority of studies stated the controls were women without any heart disease.

For the second component of comparability there was little variation among the scores and 15 out of the 19 studies got the maximum score of two awarded in this component. The most important risk factor for CAD is age and all studies except one controlled for age, either in design or analysis. Similarly, the majority of studies controlled for at least one other confounding variable. In fact, the majority of studies controlled for several confounding variables as summarized in Table 4.1.3. Jick et al [161] was the only study to mention three confounding variables and not controlling for any of them. This study did not receive any stars for this component.

Table 4.1.5 Quality assessment of case control studies using Newcastle Ottawa Scale

NOS criteria	Selection				Comparability		Exposure			
Study year	Case definition adequate	Cases representative	Community controls	Controls: no history of disease	Study controls for age	Study controls for any other risk factor	Secure record of HRT use	Same for cases and controls	Same non-response rate	Total score out of 9
Chilvers 2003	*	*	*	*	*	*	-	*	*	8
Varas Lorenzo 2000	*	*	*	*	*	*	*	*	*	9
Sidney 1997	*	-	-	*	*	*	-	*	-	5
Grodstein 1997	-	-	*	*	*	*	-	*	*	6
Heckbert 1997	*	*	*	*	*	*	*	*	*	9
Mann 1994	-	*	*	*	*	*	*	*	*	8
Rosenberg 1993	*	-	*	*	*	*	-	*	-	6
Avila 1990	*	-	*	*	*	*	*	*	*	8
Beard 1989	*	*	-	*	*	*	-	*	-	6

Table 4.1.5 Contd. Quality assessment of case control studies using Newcastle Ottawa Scale

NOS criteria	Selection				Comparability		Exposure			
Study year	**Case definition adequate**	**Cases representative**	**Community controls**	**Controls: no history of disease**	**Study controls for age**	**Study controls for any other risk factor**	**Secure record of HRT use**	**Same for cases and controls**	**Same non-response rate**	**Total score out of 9**
Croft 1989	-	*	*	*	*	*	-	*	*	7
La Vecchia 1987	*	-	-	*	*	*	-	*	*	6
Szklo 1984	*	*	-	*	*	*	-	*	-	6
Bain 1981	-	-	*	*	*	*	-	*	-	5
Ross 1981	-	-	*	*	*	*	-	*	-	5
Adam 1981	*	-	-	-	*	-	-	*	-	3
Rosenberg 1980	*	-	-	*	*	*	-	*	*	6
Pfeffer 1978	*	*	*	*	*	*	*	*	*	9
Jick 1978	*	-	-	*	-	-	-	*	-	3
Rosenberg 1976	-	*	-	-	*	*	-	*	*	4

For the third component of exposure ascertainment, two items contributed greatly to the variation in the scores. For the ascertainment of exposure only four studies had a secure record or blinded assessment of exposure while in the rest, the exposure was either self reported, or was assessed without any blinding to the status of a case or a control. For the non-response rate among cases and controls, ten studies did not mention this rate resulting in a low score on this item. The higher scoring studies either had no non-responders with complete data on all included women or indicated an equal non-response rate for both the groups. The item assessing if the method of ascertainment of exposure was the same for cases and controls provided less variation with all studies scoring a star.

The inter-rater reliability scores for these studies were calculated in three ways. The Cohen's kappa coefficient was used for agreement on each item and on average, was 0.88 (0.79, 0.96) demonstrating a high degree of agreement between the two raters beyond chance. The Cohen's kappa coefficients for the three components were 0.89 (0.68, 1.0), 0.77 (0.34, 1.0), and 0.67 (0.33, 1.0) for selection, comparability and exposure ascertainment respectively. The agreement on overall score was assessed using intraclass correlational coefficient and this was 0.96 demonstrating a high degree of agreement between raters.

4.1.3.3 Quantitative analysis of case control studies

The qualitative analysis of case control studies revealed that most studies had a comparable population of interest. For quantitative pooling, the studies were grouped according to exposure status (current, past, ever use) and outcome analyzed (overall CAD, only fatal CAD or only nonfatal CAD) in the primary study. As mentioned before, 14 of the 19 studies were included in the quantitative analysis.

Current versus never users of HRT

Ten of the 14 studies analyzed data evaluating the association between current users of HRT and risk of CAD. The outcomes studied were as follows: four studies assessed overall CAD, one assessed fatal CAD, four studies examined nonfatal CAD and one study presented separate analysis for overall CAD and nonfatal CAD.

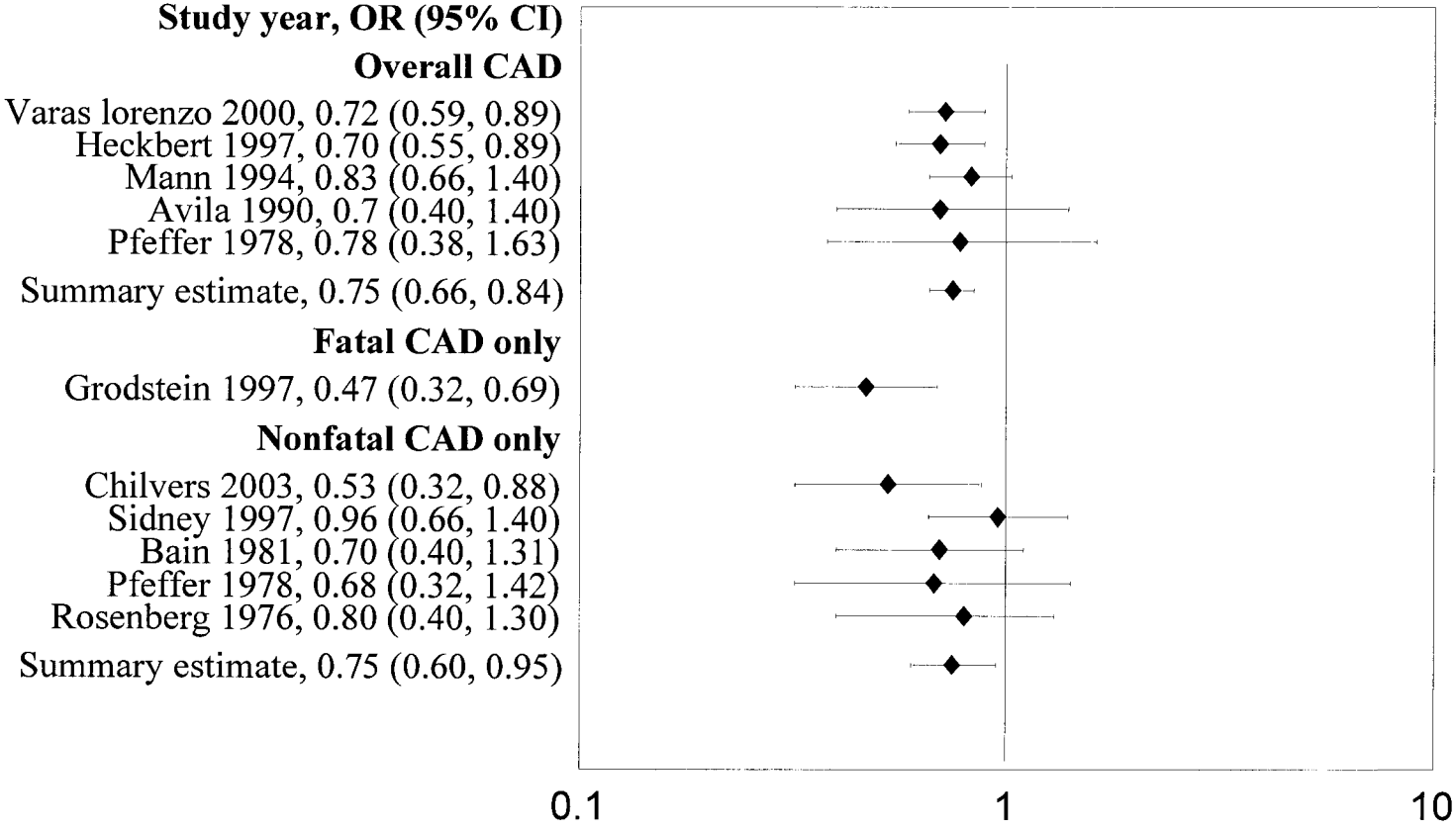
The ORs and 95% CIs from the primary studies are plotted in Figure 4.1.3. Of the five studies assessing the risks of overall CAD, two studies showed a statistically significant inverse association between current HRT use and CAD while the remaining studies did not show a significant association. These estimates were combined using a fixed effects model and yielded a summary estimate of 0.75 (95% CIs: 0.66, 0.84) demonstrating a significant decrease in risk of overall CAD in current users of HRT. Only one study examined the association between current users of HRT and risk of fatal CAD and this study demonstrated a significant inverse assocation (OR: 0.47; 95% CIs: 0.32, 0.69). All five of the studies examining the association between current use of HRT and risk of nonfatal CAD found an inverse association but it was statistically significant in only one study. A quantitative pooling of the results using a fixed-effects model yielded a summary odds ratio of 0.75 (95% CI: 0.6, 0.95) demonstrating significantly decreased risk of nonfatal CAD among current users of HRT.

Table 4.1.6 provides the results of statistical testing of the summary estimates. The results suggest the current use of HRT is protective for all CAD, whether fatal or nonfatal. The Cochran's Q test for the two pooled estimates revealed a p-value greater than 0.10 indicating that the null hypothesis of homogeneity of study results could not be rejected and a fixed effects model was appropriate.

Past versus never users of HRT

A total of eight studies examined the relation between past use of HRT and CAD. All studies defined past users as women who had used HRT in the past and were not using it at the time of their index coronary event for cases or at the time of the study for controls. The outcomes assessed in these studies were as follows: three studies assessed overall CAD, one assessed only fatal CAD and the remaining four assessed only nonfatal CAD. Chilvers et al[167] divided past users into women who stopped less than a year ago, one to four years and greater than five years ago before the index event. They found the risk of nonfatal CAD did not decrease in any of these groups of past users. The study did not provide a single estimate for all past users and its results could not be pooled quantitatively.

Figure 4.1.3 Risk of CAD in current users of HRT versus never users of HRT among postmenopausal women in case control studies: odds ratios and 95% confidence intervals



The ORs and 95% CIs for all the remaining seven studies are plotted in Figure 4.1.4. Of the three studies analyzing the risk of overall CAD, all showed decreased risk with past use of HRT but the decrease was statistically significant in only one study. The pooled estimate for these three studies using a fixed effects model was 0.73 (95% CI: 0.59, 0.89) demonstrating decreased risk of overall CAD among past users. The one study looking at fatal events alone failed to show any association between past use of HRT and risk of fatal CAD. All three studies assessing the risk of nonfatal CAD failed to show any association between past use of HRT and nonfatal CAD. The summary odds ratio, calculated by using a fixed effects model, was 0.94 (95% CI: 0.74, 1.19) demonstrating a lack of association between past users of HRT and risk of nonfatal CAD.

Chi-square tests of association and heterogeneity were performed for the pooled estimates as displayed in Table 4.1.6. These showed a significant association between past users of HRT and risk of overall CAD but no association between past users of HRT and risk of nonfatal CAD. Testing using the Cochran's Q test revealed that for the studies grouped together the null hypotheses of homogeneity could not be rejected and using a fixed effects model was appropriate.

Ever versus never users of HRT

Seven studies evaluating the association between HRT and risk of CAD classified HRT users as ever users if they had used HRT at any time after menopause. The outcome assessed in these studies were as follows: one study analyzed data for overall CAD, one assessed only fatal CAD, three assessed only nonfatal CAD, one study reported separately on overall CAD and nonfatal CAD and one reported separately on fatal and nonfatal CAD.

The ORs and 95% CIs for all included studies are shown in Figure 4.1.5. The two studies on overall CAD found ever users to be at a decreased risk of CAD which was not statistically significant. A pooled estimate using a fixed effects model yielded a summary estimate of 0.78 (95% CI: 0.53, 1.16) with a non-significant protective effect. Both studies evaluating ever use of HRT and risk of fatal CAD found a decreased risk of fatal CAD which was statistically significant. A pooled estimate using a fixed effects model showed a decreased risk in this group with an OR of 0.42 (95% CI: 0.29, 0.58). Of the

Figure 4.1.4 Risk of CAD in past users of HRT versus never users of HRT among postmenopausal women in case control studies: odds ratios and 95% confidence intervals

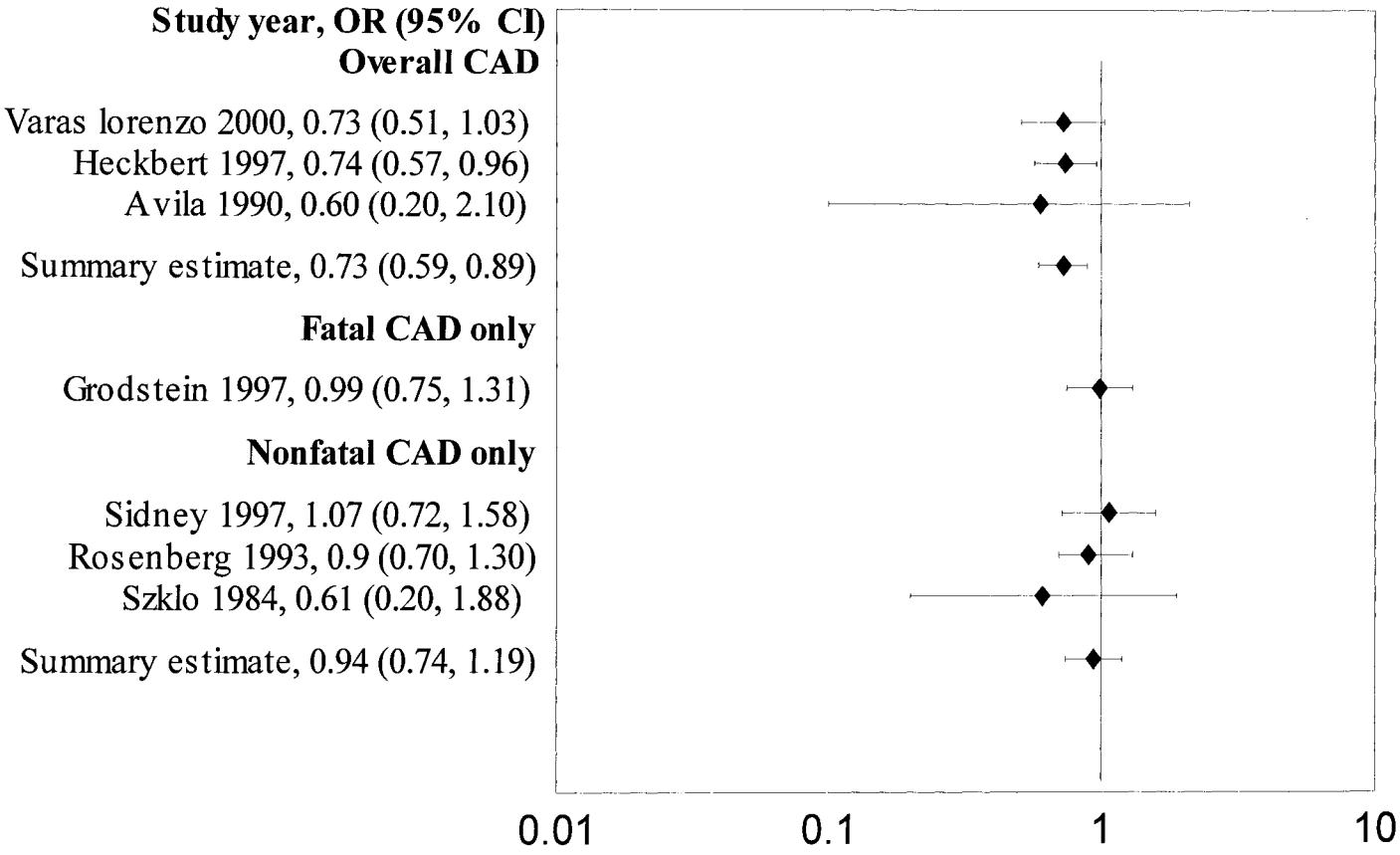
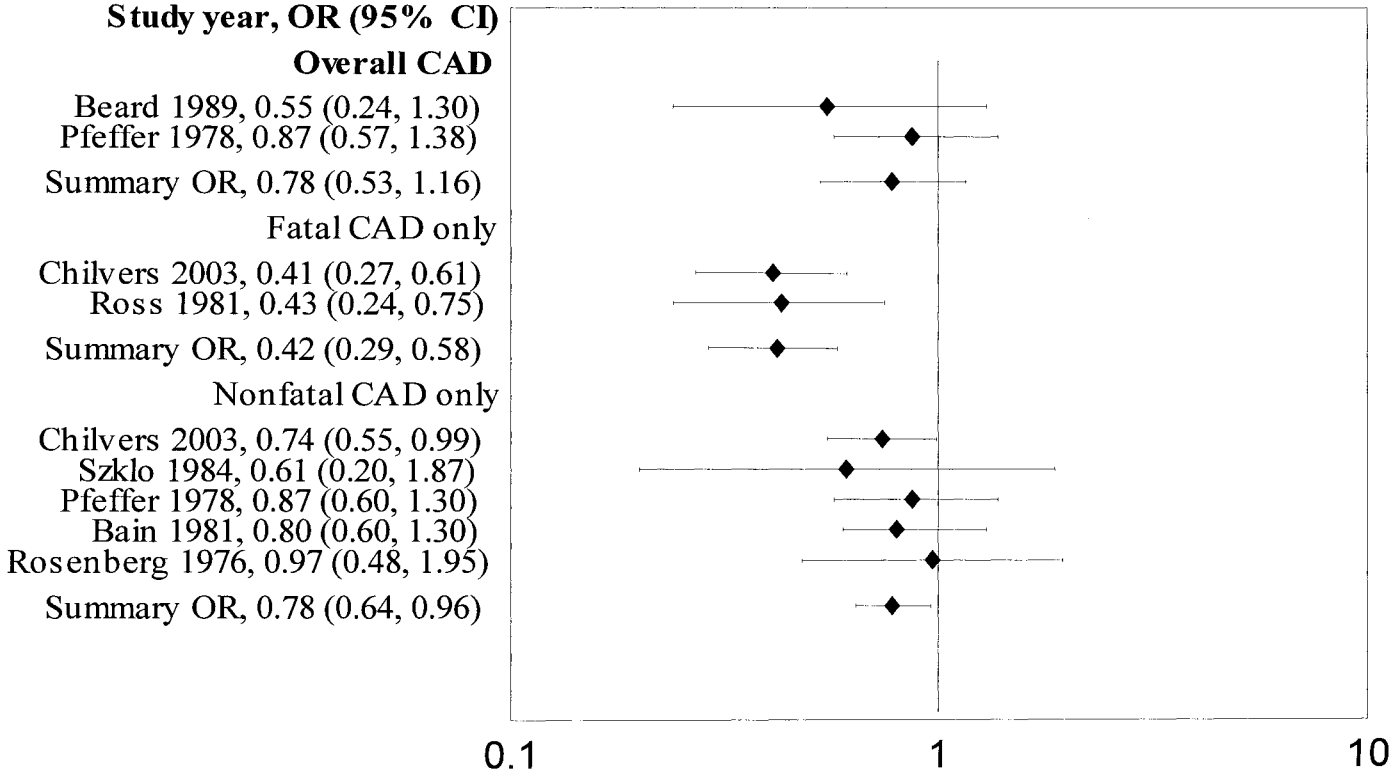


Figure 4.1.5 Risk of CAD in ever users of HRT versus never users of HRT among postmenopausal women in case control studies: odds ratios and 95% confidence intervals



five studies examining the association between ever use and nonfatal HRT the point estimates of all studies showed a decreased risk but achieved statistical significance in only one study. A pooled estimate using a fixed effects model was 0.78 (95% CI: 0.64, 0.96) demonstrating a significantly decreased risk for ever use of HRT and nonfatal CAD.

The chi-square of association and homogeneity are summarized in Table 4.1.6. The chi-square of association for overall CAD was not statistically significant indicating a lack of association between ever users of HRT and overall CAD. In contrast, ever use of HRT was associated with a decreased risk for fatal CAD alone and nonfatal CAD alone.

Table 4.1.6 Results of chi-square tests of association and homogeneity for HRT users versus nonusers in case control studies

Type of user	Number of studies	Summary OR (95% CIs)	$\chi 2$ association	$\chi 2$ homogeneity I^2 statistic (%)
Current users				
Overall CAD	N=5	0.75 (0.66, 0.84)	21.3178 (p<0.0001)	1.3192 (p=0.8581), 0%
Fatal CAD	N=1	0.47 (0.32, 0.69)	14.8371 (p<0.0001)	Not applicable
Nonfatal CAD	N=5	0.75 (0.60, 0.95)	9.4912 (p=0.0021)	3.6468 (p=0.4559), 0%
Past users				
Overall CAD	N=3	0.73 (0.59, 0.89)	8.8155 (p=0.003)	0.1168 (p=0.9433), 0%
Fatal CAD	N=1	0.99 (0.75, 1.3)	0.00513 (p=0.9429)	Not applicable
Nonfatal CAD	N=3	0.94 (0.74, 1.19)	0.2416 (p=0.6231)	1.0652 (p=0.5871), 0%
Ever users				
Overall CAD	N=2	0.78 (0.53, 1.16)	1.4166 (p=0.2340)	0.8887 (p=0.3458), 0%
Fatal CAD	N=2	0.42 (0.29, 0.58)	26.8015 (p<0.0001)	0.0178 (p=0.8939), 0%
Nonfatal CAD	N=5	0.78 (0.64, 0.96)	9.4900 (p=0.0021)	0.8563 (p=0.9307), 0%

The Cochran's Q test for homogeneity showed a p-value greater than 0.10 for both the groups implying the studies were not very different and a fixed-effects model was appropriate.

4.1.3.4 Subgroup analysis in case control studies

Analysis was performed on specific subgroups to further explore the association between HRT and CAD. For an exposure such as HRT, analysis by duration of use, estrogen alone versus estrogen and progestin and dose of estrogen in HRT preparation and risk of CAD may be insightful. The number of studies analyzing these parameters were small. Therefore, for the subgroup analysis we grouped current or ever users of

HRT as HRT users and all outcomes (overall CAD, fatal CAD or nonfatal CAD) as coronary artery disease.

Duration of HRT

Five studies analyzed the association between the duration of HRT and risk of CAD. A sixth study by Grodstein et al[169] analyzed the duration of HRT use with all cause mortality and was not included.

The duration of exposure in different studies varied from more than 3 years to more than 10 years. Three of the five studies found the risk of CAD decreased with increasing duration of HRT use, while two did not find any such association. The ORs and the 95% CIs of individual studies and the pooled estimate are depicted in Figure 4.1.6. Quantitative pooling of these studies by a fixed effects model yielded a summary estimate of 0.62 (95% CI: 0.51, 0.76) demonstrating a decrease in risk of CAD with increasing duration of HRT use. Table 4.1.7 shows the results of statistical testing of the summary estimates. The chi-square test of association for CAD was statistically significant ($\chi 2 = 21.3641$, $p < 0.001$). The p-value for the Cochran's Q test for homogeneity was greater than 0.10 implying the studies were not very different and a fixed effects model was appropriate.

Estrogen alone versus estrogen and progestin

Six studies examined if there was a difference between the risk of CAD and use of estrogen alone versus use of estrogen and progestin. The ORs and 95% CIs for the individual studies are depicted in the Figure 4.1.7. One study[178] found the users of estrogen alone had a decreased risk of CAD while women on estrogen and progestin did not benefit. Mann et al[171] found the opposite association with users of estrogen and progestin showing a significant decrease in risk of CAD. The remaining four studies had consistent results showing no significant difference between users of estrogen alone or estrogen and progestin. On pooling the results from these six studies using a fixed effects model, summary estimates of 0.78 (95% CI: 0.66, 0.93) and 0.77 (95% CI: 0.65, 0.92) were achieved for estrogen alone and estrogen and progestin and risk of CAD respectively. HRT use with estrogen alone or in combination with progestin decreased

Figure 4.1.6 Risk of CAD by duration of use of HRT among postmenopausal women in case control studies: odds ratios and 95% confidence intervals

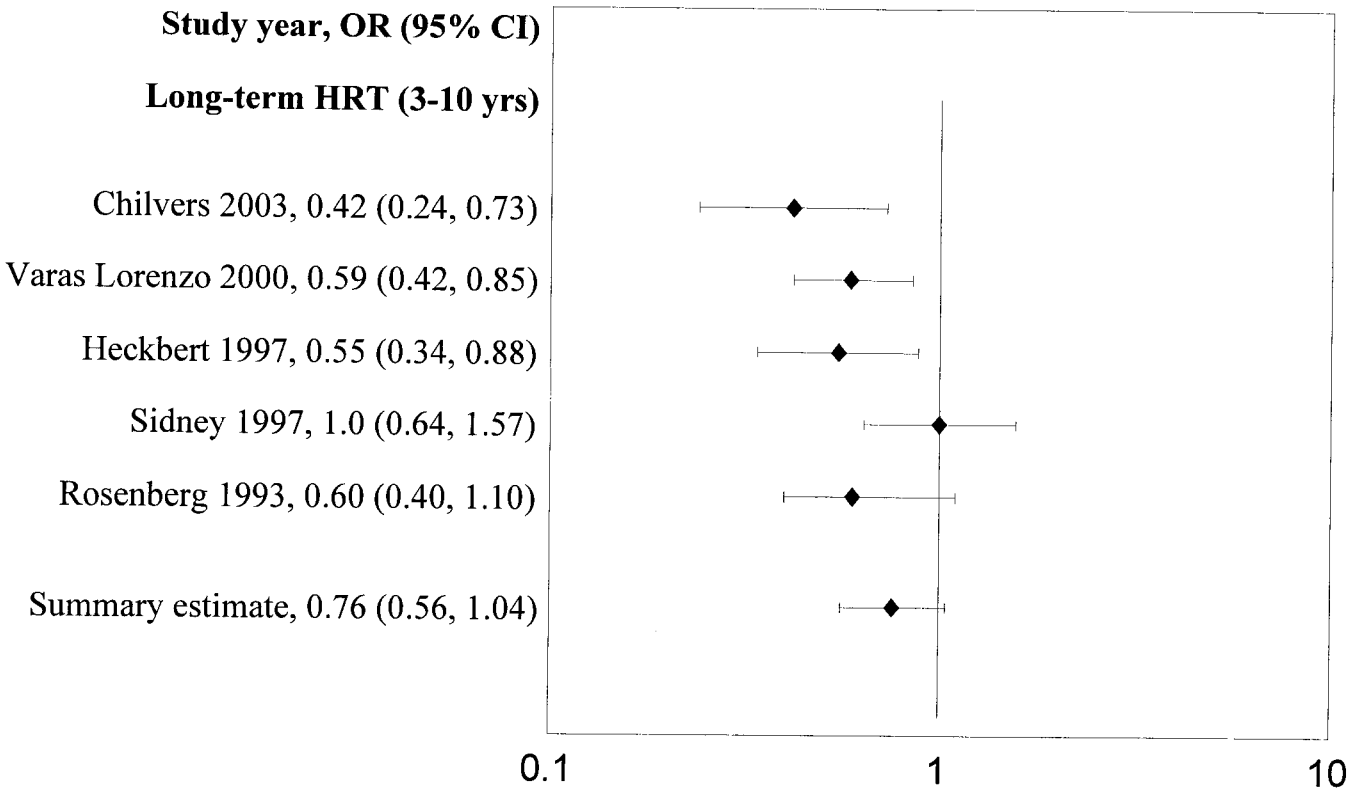
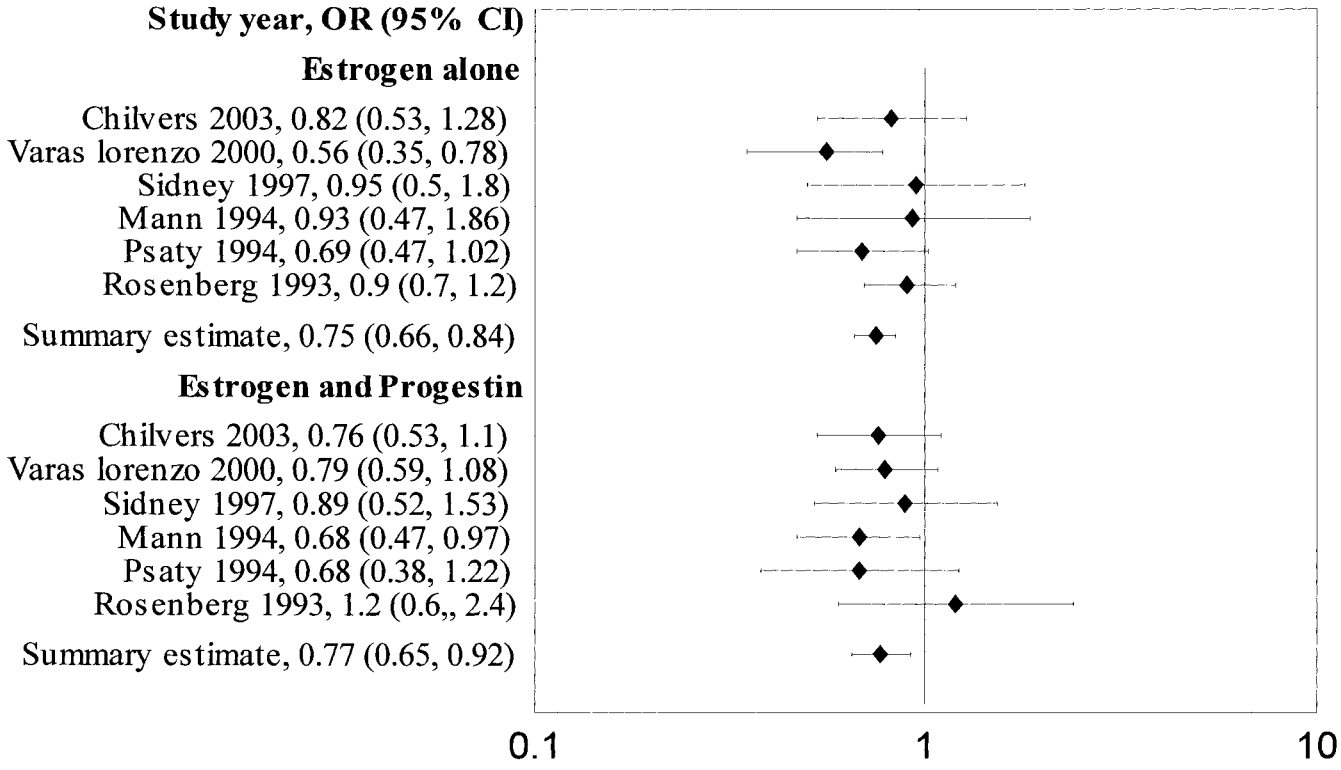


Figure 4.1.7 Risk of CAD by use of estrogen alone versus use of estrogen and progestin in case control studies: odds ratios and 95% confidence intervals



the risk of CAD to the same extent. The chi square tests for association and heterogeneity were performed with results shown in Table 4.1.7. The results suggest a decreased risk of CAD in HRT users which was similar for both regimens. Also, the Cochran's Q test for heterogeneity had a p-value greater than 0.10 suggesting little variation in the OR estimates among the studies.

Dose of estrogen in HRT

Three studies evaluated the difference in risk of CAD with differing doses of estrogen in the HRT used. Two of these divided the estrogen dose into three potency groups: low (< 0.625 mg), medium (0.625 mg) and high (≥ 1.25 mg). One study divided them into two groups, low (≤ 0.625 mg) and high (> 0.625 mg). For this review, estrogen dose greater than 0.625 mg was classified as high potency.

Two of these three studies found no difference between the different doses of estrogen and risk of CAD. The study by Varas Lorenzo[178] reported only medium potency estrogen significantly decreased the risk of CAD while the low and high potency estrogen use did not have a protective effect. Quantitative pooling was performed using a fixed effects model and summary estimates for low, medium and high potency of estrogen and risk of CAD were 0.85 (95% CI: 0.56, 1.29), 0.56 (95% CI: 0.42, 0.79) and 0.75 (95% CI: 0.51, 1.1) respectively (Figure 4.1.8). These results demonstrated that only medium potency dose of estrogen was associated with a significant decrease in the risk of CAD.

The chi square tests for association and heterogeneity were performed and results are shown in Table 4.1.7. The results suggest a decreased risk of CAD in women who used medium potency estrogen but not in women who used low or high potency estrogen. This is difficult to explain biologically. The Cochran's Q test for heterogeneity had a p-value greater than 0.10 suggesting little variation in the OR estimates among studies and the choice of a fixed effects model was appropriate.

Figure 4.1.8 Risk of CAD by differing estrogen dose in HRT among postmenopausal women in case control studies: odds ratios and 95% confidence intervals

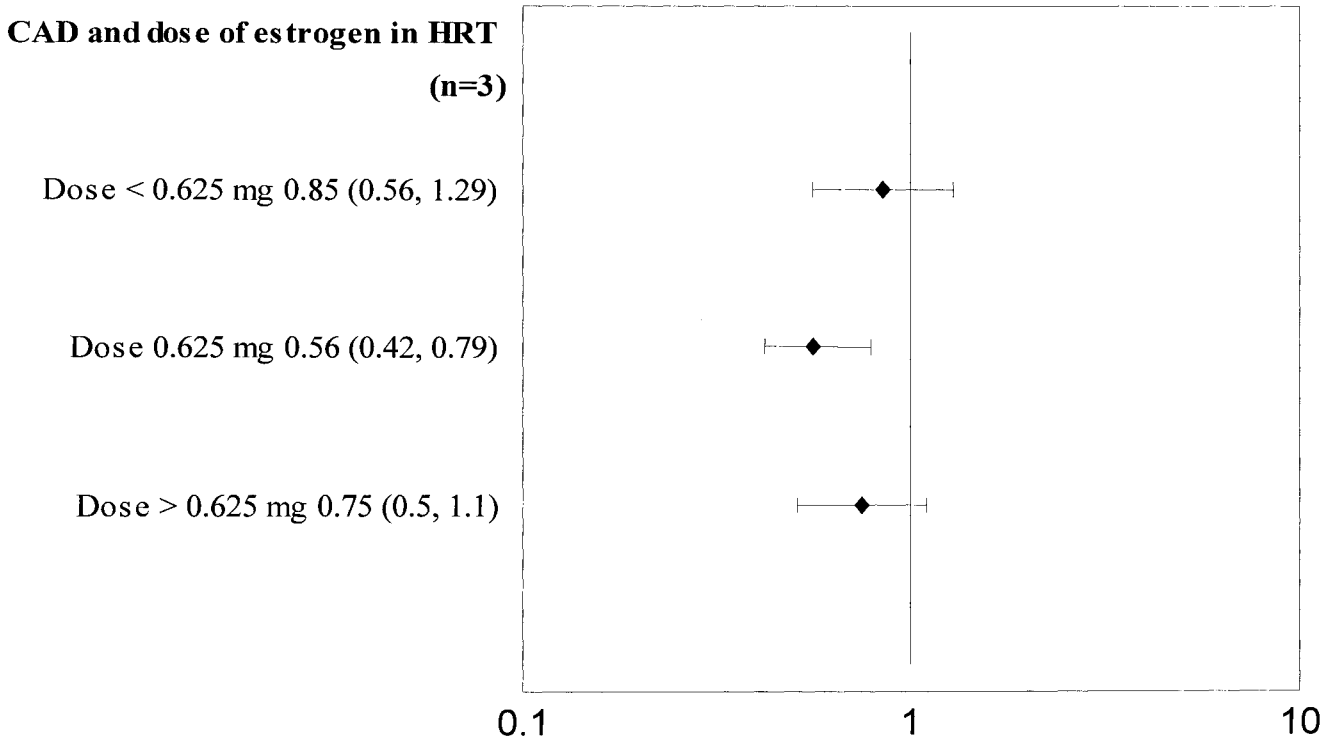


Table 4.1.7 Results of chi-square association and homogeneity for subgroup analysis in case control studies

Subgroup studied	Number of studies	Summary OR (95% CIs)	$\chi 2$ association	$\chi 2$ homogeneity, I^2 statistic (%)
Duration of HRT (3-10 years)	N=5	0.62 (0.51, 0.76)	21.3641 (p<0.0001)	6.5798 (p=0.1598), 39.2%
Estrogen vs	N=6	0.78 (0.66, 0.93)	8.2459 (p=0.0041)	4.7557 (p=0.4464), 0%
E+P	N=6	0.77 (0.65, 0.92)	8.4841 (p=0.0036)	2.5006 (p=0.7764), 0%
Dose of estrogen				
<0.625mg	N=3	0.85 (0.56, 1.29)	0.5707 (p=0.4500)	0.9985 (p=0.6070), 0%
0.625 mg	N=2	0.56 (0.42, 0.79)	11.5576 (p=0.0007)	0.1065 (p=0.7442), 0%
>0.625 mg	N=3	0.75 (0.51, 1.1)	2.1656 (p=0.1411)	0.2957 (p=0.8626), 0%

4.1.3.5 Sensitivity analysis for case control studies

Study quality

The quality scores of the 19 case control studies as determined by the NOS varied between a low score of three and a maximum score of nine. One study with a score of three did not provide summary estimates and could not be integrated into the quantitative analysis. Summary estimates were calculated using a fixed effects model for the set of studies with the same NOS scores. The summary ORs are plotted in Figure 4.1.9 in order of study quality starting with the highest quality studies, to evaluate if the variation in study quality could explain some of the variation seen in the results of the studies. The size of the RRs and the width of the confidence intervals were assessed. We could not find any consistent trend in the size of the RR or the width of confidence intervals with the study quality. All studies with scores of five and below failed to show any decrease in risk of CAD in HRT users. The one study with a score of three reported that HRT use was associated with a significant increase in the risk of CAD.

We conducted a cumulative meta-analysis starting with the highest quality studies and adding in studies with a lower score. The results of this meta-analysis are illustrated in Figure 4.1.10. All the summary estimates were associated with a decreased risk of CAD among HRT users. The summary estimates and the width of confidence intervals did not change with addition of poor quality studies. In this set of studies, quality scoring did not assist in explaining the variability of results.

Figure 4.1.9 Sensitivity analysis of study quality for risk of CAD in HRT users in case control studies: odds ratios and 95% confidence intervals

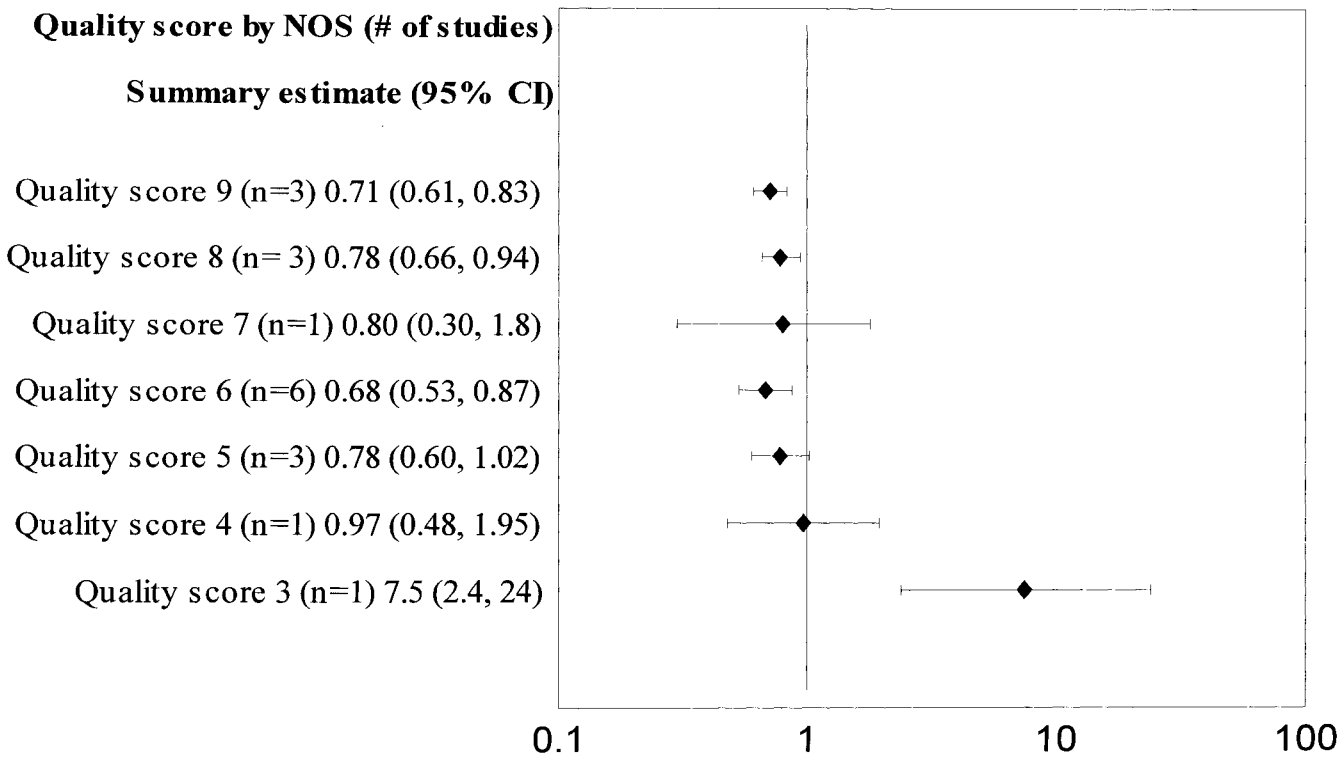


Figure 4.1.10 Cumulative* meta-analysis of case control studies according to quality scores: odds ratios and 95% confidence intervals

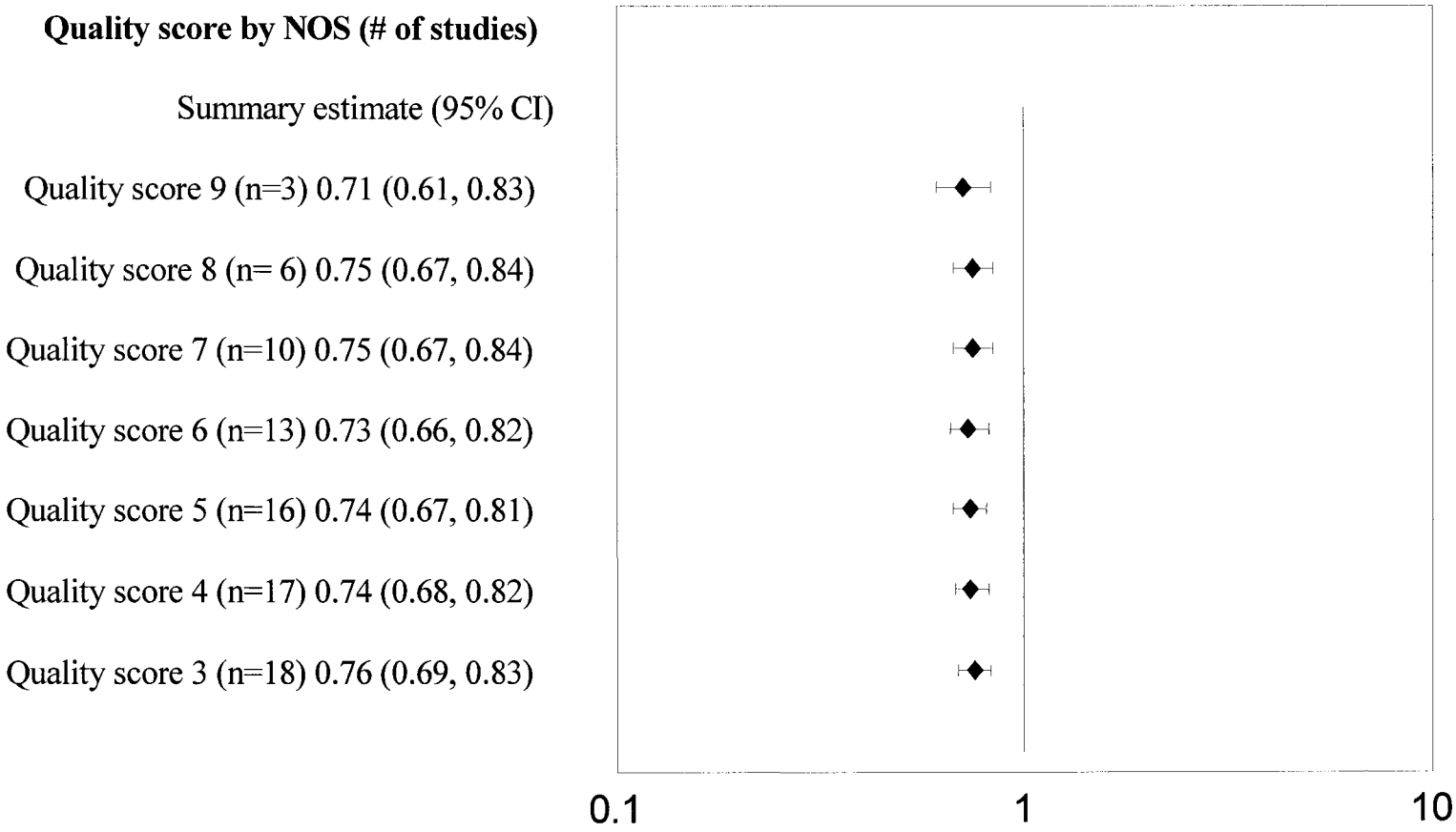
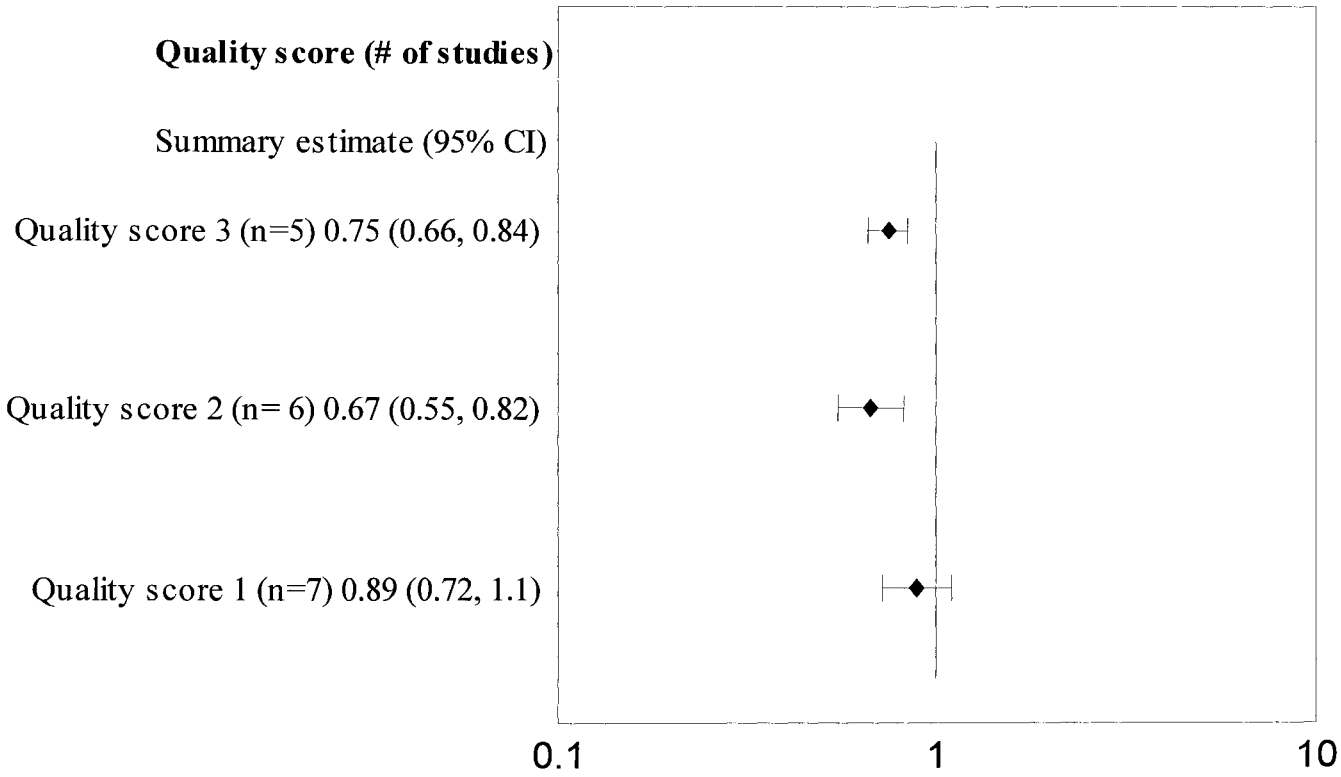


Figure 4.1.11 Sensitivity analysis of study quality for risk of CAD in HRT users in case control studies: analysis by exposure ascertainment component of the NOS



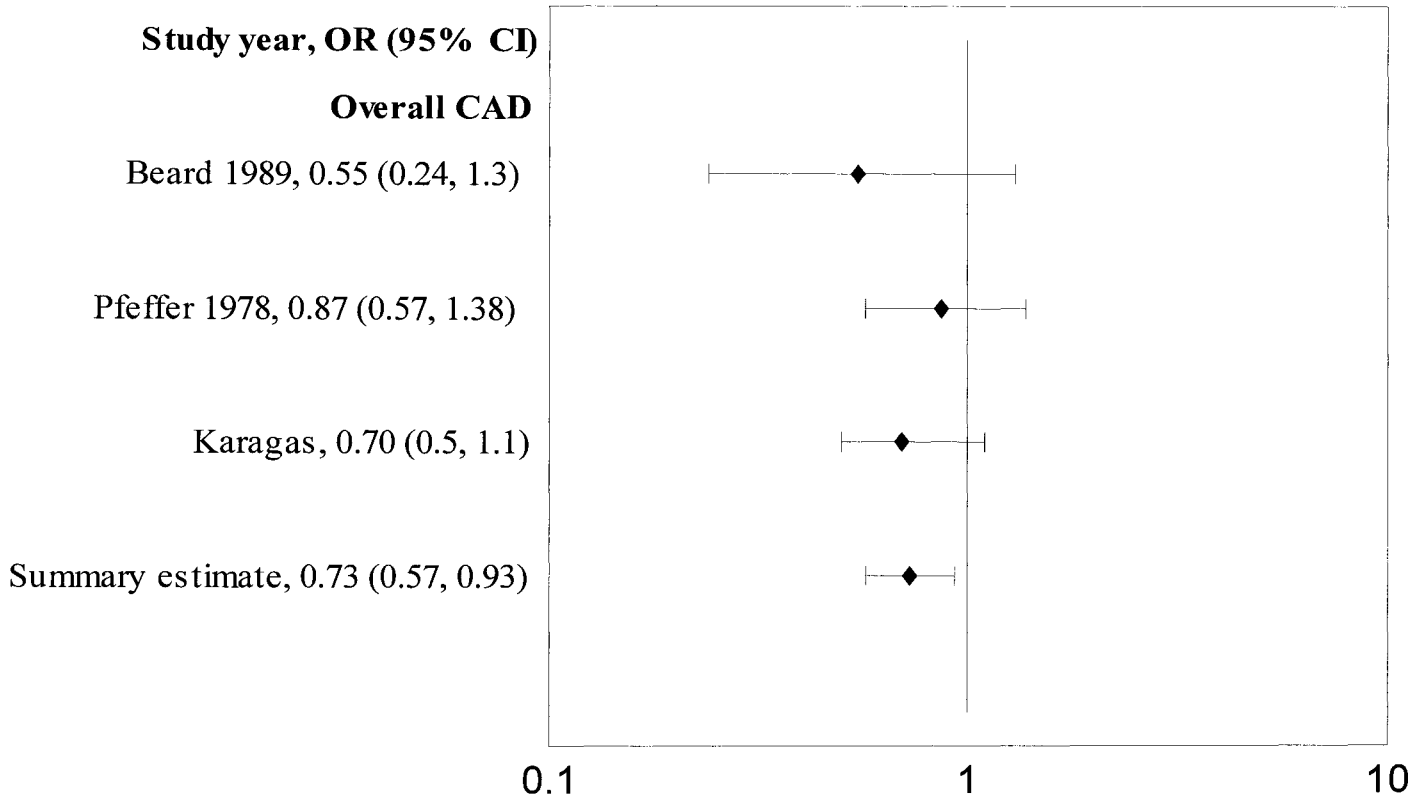
Besides the overall score, studies were also analyzed according to their scores on each of the three components. For case control studies, the exposure ascertainment is often the most difficult to measure accurately. Summary estimates were calculated using a fixed effects model for the set of studies with the same score on exposure ascertainment. The summary ORs are plotted in Figure 4.1.11 in order of study scores starting with the highest scoring studies to evaluate if some of the variation seen in the results of the studies could be explained by differences in the accuracy of exposure ascertainment. Pooled odds ratio for studies scoring high on this component had a decreased risk of CAD with narrower confidence intervals than studies which had a lower score. Pooled odds ratio for studies with a low score of one had a low risk of CAD that was not statistically significant.

Including unpublished literature

We identified one unpublished case control study from dissertation abstracts. This study was not included in the main analysis. This was a population based case control study by Karagas et al[179] done in North America. Cases were defined as women who had MI or fatal CAD. The exposure studied was ever use of HRT determined from a pharmacy database. There was no data on the preparation or route of HRT. The analysis was controlled for age, race and smoking. The authors found a slight decrease in risk of CAD in users of estrogen alone which was not statistically significant (OR: 0.7; 95% CI: 0.6, 1.1). For users of estrogen and progestin, the effect was more neutral (OR: 0.9; 95% CI: 0.6, 1.4). The authors also found long term use and early age at first use were associated with a particularly reduced risk of CAD risk but no ORs were mentioned. They found women with a surgical menopause showed a greater decrease in risk of CAD than those with natural menopause. None of the ORs achieved statistical significance.

The results of the unpublished study were pooled with the two published case control studies assessing the association between ever use of HRT and overall CAD using a fixed effects model (Figure 4.1.12). A summary estimate of 0.73 (95% CI: 0.59, 0.93) was obtained showing a decreased risk of overall CAD in ever users which was statistically significant. The summary estimate for the two published studies only was

Figure 4.1.12 Sensitivity analysis by including unpublished literature for the risk of CAD in ever users of HRT in case control studies: odds ratios and 95% confidence intervals



0.78 (95% CI: 0.53, 1.16) with a lack of statistical significance. The addition of gray literature accentuated the treatment effect.

Results of chi-square tests of association and heterogeneity are shown in Table 4.1.8. The chi-square for association was statistically significant suggesting a decreased risk of CAD in ever users of HRT. The chi-square test for homogeneity was not significant implying lack of statistical heterogeneity among the studies.

Including studies adjusting for socioeconomic status

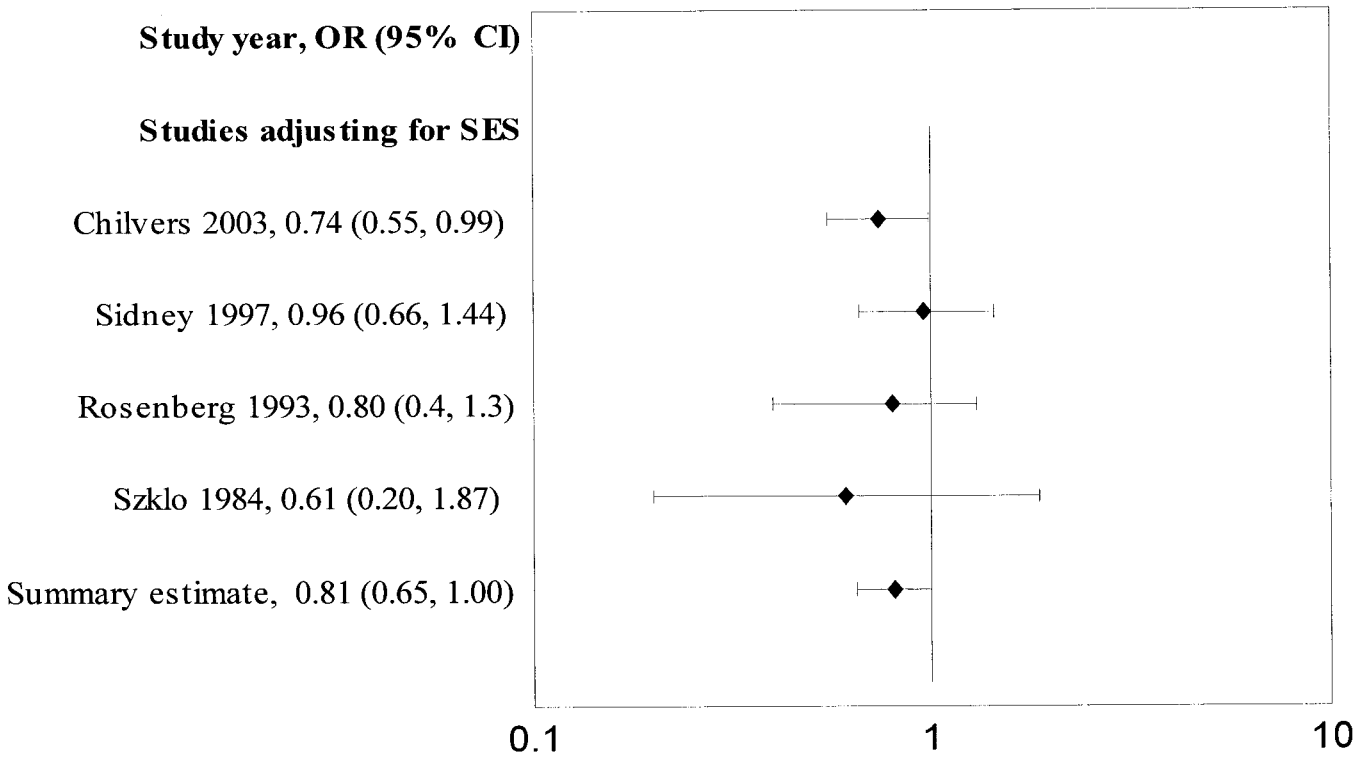
A large number of non-randomized studies commented on the differences in baseline characteristics between the cases and the controls. A prominent hypothesis was women who agree to take HRT are healthier to begin with and the lower risk of CAD in them is a reflection of 'healthy user' effect rather than HRT use. Several factors may contribute to this effect: more exercise, better nutrition, and awareness of healthy life styles. All these factors are associated with the socioeconomic status of women. Socioeconomic status is a composite factor and has been addressed in different studies by studying social class, income, or education. We examined studies adjusting for any of these factors and the association between HRT and risk of CAD.

Four of the 14 studies examined the association between HRT and CAD after adjusting for socioeconomic status. Of these, two studies adjusted for education, and two for social class. The results of the individual studies are shown graphically in Figure 4.1.13. Three of the studies showed a decreased risk of CAD which was not statistically significant while one study showed a significant decrease. A pooled estimate using a fixed effects model was 0.81 (95% CI: 0.65, 1.00), a value closer to the null and the upper confidence interval was 1.00 showing the effect was borderline significance. Adjusting for socioeconomic status nullified the cardio-protective effect seen among HRT users.

Table 4.1.8 Results of chi-square association and homogeneity for sensitivity analysis in case control studies

Sensitivity analysis	Number of studies	Summary OR (95% CIs)	$\chi 2$ association	$\chi 2$ homogeneity, I2 statistic (%)
Including unpublished literature	N=3	0.73 (0.57, 0.93)	6.5163 (p=0.0107)	1.1098 (p=0.5741), 0%
Studies that adjusted for SES	N=4	0.81 (0.65, 1.00)	3.9851 (p=0.0.0459)	1.3947 (p=0.7068), 0%

Figure 4.1.13 Sensitivity analysis of studies adjusting for socioeconomic status as a confounding variable in case control studies: odds ratios and 95% confidence intervals



Results of chi-square test of association and heterogeneity are shown in Table 4.1.8. The chi-square for association was just significant with the p-value at 0.045. The chi- square for homogeneity was not significant implying lack of statistical heterogeneity among studies.

4.1.4 Cohort studies on HRT and CAD in postmenopausal women

4.1.4.1 Characteristics of cohort studies

The inclusion/exclusion criteria for the review were met by a total of 30 published cohort studies and one unpublished cohort study. The unpublished study was included only in a sensitivity analysis and is described in section 4.1.4.5. Of the published studies, eight included women with pre-existing heart disease only and studied the association between HRT and recurrent CAD. In the remaining 22 studies, 17 studies excluded women with pre-existing heart disease while four studies did not exclude women with pre-existing heart disease but adjusted for it in the analysis.

Among the 22 studies, only three[180-182] were retrospective longitudinal studies while the remaining were prospectively designed studies. In addition, three studies published several follow up studies. First, Grodstein et al published data on the Nurses Health Study cohort in 1985[183],1996[184] and most recently in 2000[185]. The results from the most recent publication were included and a pertinent sub-group analysis from the 1996 publication analyzing the lack of continued benefit of HRT after stopping HRT was used in the descriptive review. Secondly, Henderson et al published results from their cohort three times between 1986 to 1991[186-188]. The publication from 1988 is being used for this meta-analysis as the most recent publication does not present 95 % confidence intervals for the summary statistic of CAD mortality with use of HRT. Third, results from the Framingham cohort[189;190] have been published twice, once in a peer reviewed journal and again as a special report. The article in the peer reviewed journal was retained for this review. Thus, there were 17 unique cohort studies.

The main characteristics of the 17 unique studies[164;180;181;185;187;190-201] are summarized in Tables 4.1.9. Four of the 17 studies were conducted in Europe and the remainder in the United States. The majority of cohorts were population based with

exceptions being Ettinger et al[193] who assembled a cohort of users and non-users from a pharmacy record, and Lafferty et al[196] and Lauritzen et al[197] who used their private practices for their cohorts.

The average age at the inception of cohorts is shown in Table 4.1.9. Three studies out of 17 were excluded from the quantitative review because in one study women were included who were not all menopausal and in two studies relative risks and 95% CIs for the results were not provided. These studies are described briefly.

Pettiti et al[199] included women between the ages of 18 to 54 years who used hormones other than oral contraceptive pills but did not state whether the women were menopausal. Most of the women used oral estrogen alone with a few using it in combination with progestins (numbers not specified). Outcomes of interest included all cause mortality, mortality due to cancers, cardiovascular disease, and accidents. The age-adjusted mortality rates for CAD and HRT use was 1.3 (95% CI: 0.2, 7.7). Multivariate analysis with adjustment for confounding variables was not presented for CAD. Mortality from accidents, homicide, and suicide was lower in HRT users making interpretation of overall death rate problematic. In the discussion, the study cautioned against the advocacy of HRT as an 'all purpose anti-aging' drug until its effect on cardiovascular system was clearer.

Hammond et al[181] sought all women with symptoms of estrogen deficiency including women with gonadal dysgenesis and hypopituitarism. The average age for women at the time of entry into study was 49.6 ± 0.58 years in nonusers and 42.9 ± 0.88 years in users indicating most women were peri-menopausal. The study included only those women who had taken oral estrogens for at least 5 years. The dose and the preparations are not mentioned. The study evaluated long-term effects of HRT and cardiovascular events were one of the outcomes studied. The unadjusted RR for the risk of new events of CAD in HRT users was 0.32 ($p<0.01$) but the confidence intervals were not provided and this study was not pooled quantitatively.

Lauritzen et al[197] conducted a study in their private clinic of 458 women half of who received estrogen and half of who did not receive estrogen therapy. The two groups were comparable in age, parity, weight, social status and race. In this cohort, the HRT used was oral CEE (28%), CEE with progestin (26%), intramuscular(I/M) estriol (29%),

Table 4.1.9 Characteristics of cohort studies on the association between HRT use and CAD in postmenopausal women

Study, year, Location	Age at inception (yrs)	Outcome	Exposure status	No. of cases or deaths	No. of women or person-years (py)	RR (95% CI)	Variables adjusted	Duration of FU (yrs)
Lokkegard 2003 Denmark	≥ 45 yrs	Overall CAD	Current Ever	81 153	13084 women	0.68 (0.30, 1.49)* 0.7 (0.33, 1.54)*	Age, DM, smoking, HT, BMI, FH, alcohol use, h/o angina, physical activity, self rated health	5 yrs
Hedblad 2002 Sweden	54 yrs (mean)	Overall CAD	Current	97	5721 women	0.37 (0.15, 0.90)	Age, DM, Smoking, HT, BMI, FH, h/o CAD, social class, age at mp	9.3 yrs
Rodriguez 2001 USA	60-80 years	Fatal CAD	Current	227	290827 women	0.66 (0.58, 0.77)	Age, DM, smoking, BMI, alcohol use, education, age at mp, race	9 yrs
Grodstein 2000 USA	35-50 years	Overall CAD	Current	259	265203 py	0.61 (0.52, 0.71)	Age, DM, smoking, BMI, HC, FH, age at mp	20 yrs
Sourander 1998 Finland	60 yrs (mean)	Nonfatal CAD Fatal CAD	Current Current	48 3	988	1.05 (0.76, 1.36) 0.19 (0.05, 0.77)	Age, DM, smoking, BMI, SES, HT, CHF, CAD	8 yrs
Cauley 1997 USA	≥ 65 yrs	Fatal CAD	Current	144	1258	0.49 (0.26, 0.93)	Age, DM, smoking, BMI, education, HT, exercise, surgical mp	6 yrs
Ettinger 1996 USA	59 years (mean)	Fatal CAD	Ever	10	232	0.4 (0.16, 1.02)	Age, smoking, BMI, HC, HT, alcohol use, abnormal ECG	12 yrs

DM: diabetes mellitus; HT: hypertension; FH: family history; HC: high cholesterol; mp: menopause

* Excluding women with diabetes

Table 4.1.9 Contd. Characteristics of cohort studies on the association between HRT use and CAD in postmenopausal women

Study, year, Location	Age at inception (yrs)	Outcome	Exposure status	No. of cases or deaths	No. of women or person-years (py)	RR (95% CI)	Variables adjusted	Duration of FU (yrs)
Folsom 1995, USA	55-69 yrs	Fatal CAD	Current	24	25356 py	0.82 (0.47, 1.43)	Age, smoking, DM, BMI, HT, physical activity	6 yrs
Beard 1995 USA	40-59 yrs	Overall CAD Nonfatal CAD	Ever Ever	2 9	7425 py (both users and nonusers)	0.25 (0.05, 1.22) 1.34 (0.52, 3.46)	Age, smoking, renal disease	15 yrs
Lafferty 1994, USA	52 yrs (mean)	Nonfatal CAD	Ever	1	81	0.34 (0.09, 1.34)	Age	25 yrs
Avila Hernandez 1990, USA	50-64 yrs	Overall CAD	Ever	12	17513 py	0.7 (0.4, 1.3)	Age	6 yrs
Henderson 1988, USA	73 yrs (median)	Nonfatal CAD	Ever		79	0.72 (0.53, 0.97)	Age,DM, smoking, HT, BMI,exercise, previous CAD, age at mp, type of mp	6 yrs
Criqui 1988 USA	50-79 yrs	Fatal CAD	Ever	95	734	0.99 (0.59, 1.67)	Age, smoking, DM, HC, BMI, SES, HT	12 yrs
Pettiti 1987 USA	18-54 yrs	Fatal CAD	Ever	9	36100 py	1.3 (0.2, 7.7)	Age	10 yrs
Wilson 1985, USA	50-83 yrs	Overall CAD	Ever	116	Not presented	1.9	Age, smoking, BMI, HC, HT, alcohol use	8 yrs
Lauritzen 1979, UK	Not stated	Nonfatal CAD	Ever	6	2786 py	P<0.01 , RR not presented	Matched on age, body weight, social status	5.3 yrs
Hammond 1979, USA	43 yrs (mean)	Overall CAD	Ever	14	301	P< 0.01, RR not presented	None	1.1 yr

DM: diabetes mellitus; HT: hypertension; FH: family history; HC: high cholesterol; mp: menopause

and I/M estriol with testosterone in the remainder. The mean duration of treatment was 5.3 years. They reported a lower incidence of MI among HRT users when compared to nonusers ($p<0.01$). The study suggested long-term use of HRT did not appear to increase cardiovascular disease. Summary estimate and confidence intervals were not presented.

Almost all studies included women with natural or surgical menopause with the exception of one study by Beard et al[180] which included women with a surgical menopause only.

The details of exposure ascertainment for the included 14 studies are summarized in Table 4.1.10. The following section describes the method, frequency and classification of HRT use and the route and preparations used in the cohort studies. Publications used self-reporting, interviews with women and pharmacy or medical records for ascertainment of exposure to HRT. Some attempts to improve the recall for HRT usage included mailing pictures of the pill along with the questionnaire[187].

The exposure status was classified as current by eight studies, and ever by the remaining six. Two studies, Lokkegaard et al[198] and Avila et al[164], analyzed data for current and ever users. All the eight studies defined current users as women using HRT at the time of initiation of study or at baseline. Some of the studies further classified current users as women who had been on HRT for a year at the time of ascertainment. Among the six studies classifying exposure as ever use, four did not define any duration of use. One study included women who had taken HRT for at least five years and one study defined users as those that had taken HRT for one year. Most studies did not verify whether women who were classified as users of HRT continued to be users during the follow-up period. Seven out of the 14 studies had a single ascertainment of exposure and assumed users continued to be users. Two studies verified the exposure classification a second time during the study and five studies repeatedly collected information on HRT use and re-classified women accordingly. Mis-classification bias would be more likely in studies who determined HRT status only once. Frequent ascertainment would allow the exposure status to be determined more accurately and allow analysis of past users and current users with nonusers. Criqui et al[192] reported that a repeat questionnaire after 9 years revealed only 198 out of 650 women classified as users at first ascertainment

Table 4. 1. 10 Ascertainment of exposure in cohort studies examining the association between CAD and HRT in postmenopausal women

Author year	Ascertainment of HRT	Frequency of ascertainment	Data on dose, route, regimen, duration	Definitions of HRT use
Lokkegard 2003	Self reported questionnaires	Single ascertainment	Oestradiol and NETA used by 85%, CEE: 15%. Analysis by regimen	Current, past, ever and never: not defined
Hedblad 2002	Self administered questionnaire	Single ascertainment	No data on dose, duration, type or route of administration	Current /never not defined.
Rodriguez 2001	Self reported questionnaires	Single ascertainment	Included women on oral HRT only, excluded women with estrogenic creams or injections	Baseline: women using at baseline or within a year. Former users: stopped at least a year ago
Grodstein 2000	Self reported	Updated every 2 years: done 10 times	Oral CEE used. Analysis by dose, duration, regimen	Current, past and never users: not defined
Sourander 1998	Questionnaire filled with the help of trained nurses	At baseline and every 2 years : done 3 times	Oral estradiol used; 139 women used E + P. No analysis on dose, duration or regimen.	Current: using at enrollment or within a year, former: stopped over a year ago and never users
Cauley 1997	Interview	Updated once during the study	Oral preparations only, analysis by duration	Current: use >1 yr; past use: stopped a year ago; never users: < 1 year use
Ettinger 1996	Computer pharmacy records	Single ascertainment	Oral estrogen used. Mean daily dose: 0.625 mg; few used E+ P;	Ever user: Used ERT for at least 5 years.
Folsom 1995	Self –reported by questionnaire.	Three follow up questionnaires in 1987, 1989 & 1992	Oral HRT, analysis by duration	Current, former or never use at baseline

Table 4.1.10 Contd. Ascertainment of exposure in cohort studies examining the association between CAD and HRT in postmenopausal women

Author year	Ascertainment of HRT	Frequency of ascertainment	Data on dose, route, regimen, duration	Definitions of HRT use
Beard 1995	Abstracted from medical records	Single abstraction	Oral CEE, 82% unopposed, no separate analysis on dose, duration or regimen	Estrogen ever users and never users.
Lafferty 1994	Interview at baseline	Two yearly follow up visits till 1989	Oral CEE: 0.625 mg daily. MPA 5 mg to women after 1983	Ever and never use: not defined
Hernandez Avila 1990	Pharmacy records	Single ascertainment	Oral CEE or EE, analysis on dose and duration given	Current: >1 year use, ever and never users.
Henderson 1988	Self reported questionnaire	Single ascertainment	Injectable, oral and vaginal estrogen use included, analysis by duration and dose	Ever and never users. Ever users had used it more than 1 year
Criqui 1988	Interview at baseline	Repeat interview after 9 years	Oral estrogens, no data on dose, duration of regimen	Ever and never users.
Wilson 1985	Interview at baseline of exposure over a 10 year span	HRT exposure used as an interval variable	Oral CEE, No analysis on dose, duration or regimen	Estrogen users and nonusers

continued to use HRT. The misclassification was differential as 95% of non-users remained non-users.

Oral HRT was the most common route of administration in the cohort studies. None of the studies provided a separate analysis by route of administration. The most common oral preparation was CEE (9 studies) and then EE (3 studies). In studies where women used EE, a proportion of women also used CEE. Several studies have not mentioned the exact preparation used. Rodriguez et al[200] explicitly stated women who used any route other than oral were excluded. Earlier studies used estrogen alone but with studies demonstrating increased incidence of endometrial cancer among estrogen users, the later studies mention the concomitant use of progestins. The most common progestin used was MPA. The actual number of users was not mentioned in majority of the studies.

The total duration of exposure varied greatly between studies ranging from 2 to 18 years in different cohorts. Several cohorts did not express their results as mean duration of exposure, for example, five studies classified duration as more than five, eight or ten years. Most studies were long-term follow up studies with exposure to HRT over several years. The association between duration of exposure, dose of estrogen in HRT and regimen of HRT was explored in subgroup analysis.

Cauley et al[191] was the only study to assess the association between the length of time to initiation of HRT after menopause and risk of CAD. The authors reported a significant decrease in all cause mortality for HRT users who initiated HRT within five years of menopause (RR: 0.47, 95% CI: 0.29, 0.75) than those who started after a gap of five years after menopause (RR: 0.60, 95% CI: 0.29, 1.30). Several studies[180-182;188;190;192;193], recruited women between 50 to 55 years of age ensuring early initiation of HRT after menopause. Results from these cohorts were examined in a sensitivity analysis to study if early initiation of HRT has a different effect on the risk of CAD in postmenopausal women.

Almost all cohort studies assessed several outcomes with CAD being one of them. For CAD, the outcomes include overall CAD, only fatal CAD and only nonfatal CAD. Five of the 14 studies presented their analysis for overall CAD, five presented results for only fatal CAD, one study analyzed only nonfatal CAD and three studies

presented data for more than one outcome. These included the study by Beard et al[180] (overall CAD and nonfatal CAD), Sourander et al[201] and Henderson et al[187] who presented data separately for fatal and nonfatal CAD.

There are several known risk factors which could potentially confound the association between HRT use and CAD. In order to study the association between HRT and CAD in non-randomized studies, it is important that studies control for pertinent confounding variables. The majority of studies controlled for some of the confounding variables as summarized in Table 4.1.9. A few studies provided adjusted relative risks for all cause mortality or cardiovascular disease but did not provide adjusted risk for CAD separately. Of the 14 studies, 12 adjusted for several confounders while Lafferty et al[196]and Avila et al[164] adjusted only for age.

Five studies reported significant interactions between confounding variables, HRT use and risk of CAD. Lokkegard et al[198] reported the risk of CAD was marginally increased in ever users of HRT when compared to nonusers. This became insignificant if ever users were classified as current and past users. The effect of HRT was modified by diabetes, with a significantly increased relative risk (RR: 2.9, 95% CI: 1.7, 7.9) in ever users with diabetes while the RR for ever users without diabetes had no increased risk of CAD (RR: 0.7, 95% CI: 0.33, 1.54). The relative risk was further increased in current users of HRT with diabetes. These findings persisted after adjusting for smoking, hypertension, high BMI, low self-rated health and previous angina.

Rodriguez et al[200] evaluated the effect of body mass on the association between HRT and CAD mortality in women older than 60 years. They reported fatal CAD increased with obesity (BMI >30 kg/m^2) and decreased with estrogen use. There was a significant interaction between BMI and HRT use on the risk of CAD ($p<0.02$), with HRT use decreasing the risk of CAD in thinner women (RR: 0.49, 95% CI: 0.37, 0.65). This protective effect was lost in obese women (RR: 0.95, 95% CI: 0.65, 1.39).

Criqui et al[192] reported results on postmenopausal estrogen use and all-cause mortality and specifically CAD mortality. They found all-cause mortality was doubled in current smokers after adjusting for confounding variables. They also reported a significant interaction between estrogen use and smoking, with estrogen use being associated with a decreased risk in current smokers (RR 0.35, 95% CI: 0.10, 1.26) and an

increased risk in past smokers (RR: 2.34, 95% CI: 0.83, 6.6). The authors could not explain this interaction physiologically and commented the population of women who had stopped smoking and used HRT in their sample was perhaps, not random. A study in the same year by Henderson et al[187] reported contrary results. This study evaluated the effect of HRT on mortality, and reported a significantly decreased incidence of fatal CAD in estrogen users. This protective effect persisted after adjusting for known risk factors for CAD except in the subgroup of current smokers. The authors hypothesized smoking had a major impact on estrogen metabolism which may be responsible for the lack of protection. The Framingham study[190] on postmenopausal estrogen use, smoking and cardiovascular disease reported an apparent interaction between cigarette smoking and estrogen use. They performed a separate multivariate analysis for smokers and nonsmokers and reported adjusted relative risks for CAD was much higher in smokers (4.17, $p < 0.01$) than in nonsmokers (1.44, p-value non-significant).

Almost all studies reported a difference in baseline characteristics between users and nonusers of HRT. The baseline characteristics generally reported included age, BMI, lipid levels, smoking status, social class or education, blood pressure, and alcohol use. In these studies, estrogen users tended to be leaner (either determined by BMI or body weight), and of younger age. Two of the studies found estrogen users were more likely to be Caucasians[193;200]. Estrogen users tended to have lower blood pressure, belonged to a higher socioeconomic status, performed more exercise, had a healthier lipid profile and a lower blood sugar level. Smoking was variable as three studies reported estrogen users smoked more often than their nonuser counterparts and one found the contrary. Presence of such confounding variables emphasizes the need to adjust for these factors to study the association between HRT and CAD.

The Framingham study[190] evaluated postmenopausal estrogen use and cardiovascular morbidity in women over 50 years of age. In this study, only oral estrogen was used and women were classified as either users or nonusers. CAD outcomes consisted of silent angina, MI, MI requiring hospitalization, and fatal CAD. Estrogen users had a statistically significant increase in the risk of CAD with an adjusted RR of 1.9 ($p< 0.01$) when compared to nonusers of HRT in this report. In comparison to other studies, this study had two differences. First, silent episodes of angina were included in

the outcome. These episodes were substantiated by the appearance of new q waves on the electrocardiogram since the last examination. Second, exposure status was updated five times during the study and the analysis was based on the last ascertainment. The overall rate of estrogen use was 24.4%. Both these issues were explored in separate analyses to make this study more similar to other studies. When angina was excluded, estrogen users between the ages of 50 to 59 years had a decreased risk of CAD which was statistically significant and there was no effect in older women. When women were re-classified as users and nonusers by a single visit, the authors found total cardiovascular mortality was decreased in estrogen users and it was not statistically significant. Therefore, analyzing HRT use based on a single ascertainment versus frequent ascertainments had a marked impact on the results. Although this study did perform a multivariate analysis and controlled for smoking, hyperlipidemia, BMI, systolic BP, alcohol use, and age, confidence intervals were not provided for the summary statistic. This study was not included in the formal meta-analysis. In summary, a total of 13 cohort studies were pooled for quantitative synthesis.

4.1.4.2 Quality assessment of cohort studies

The NOS was applied to the 17 cohort studies of healthy postmenopausal women for objective assessment of the quality of these studies. The quality scores ranged from a low of four to a maximum score of nine with two studies getting the highest score and one study getting the low score of four. Unlike the case control studies the quality scores did not improve in recent publications. The results of NOS are summarized in Table 4.1.11.

For the first component of selection of the cohort, only two studies recruited women who were representative of postmenopausal women in the community. Another five studies selected cohorts somewhat representative of the population at large. The rest of the studies recruited a select population, for example nurses or volunteers. Almost all of the studies selected the non-exposed cohort from the same community except Beard et al[180]. This study did not describe the source of the non-exposed cohort clearly. Ascertainment of exposure was ranked high in studies using complete pharmacy records

Table 4.1.11 Quality assessment of cohort studies using Newcastle Ottawa Scale

NOS criteria	Selection				Comparability		Exposure			
Study year	**Exposed group representative of community**	**Selection of non exposed**	**Secure HRT exposure records**	**Outcome not present at the start**	**Study controls for age**	**Study controls for any other risk factor**	**Blinded outcome assessment**	**Adequate follow up for outcome**	**Adequate follow up of cohort**	**Total score out of 9**
Lokkegaard 2003	-	*	-	*	*	*	*	*	-	6
Hedblad 2002	-	*	-	*	*	*	*	*	*	7
Rodriguez 2001	-	*	-	*	*	*	*	*	*	7
Grodstein 2000	-	*	-	*	*	*	*	*	*	7
Sourander 1998	-	*	*	*	*	*	*	*	-	7
Cauley 1997	*	*	*	*	*	*	*	*	*	9
Ettinger 1996	-	*	*	*	*	*	*	*	*	8
Folsom 1995	*	*	-	*	*	*	*	*	*	8
Beard 1995	*	-	*	-	*	*	*	*	-	6

Table 4.1.11 Contd. Quality assessment of cohort studies using Newcastle Ottawa Scale

NOS criteria	Selection				Comparability		Exposure			
Study year	**Exposed group representative of community**	**Selection of non exposed**	**Secure HRT exposure records**	**Outcome not present at the start**	**Study controls for age**	**Study controls for any other risk factor**	**Blinded outcome assessment**	**Adequate follow up for outcome**	**Adequate follow up of cohort**	**Total score out of 9**
Lafferty 1994	-	*	*	*	*	-	*	*	*	7
Avila Hernandez 1990	-	*	*	*	*	-	*	*	*	7
Henderson 1988	-	*	-	*	*	*	*	*	*	7
Criqui 1988	*	*	*	-	*	*	*	*	*	8
Pettiti 1987	-	*	-	-	*	-	*	*	-	4
Wilson 1985	*	*	*	*	*	*	*	*	*	9
Lauritzen 1979	-	*	*	-	*	*	-	*	-	5
Hammond 1979	*	*	*	*	-	-	*	-	-	5

or those conducting a structured interview. Interestingly, all the recent studies conducted after the year 2000 ascertained exposure by self-reports from the postmenopausal women which is considered a weaker method and does not score a star on the NOS. Almost all studies stated the cohort was either free of the outcome of interest at the beginning of the study or if women with previous heart disease were included, it was controlled for in the analysis.

The second component on the comparability of the two groups provided little variation. All studies except one controlled for age as a confounding variable. In one study, the authors controlled for age for overall cardiovascular mortality but not for the CAD outcome alone. Most of the studies controlled for a second confounding risk factor for CAD except three studies. The majority of studies controlled for several confounding variables. A detailed list of the variables controlled by cohort studies is provided in Table 4.1.9. Studies adjusting for two or more factors received the same credit and a maximum of two stars in this category.

For the third component of outcome assessment, maximum variation existed for the last item (i.e. adequacy of follow up of the cohorts). Six studies either did not provide any data on the adequacy of follow up or had an attrition rate greater than 20% over the years of follow up. The average duration of follow up of the cohorts ranged from one to 25 years with all studies but one having a follow up greater than 5 years. In almost all studies, some form of record linkage assessed the outcomes. In five of the studies independent blind assessment had been done.

The inter-rater reliability scores for these 17 studies were calculated in three ways. The Cohen's kappa coefficient was utilized for agreement on each item and averaged 0.78 (0.66, 0.90) demonstrating a high degree of agreement between the two raters beyond chance. The Cohen's kappa coefficients for individual components were 0.76 (0.45, 1.0), 0.71 (0.35, 1.0) and 0.76 (0.48, 1.0) for selection, comparability and outcome ascertainment respectively. The agreement on overall score was assessed using intraclass correlation coefficient and this was 0.82 demonstrating a high degree of agreement between raters.

4.1.4.3 Quantitative analysis of cohort studies on HRT and CAD

For the quantitative analysis, the 13 cohort studies were grouped according to exposure status (current, past, ever use) and outcome analyzed (overall CAD including fatal and nonfatal events, only fatal CAD or only nonfatal CAD) in the primary study.

Current versus never users of HRT and CAD

A total of eight out of 13 cohort studies analyzed data comparing current users with never users. Of these eight studies, four analyzed results for overall CAD, three for fatal CAD alone and one study for fatal CAD and nonfatal CAD separately. In all studies the current users were taking HRT at the initiation of the cohort. Only Grodstein et al[185] and Folsom et al[194] updated the exposure status more than once during the study. The results of these studies are summarized in Figure 4.1.14.

Two of the four studies analyzing results for overall CAD found a significant decrease in the risk of CAD among current users of HRT. The remaining two studies also found a statistically insignificant decrease in risk of overall CAD. Quantitative pooling using a fixed effects model yielded a summary relative risk of 0.60 (95% CI: 0.52, 0.70).

All four studies examining the association between current users of HRT and fatal CAD found a decreased risk of fatal CAD. This was statistically significant in three of the four studies. The results were combined using a fixed effects model and yielded a summary relative risk of 0.65 (95% CI: 0.57, 0.75). The only study to examine the association between current use of HRT and nonfatal CAD reported no association between the two variables. The same study reported a decreased risk of fatal CAD among current users. The authors hypothesized HRT may be more protective in halting the progression of pre-existing disease and hence reducing fatal events.

Statistical testing of the summary estimates is illustrated in Table 4.1.12. The results suggest a decreased risk of overall CAD and fatal CAD among current users of HRT. The Cochran's Q test had a p-value > 0.1 suggesting a lack of statistical heterogeneity among studies.

Figure 4.1.14 Risk of CAD in current users of HRT versus never users of HRT among postmenopausal women in cohort studies: relative risks and 95% confidence intervals

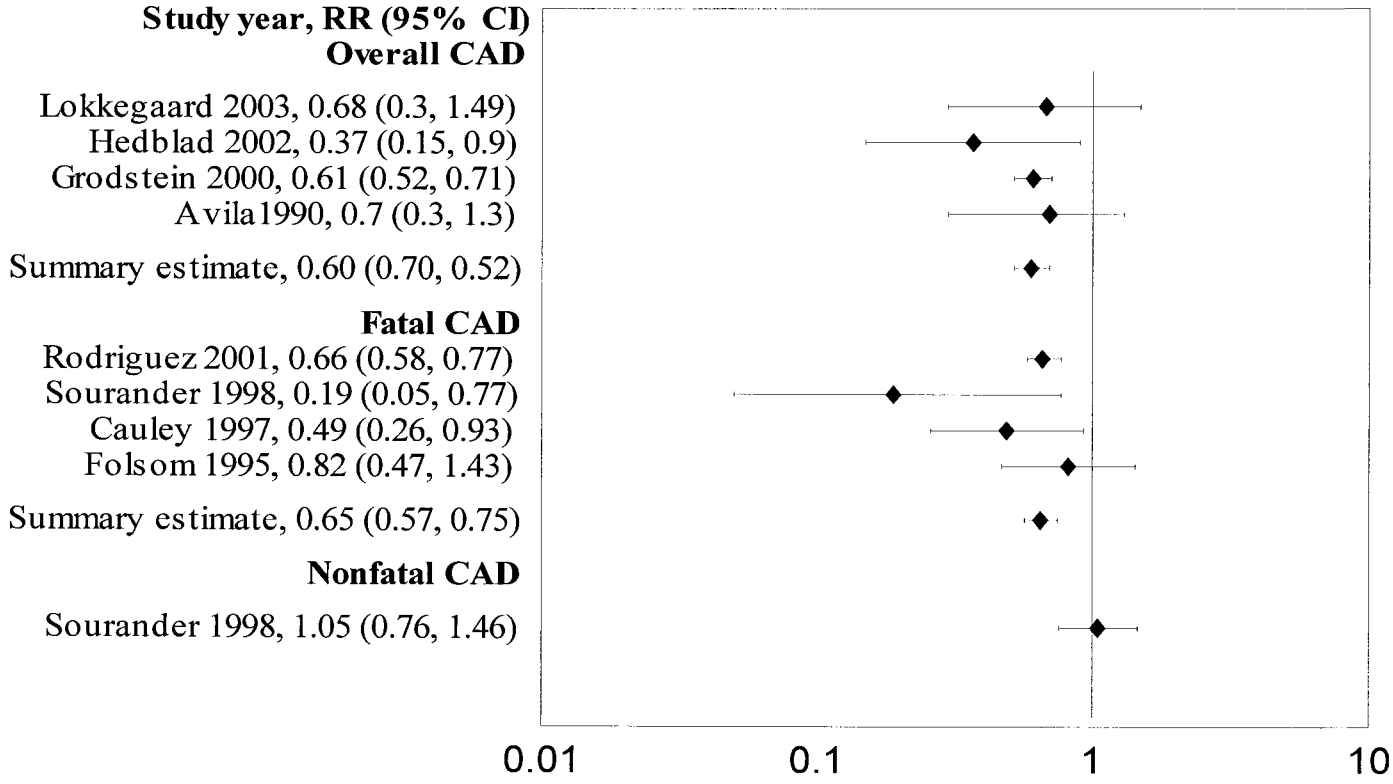
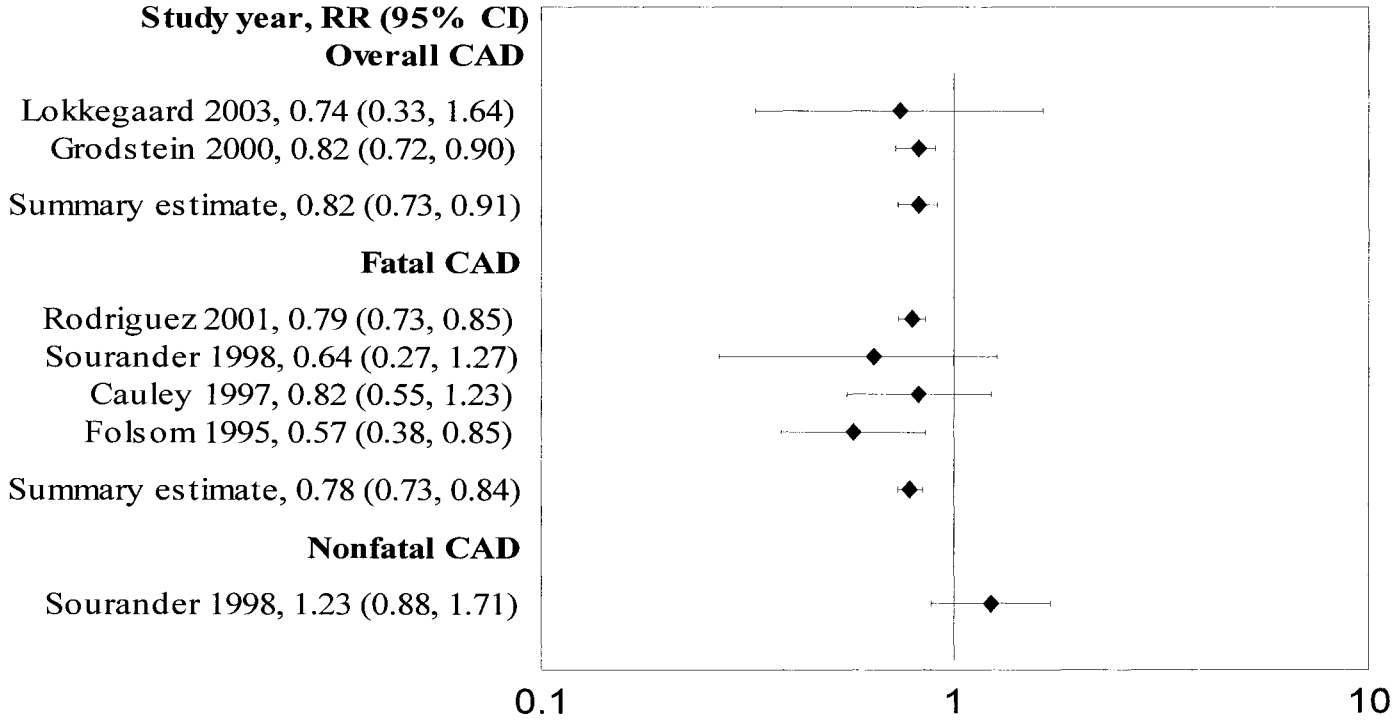


Figure 4.1.15 Risk of CAD in past users of HRT versus never users of HRT among postmenopausal women in cohort studies: relative risks and 95% confidence intervals



Past versus never users of HRT and CAD

A total of six studies out of 13 examined the association between past use of HRT and CAD. Of these, two studies analyzed data for overall CAD, three for only fatal CAD and one for fatal CAD and nonfatal CAD separately. The RRs and the 95% confidence intervals of individual studies are illustrated in Figure 4.1.15.

Both the studies analyzing data for overall CAD and past use of HRT found a decreased risk of CAD but it was significant in only one of the studies. Pooling of studies by a fixed effects model yielded a summary RR of 0.82 (95% CI: 0.73, 0.91). All of the four studies examining the association between past users of HRT and fatal CAD found hormone therapy to be cardioprotective. As shown in Figure 4.1.14, two of the studies achieved statistical significance and two did not. The summary RR calculated by a fixed effects model was 0.78 (95% CI: 0.72, 0.84). The single study evaluating past use of HRT with nonfatal CAD found an insignificant increase in the risk of heart disease.

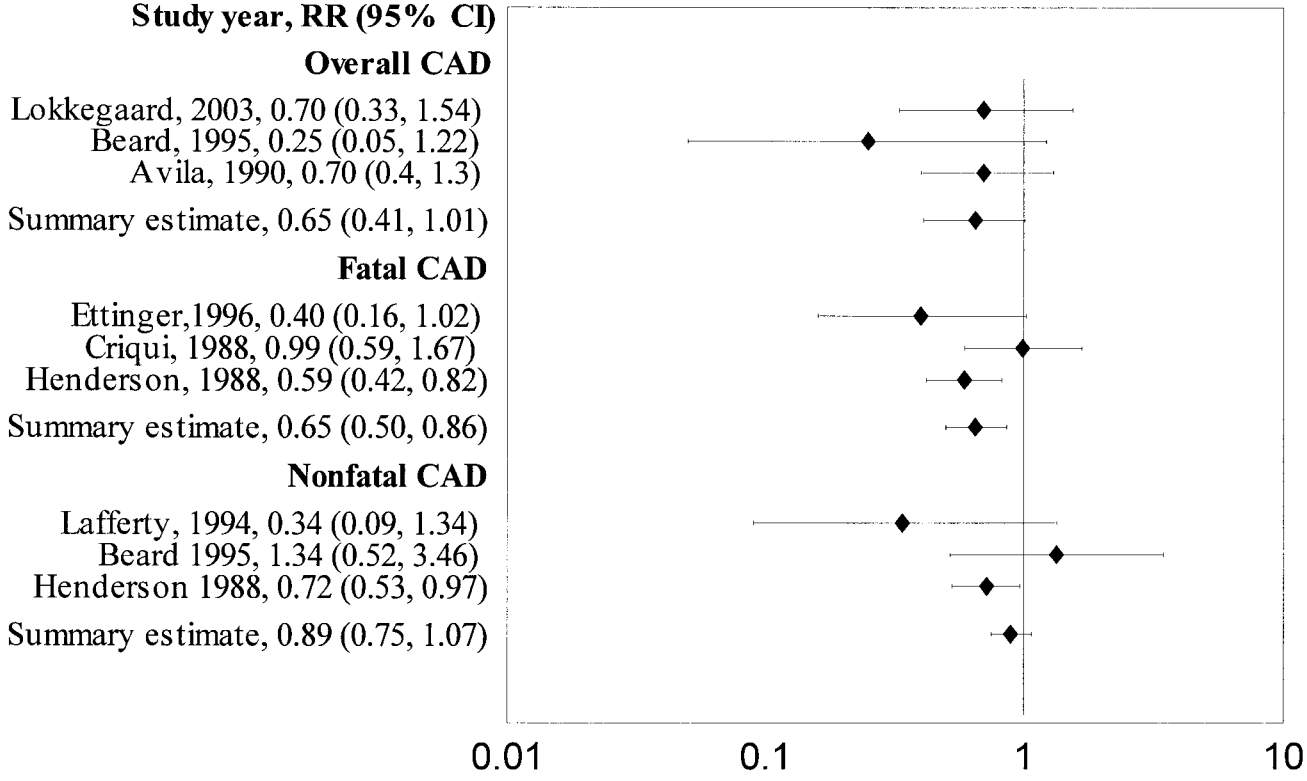
Testing of summary statistics demonstrated a decreased risk of overall CAD and fatal CAD which was statistically significant among past users. The Cochran's Q test for homogeneity had a p-value greater than 0.10 implying the estimate among the studies were not different from each other.

Ever versus never users of HRT and CAD

Seven studies evaluated the association between ever use of HRT and CAD. These are shown graphically in Figure 4.1.16. Two studies analyzed data for overall CAD, two for only fatal CAD, one for only nonfatal CAD, one study for overall CAD and nonfatal CAD separately and one presented separate results for fatal and nonfatal CAD.

All three studies analyzing data for overall CAD found an insignificant decrease in risk of overall CAD among ever users of HRT. Pooled analysis using a fixed effects model provided a RR of 0.65 (95% CI: 0.41, 1.01). Among the three studies evaluating the association between ever use of HRT with fatal CAD, all found a decreased risk but only one found it to be statistically significant. Quantitative pooling using a fixed effects model yielded a RR of 0.65 (95% CI: 0.5, 0.86). Of the three studies analyzing results for nonfatal CAD, two found a decrease in risk while one study found an increase in risk.

Figure 4.1.16 Risk of CAD in ever users of HRT versus never users of HRT among postmenopausal women in cohort studies: relative risks and 95% confidence intervals



Only one study reporting a decreased risk achieved statistical significance. Statistical pooling using a fixed effects model yielded a summary RR of 0.74 (95% CI: 0.56, 0.98).

Table 4.1.12 Results of chi-square tests of association and homogeneity for HRT users versus nonusers in cohort studies

Type of user	Number of studies	Summary OR (95% CIs)	$\chi 2$ association	$\chi 2$ homogeneity, I^2 statistic (%)
Current users				
Overall CAD	N=4	0.6 (0.52, 0.7)	43.0809 (p<0.0001)	1.2490 (p=0.7413), 0%
Fatal CAD	N=4	0.65 (0.57, 0.75)	39.435 (p<0.0001)	4.577 (p=0.2055), 34.4%
Nonfatal CAD	N=1	1.05 (0.76, 1.46)	0.0858 (p=0.7696)	Not applicable
Past users				
Overall CAD	N=2	0.82 (0.73, 0.91)	11.5803 (p=0.0007)	1.1154 (p=0.2909), 10.3%
Fatal CAD	N=4	0.78 (0.73, 0.84)	41.2509 (p<0.0001)	2.7491 (p=0.4319), 0%
Nonfatal CAD	N=1	1.23 (0.88, 1.71)	1.4921 (p=2219)	Not applicable
Ever users				
Overall CAD	N=3	0.65 (0.41, 1.01)	3.6546 (p=0.0559)	1.4700 (p=0.4795), 0%
Fatal CAD	N=3	0.65 (0.5, 0.86)	9.4209 (p=0.0021)	3.8973 (p=0.1425), 48.7%
Nonfatal CAD	N=3	0.74 (0.56, 0.98)	4.5447 (p=0.0330)	2.81296 (p=0.2450), 44.8%

Testing of summary statistics demonstrated a statistically significant decrease in risk of fatal and nonfatal CAD among ever users of HRT. This decrease was not seen for overall CAD among ever users. The Cochran's Q test for homogeneity had a p-value greater than 0.10 demonstrating the study estimates were not different from each other.

4.1.4.4 Subgroup Analysis in cohort studies

Duration of HRT

Five cohort studies evaluated the association between duration of HRT use with the risk of CAD. First, Grodstein et al[185] analyzed the risk of CAD for different durations (< 1 year, 2 to 4.9 yrs, 5 to 9.9 yrs and more than 10 years) and found a decreased risk in all users irrespective of the duration of exposure. Second, Cauley et al[191] dichotomized their data as long-term users (greater than 10 years) and short term users (1 to 9 yrs) and found HRT to be protective only in long-term users. Third, Folsom et al[194] categorized women as having used HRT for $\leq$ 5 years and > 5 years and found no association between HRT and risk of CAD for either duration. Fourth, Henderson et al divided their data on duration of exposure as greater than or less than 8 years based on the median duration of exposure. They found HRT users had a decreased risk of CAD and this did

not change with the length of duration of use. Lastly, Avila et al[164] divided their data on a yearly interval till 5 years and found no association between duration of HRT and risk of CAD. The number of women in each group were too few for any meaningful analysis. Cauley et al[191] was the only article to report longer duration of HRT use may confer protection against risk of CAD. The RRs and the 95% confidence intervals for these five studies are plotted in Figure 4.1.16. Quantitative pooling using a fixed effects model yielded a summary estimate of 0.70 (95% CI: 0.59, 0.84) for the risk of CAD in long-term HRT users and of 0.64 (95% CI: 0.45, 0.90) in short term HRT users demonstrating no difference between the two groups.

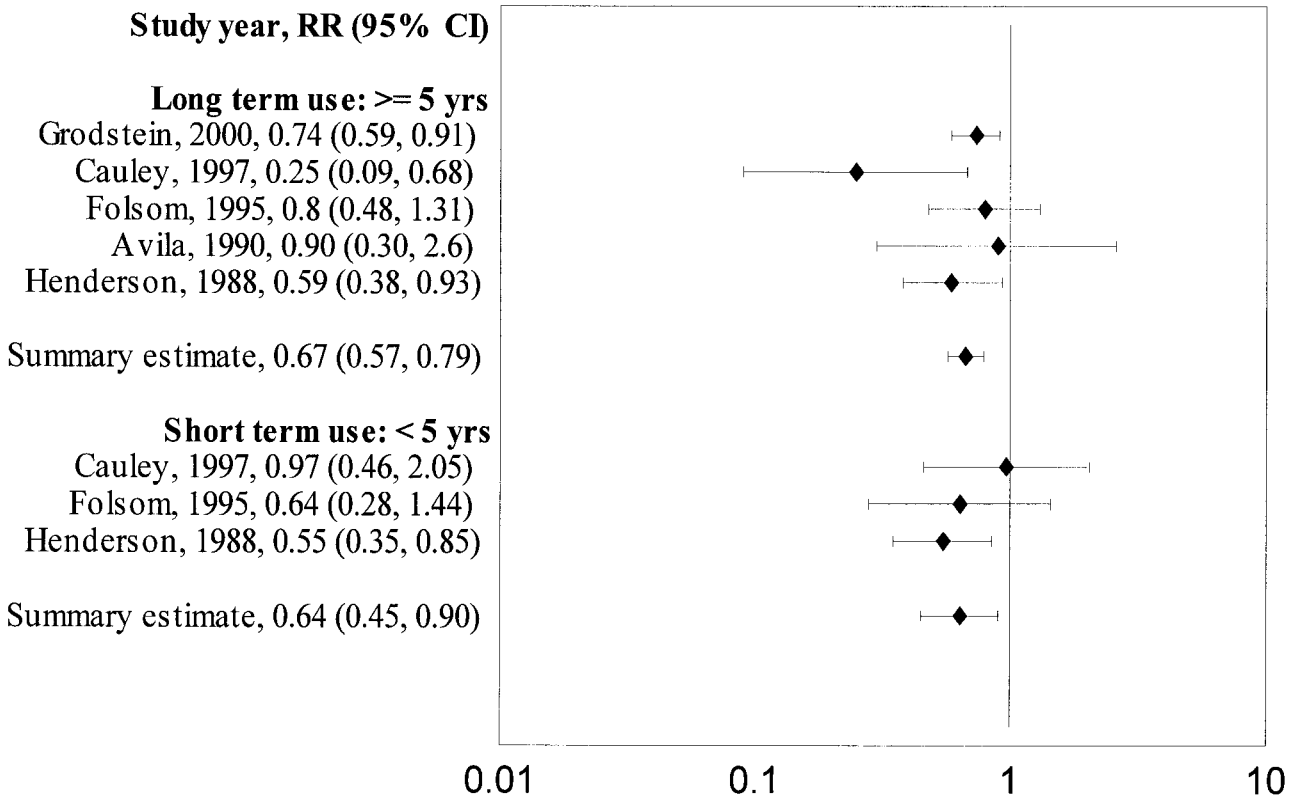
Testing of summary statistics demonstrated a decreased risk of CAD which was statistically significant for both short and long term users of HRT. The Cochran's Q test for homogeneity had a p-value greater than 0.10 and implying the study estimates were not different from each other.

Estrogen alone versus estrogen and progestin

Two cohort studies examining the relation between use of estrogen alone and the risk of CAD versus estrogen and progestin and the risk of CAD gave qualitatively opposite results. Lokkegard et al[198] found any HRT use in postmenopausal women, either as estrogen alone or as estrogen and progestin did not decrease the risk of CAD. Grodstein et al[185], on the other hand, concluded the use of either regimen was protective against CAD. The studies were compared in detail to understand the clinical heterogeneity between the two cohorts.

The geographical location of the two studies was different, one was in Denmark and the other in United States. The studies were similar on the population included, exposure ascertainment and outcome measurement. The sample size was 13,084 women in one and 70,533 in the other. In both studies, the final RRs were adjusted for major confounders. The preparation of estrogen and progestin in the two studies was different. The European study used 17 beta estradiol and NETA while the North American study used CEE and MPA.

Figure 4.1.17 Risk of CAD by duration of use of HRT among postmenopausal women in cohort studies: relative risks and 95% confidence intervals



Lokkegard et al reported a significant interaction between HRT use, the risk of CAD and presence of diabetes in postmenopausal women. They found a significantly increased relative risk (RR: 2.9; 95% CI: 1.7 to 7.9) in ever users with diabetes while the RR for ever users without diabetes had a decreased risk of CAD (RR: 0.70; 95% CI: 0.33, 1.54) which was not significant. The relative risk was further increased in current users of HRT with diabetes. These findings persisted after adjusting for smoking, hypertension, high BMI, low self-rated health and previous angina. Pooling of results obtained after excluding diabetic women did not show statistical heterogeneity. This explained the discrepancy between the results of the two studies. The authors did not report separate results for use of estrogen alone and estrogen with progestin after excluding diabetic women. Thus, statistical pooling of the results was not done.

Dose of estrogen in HRT

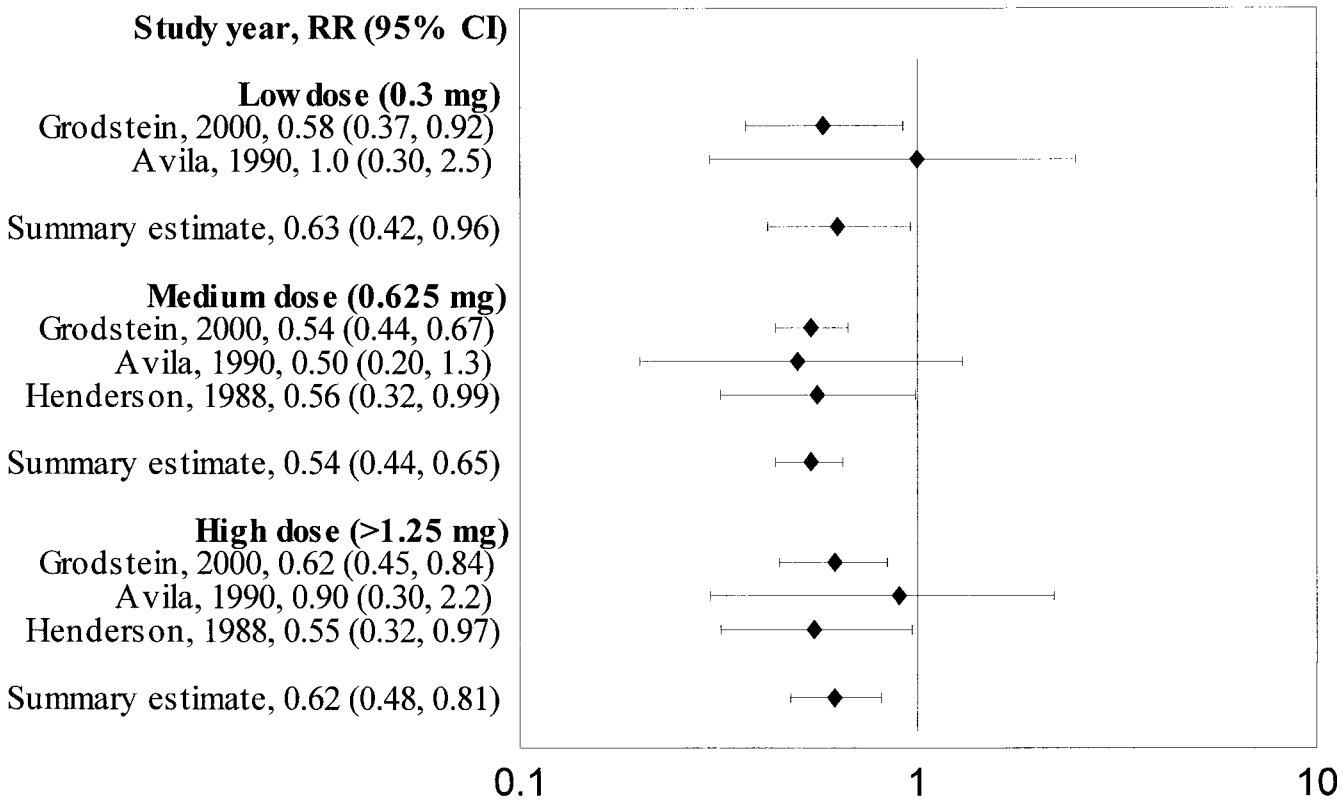
Three cohort studies examined the difference in risk of CAD with differing potency of estrogen in the HRT used. Two studies classified the potency as low (< 0.625 mg), medium (0.625 mg) and high (≥ 1.25 mg) and one classified the potency as medium or high using the same dose criteria. The results from these studies are depicted graphically in Figure 4.1.18. Two of the cohort studies reported a decreased risk in users of HRT at any potency while one study found no association between HRT and risk of CAD at low or high potency, but found a protective effect at medium potency that did not reach statistical significance. The results were pooled using a fixed effects model. The summary estimates for the low, medium and high dose were 0.63(0.42, 0.95), 0.54(0.45, 0.65) and 0.62(0.47, 0.86) respectively.

The chi square tests for association were significant for all three groups indicating a protective association between HRT use and CAD for all doses of estrogen in HRT. The Cochran's Q test for homogeneity had a p-value greater than 0.10 implying lack of statistical heterogeneity among the studies.

Effect of early initiation of HRT after menopause and CAD

Six cohort studies included young postmenopausal women (50 to 55 years) as their study population. The results from these studies would reflect the effect of initiating

Figure 4.1.18 Risk of CAD by differing estrogen dose in HRT among postmenopausal women in cohort studies: relative risks and 95% confidence intervals



HRT early after menopause. Cauley et al[191] was the only study to compare the relative risks of all cause mortality in women who initiated HRT early (within 2 years to 5 years) after menopause with those who started it 5 years after menopause. This study found both groups had a decreased all cause mortality but it achieved statistical significance among the early initiators only.

The relative risks from the six studies are plotted in Figure 4.1.19. The pooled estimate using a fixed effects model yielded a RR of 0.63 (95% CI: 0.5, 0.78) demonstrating a significant decrease in the risk of CAD among HRT users. The Cochran's Q test for homogeneity had an insignificant p-value implying lack of statistical heterogeneity among the studies (Table 4.1.13).

Table 4.1.13 Results of chi-square tests of association and homogeneity for subgroup analysis in cohort studies

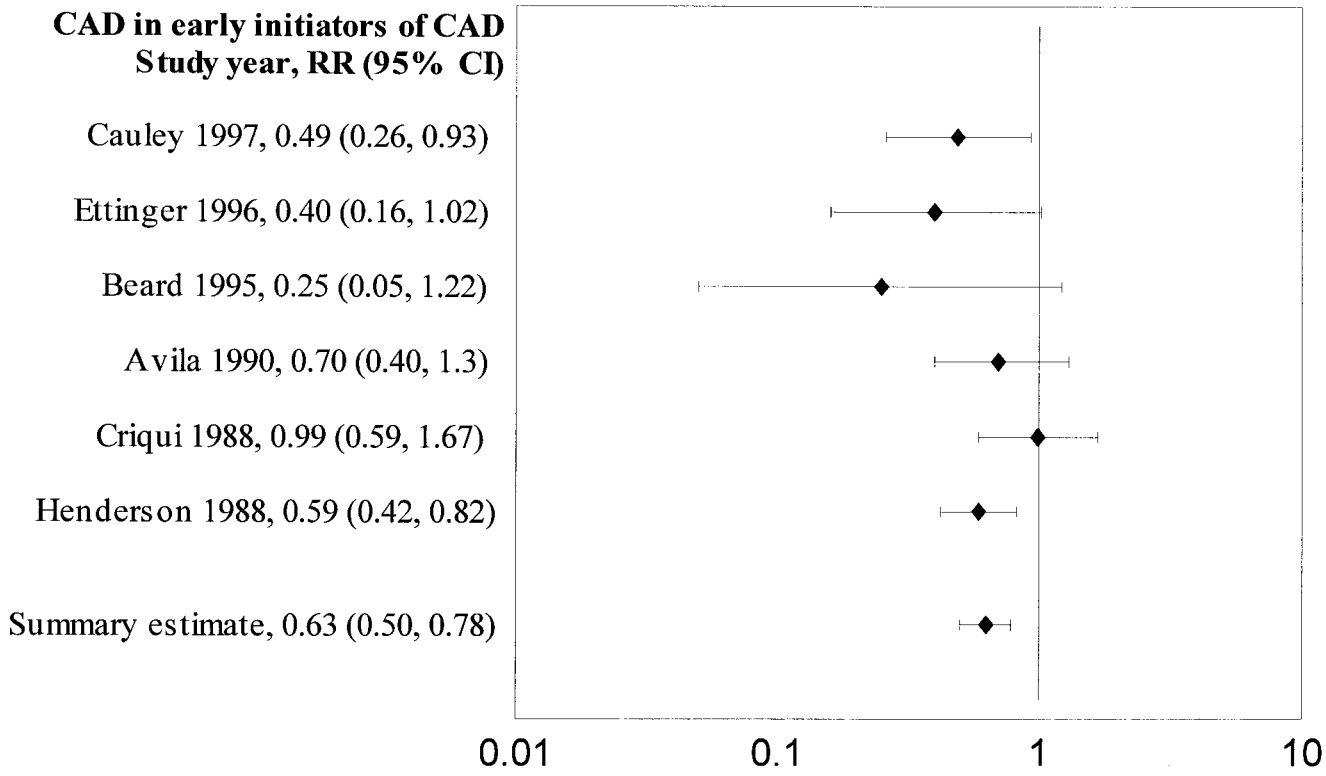
Subgroup analysis	Number of studies	Summary OR (95% CIs)	$\chi 2$ association	$\chi 2$ homogeneity, I^2 statistic (%)
Duration of HRT				
≥ 5 years	N=5	0.70 (0.59, 0.84)	15.5008 (p<0.0001)	5.2760 (p=0.2601), 24.2%
< 5 years	N=3	0.64 (0.45, 0.90)	6.4856 (p=0.0109)	1.6378 (p=0.4409), 0%
Potency of Estrogen				
Low	N=2	0.63 (0.42, 0.96)	4.6394 (p=0.0312)	0.8562 (p=0.3548), 0%
Medium	N=3	0.54 (0.44, 0.65)	39.1113 (p<0.0001)	0.0418 (p=0.9793), 0%
High	N=3	0.62 (0.48, 0.81)	12.8056 (p=0.0003)	0.7168 (p=0.3972), 0%
Early initiation of HRT after menopause	N=6	0.63 (0.5, 0.78)	16.4579 (p<0.0001)	5.9752 (p=0.3086), 16.3%

4.1.4.5 Sensitivity Analysis

Study Quality

The quality scores of the 17 cohort studies as determined by the NOS, varied between a low score of four and a maximum score of nine. One study with a score of nine and two studies with scores of five did not provide summary estimates and could not be integrated in the quantitative analysis. Summary estimates were calculated using a fixed effects model for the set of studies that achieved the same NOS score. The summary RRs are plotted in Figure 4.1.20 in order of study quality starting with the highest quality studies, to evaluate if the variation in study quality could explain some of the variation seen in the results of the studies. The size of the RRs and the width of the

Figure 4.1.19 Risk of CAD by early initiation of HRT after menopause in cohort studies: relative risks and 95% confidence intervals



confidence intervals were assessed. We could not find any trend in the size of the RR or the width of confidence intervals with the study quality. The one study with a score of nine found that users of HRT had a significant decrease in the risk of CAD. The studies with a score of eight found a much less decrease that was not statistically significant. Studies with scores of seven and six both showed a significant decreased risk while the one study with a score of four found an increase in the risk of CAD but the confidence intervals were quite wide.

We conducted a cumulative meta-analysis starting with the highest quality studies and adding in studies with a lower score. The results of this meta-analysis are illustrated in Figure 4.1.21. All the summary estimates were associated with a decreased risk of CAD among HRT users. There was little variation in the risk of CAD among HRT users with quality scores of the studies.

Besides the overall score, studies were also analyzed according to their scores on each of the three components. For cohort studies, the selection of a representative cohort as well as exposure measurement are often the most difficult to measure accurately. Summary estimates were calculated using a fixed effects model for the set of studies with the same score on the selection component of NOS. The summary RRs are plotted in Figure 4.1.22 in order of study scores starting with the highest scoring studies to evaluate if some of the variation seen in the results of the studies could be explained by differences in the scores in selection component. One study had a high score of four and one had a low score of one. The study with a score of four as well as the summary relative risk of studies with scores of three and two demonstrated a decreased risk of CAD. However, the one study with a low score of one had wide confidence intervals and showed an increase in the risk of CAD that was statistically insignificant.

Including unpublished literature

A retrospective cohort study[202] assessing postmenopausal hormone use and coronary heart disease was identified in the database on dissertation abstracts. In this study, a stratified random sample of women between the ages of 50 to 59 years was assembled from National Breast Screening Study in Ontario. Estrogen exposure was documented only once at baseline. After a follow up of 8.9 years, a questionnaire was

Figure 4.1.20 Sensitivity analysis of study quality using the NOS for risk of CAD in HRT users in cohort studies: relative risks and 95% confidence intervals

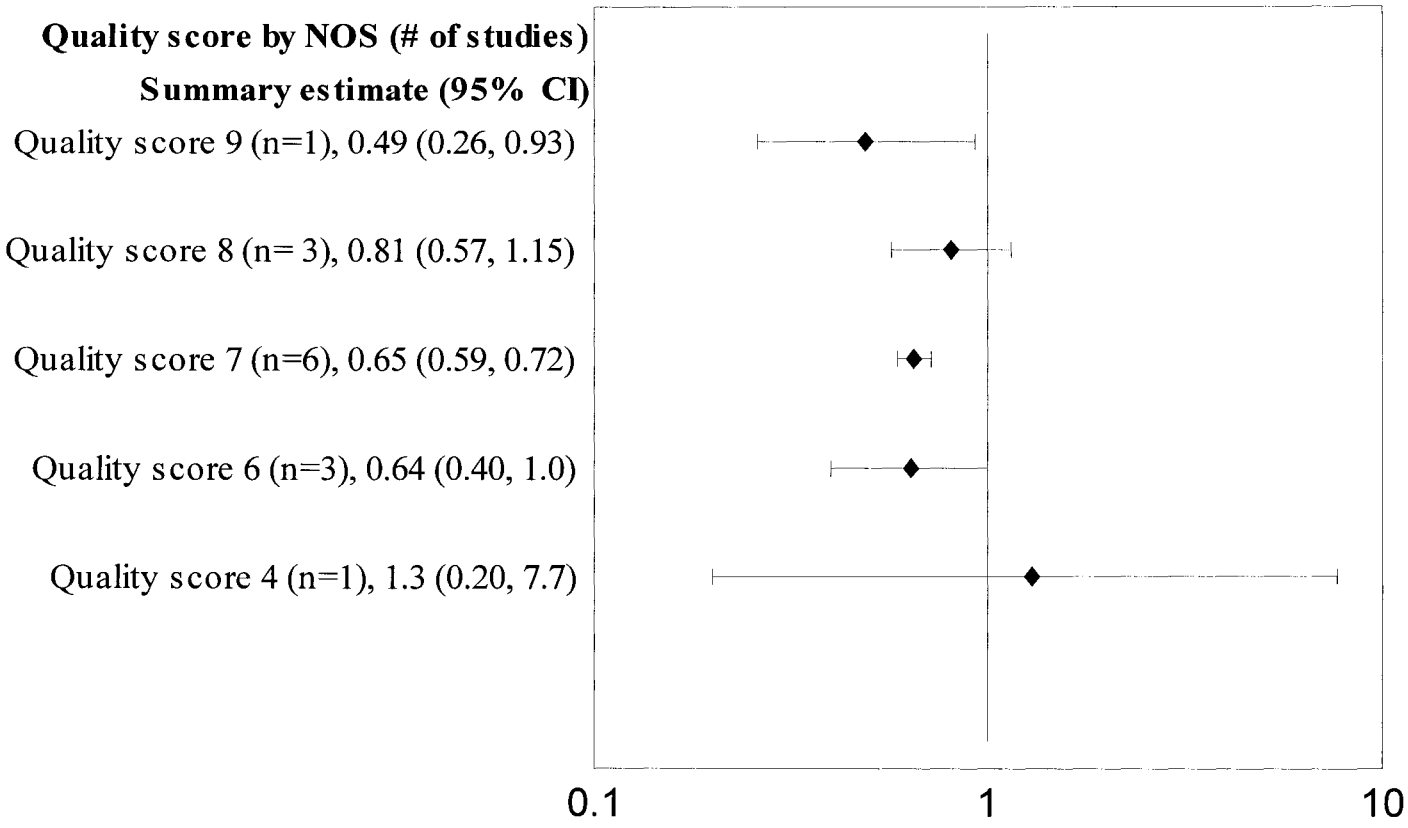


Figure 4.1.21 Cumulative* meta-analysis of cohort studies according to quality scores: relative risks and 95% confidence intervals

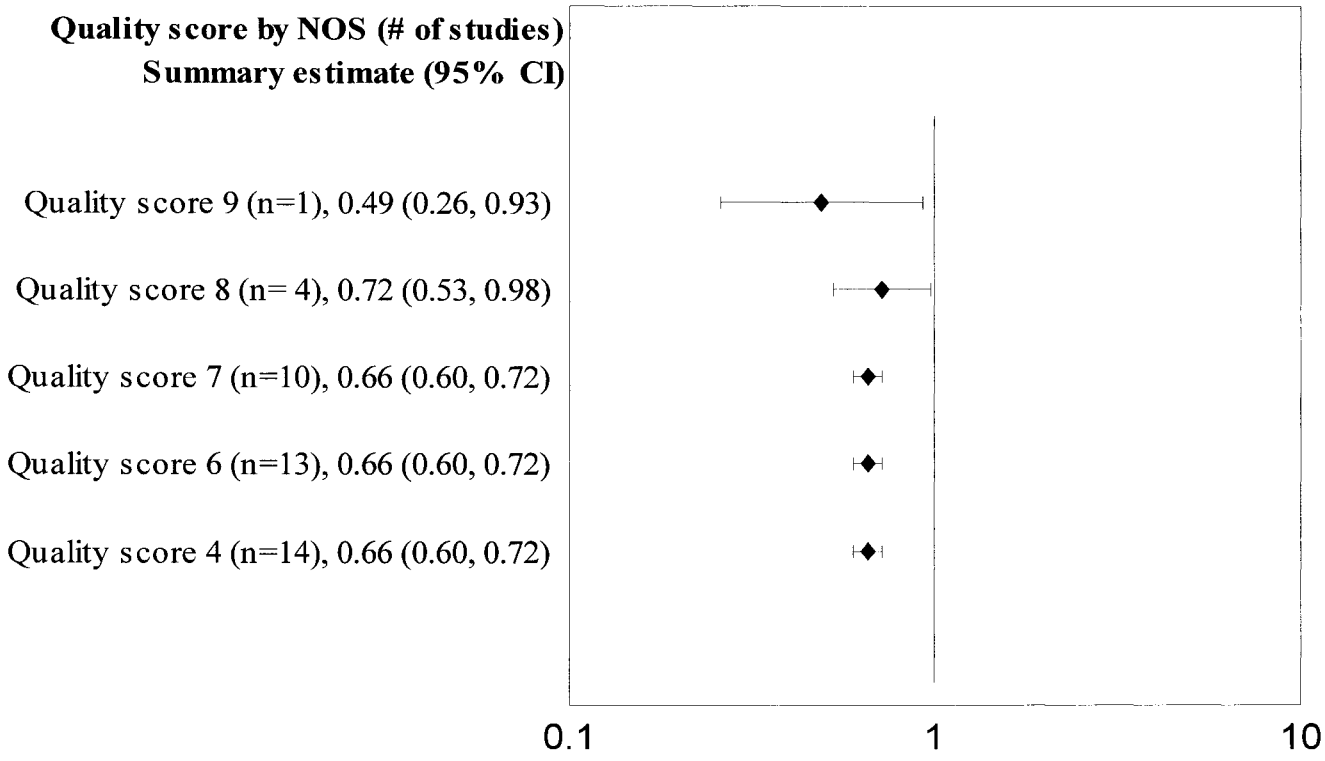
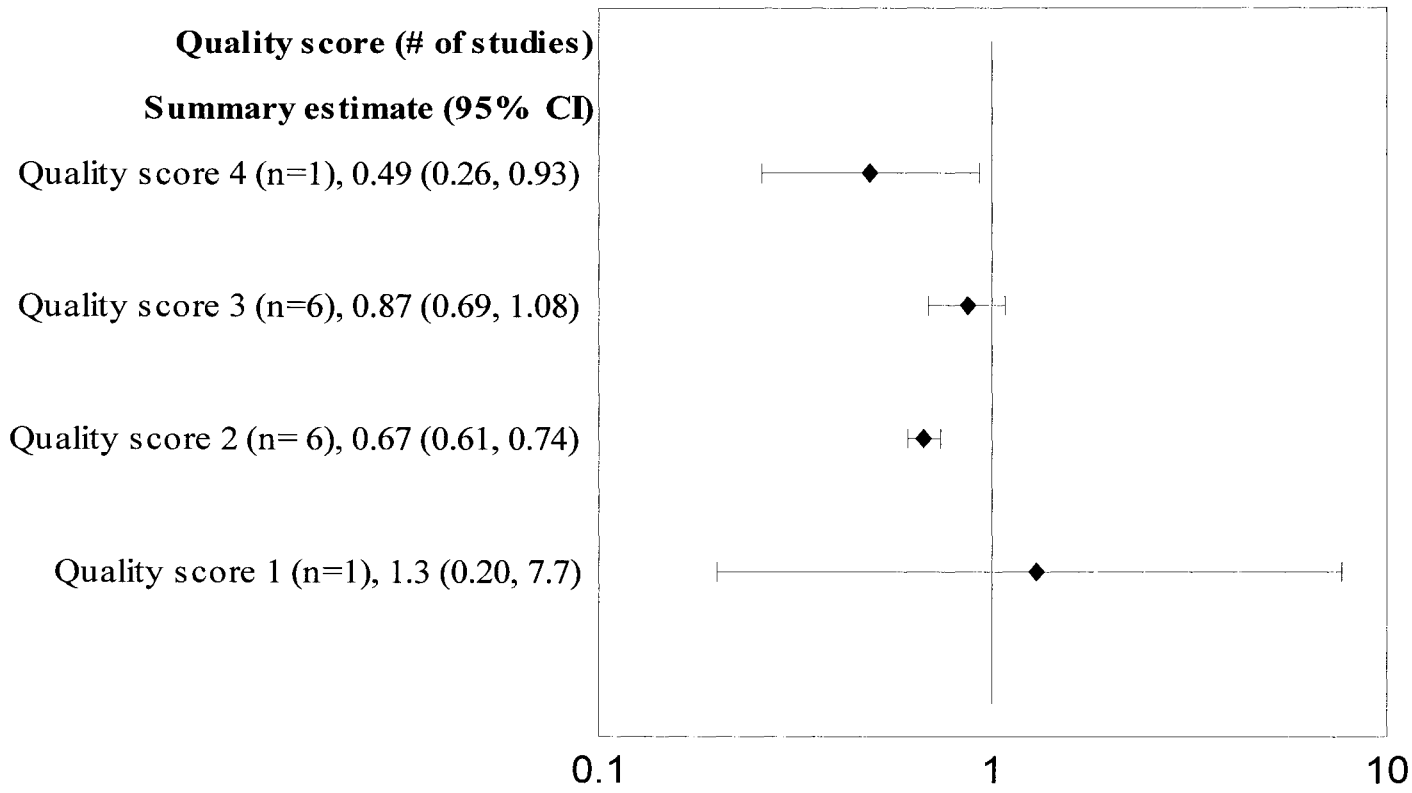


Figure 4.1.22 Sensitivity analysis of study quality for risk of CAD in HRT users in cohort studies: analysis by selection component of the NOS



sent to document nonfatal CAD events while record linkage with the Ontario and Canadian Mortality Data Base was used to document fatal CAD events. The analysis was adjusted for smoking, age and type of menopause. The authors reported an adjusted RR of 1.51 (95% CI: 0.90, 2.54) for overall CAD. Stratification for type of menopause suggested an increase in risk of CAD in postmenopausal women on HRT who achieved a natural menopause but not in those that achieved a surgical menopause.

The results from this study were pooled with other cohort studies that assessed ever use of HRT and CAD using a fixed effects model (Figure 4.1.23). A summary estimate of 0.93 (95% CI: 0.66, 1.33) was obtained. The estimate for published studies for ever use of HRT and overall CAD was 0.65 (95% CI: 0.41, 1.01). The pooled estimate for published studies showed a lower risk of CAD in ever users of HRT than the pooled estimate with inclusion of unpublished study. None of these achieved statistical significance.

Chi-square test for association demonstrated a statistically insignificant decrease in risk of overall CAD among ever users of HRT including gray literature (Table 4.1.13). The p-value for the chi-square test for homogeneity of studies was 0.06 suggesting some heterogeneity among the unpublished and published studies.

Including only those studies that adjusted for socioeconomic status

Among the cohort studies, five studies examined the association between use of HRT and risk of CAD adjusted for SES. Of these, four studies adjusted for education and one for social class. Two of the five studies reported a significant decrease in the risk of CAD with HRT use, while the remaining reported an insignificant decrease in risk of CAD. Both studies that reported a significant decrease in risk of CAD had evaluated current use of HRT. The results of these five studies are plotted in Figure 4.1.24. Statistical pooling was done using a fixed effects model. The summary estimate was found to be 0.65 (95% CIs: 0.57, 0.74) demonstrating a significant decrease in risk of CAD in study populations that initiated HRT early after menopause. The Chi-square test for association was significant (Table 4.1.13). The p-value for chi- square test for homogeneity of studies was 0.10 suggesting little variation among the studies.

Figure 4.1.23 Sensitivity analysis by including unpublished literature for risk of CAD in ever users of HRT in cohort studies: relative risks and 95% confidence intervals

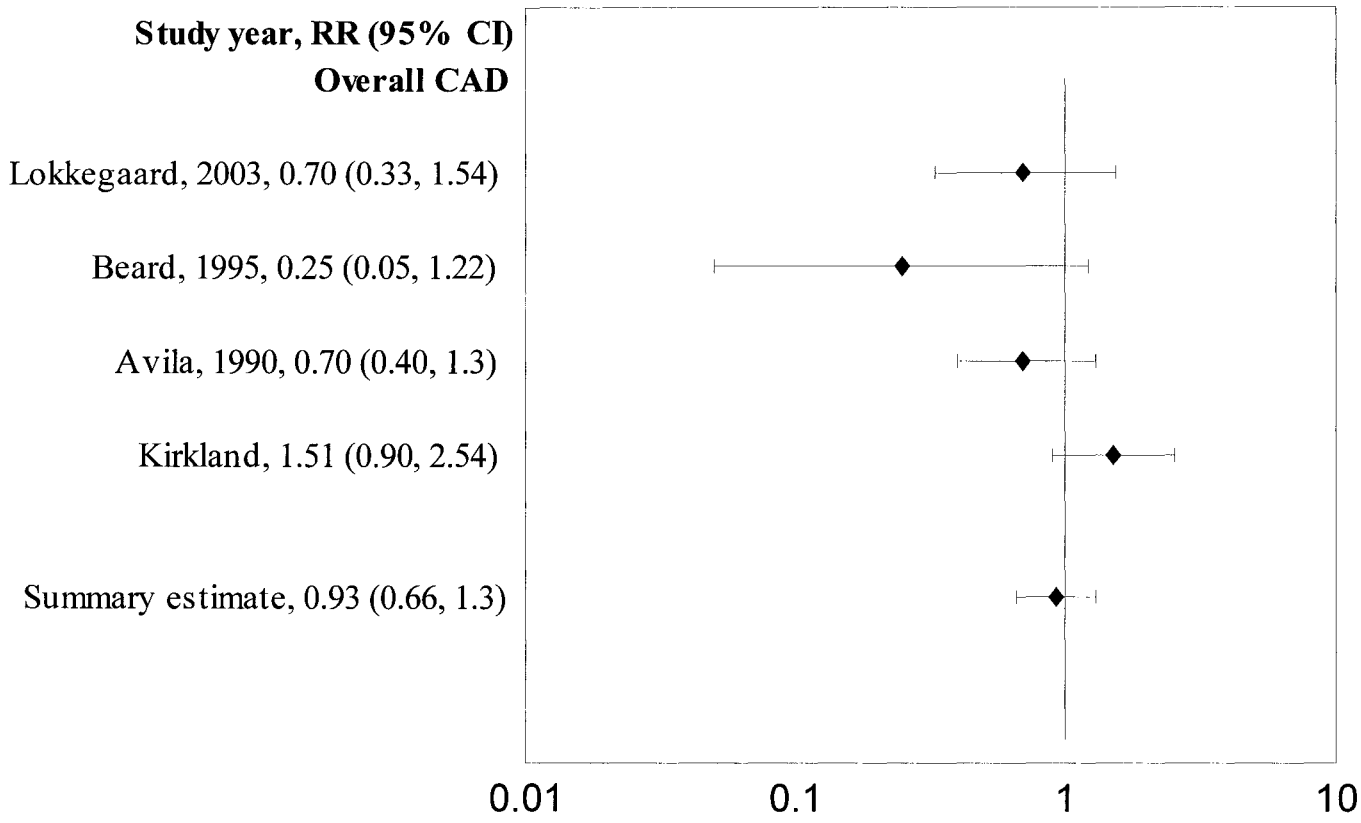


Figure 4.1.24 Sensitivity analysis of studies adjusting for socioeconomic status as a confounding variable in cohort studies: relative risks and 95% confidence intervals

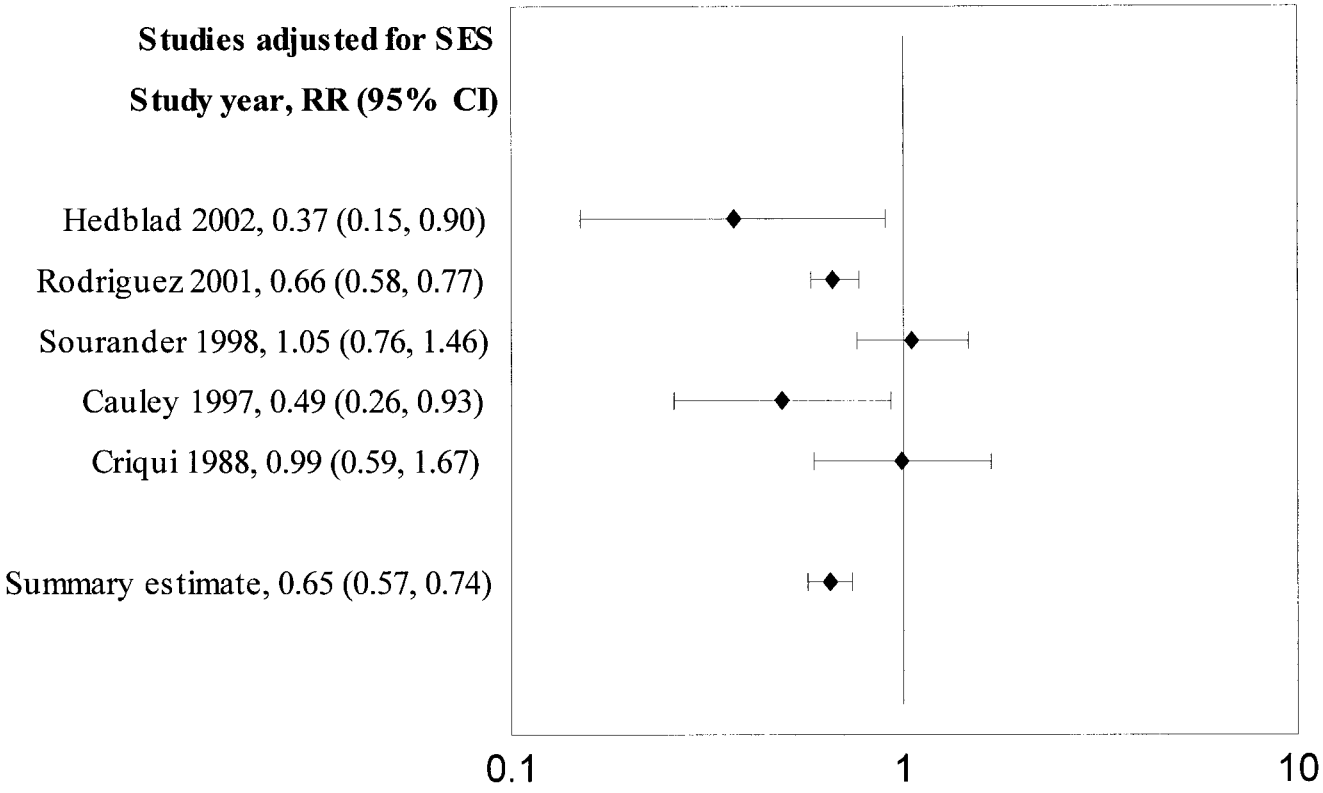


Table 4.1.14 Results of chi-square association and homogeneity for sensitivity analysis in cohort studies

Sensitivity analysis	Number of studies	Summary OR (95% CIs)	$\chi 2$ association	$\chi 2$ homogeneity, I^2 statistic (%)
Including unpublished literature	N=4	0.93 (0.66, 1.3)	0.18155 (p=0.6700)	7.3675 (p=0.0611), 59.2%
Studies that adjusted for SES	N=5	0.65 (0.57, 0.74)	40.3200 (p<0.0001)	7.737 (p=0.1017), 48.3%

4.1.5 Cohort studies on secondary prevention of CAD and HRT

4.1.5.1 Characteristics of cohort studies on secondary prevention

Seven cohort studies evaluated the association between HRT and risk of CAD in women with pre-existing heart disease. However, the study by Newton et al[203] described the same population from the Group Health Cooperative (GHC) of Puget Sound as by Heckbert et al[204]. The earlier study enrolled patients from January 1980 to December 1991 and the more recent study enrolled women from January 1986 to December 1996. The more recent study was included for the quantitative review.

All studies were conducted in the United States. The cohorts for these studies were derived in different ways. Grodstein et al[205] followed the Nurses Health Cohort to identify all women with known cardiac disease, while Alexander et al[206] identified a post MI cohort from the trial participants for the Coumadin Aspirin Reinfarction Study. Three of the other studies identified their cohorts from hospital centres that maintained comprehensive prospectively collected database on all patients attending those centres while one study was done from the GHC in Puget Sound.

The main characteristics of these studies are summarized in Table 4.1.14. Four of the six studies were prospective while the remaining two were retrospective. All studies included postmenopausal women with pre-existing CAD. Grodstein et al[205] included women with CABG, MI, revascularization procedure, or angiographic evidence of greater than 70% stenosis of a major artery. Two studies[204;206] included women who were post MI, two studies included women who underwent a percutaneous transluminal coronary balloon angioplasty (they both excluded women with a previous PTCA or

Table 4.1.15 Characteristics of cohort studies examining the association between HRT and CAD in postmenopausal women with pre-existing heart disease

Study year	Age at inception (yrs)	Outcome	Exposure status	No. of cases or deaths	No. of women or person-years	RR (95% CI)	Variables adjusted	Duration of FU
Alexander 2001	≥50 yrs	Combined CAD	Current New users	28 41	413 111	0.94 (0.75, 1.18) 1.44 (1.05, 1.99)	Age, smoking, BMI, previous event	15 mo (8-23 mo)
Heckbert 2001	67.8 yrs	Combined CAD	Users	25	686	0.96 (0.62, 1.5)	Age, DM, previous event and medication	3.5 yrs
Grodstein 2000	30-55 yrs	Combined CAD	Current	42	4997 py	0.65 (0.45, 0.95)	Age, smoking, DM, HC, HT, BMI, age at mp, FH	20 yrs
Khan 2000	52 yrs	Coronary event free survival	Users	67%	-	P=0.02	Stated that groups were similar on age, smoking, DM, HC	3 yrs
Abu-Hawala 1998	≥ 55 yrs	Combined CAD	Users	3	59	P=0.02	Age, smoking, DM, diseased vessels	22 +/- 17 mo
O'Keefe 1997	60 yrs	Combined CAD	Users	13	137	0.38 (0.19, 0.79)	Age, smoking, DM, HT, previous event	65 +/- 35 mo

DM: diabetes mellitus; HT: hypertension; FH: family history; HC: high cholesterol; mp: menopause

CABG) and one study defined their cohort as women who underwent single coronary artery stenting.

The average age of postmenopausal women was between 50 to 55 years for four of the studies suggesting an early initiation of HRT after menopause. In two studies[204;207], the average ages of women were 67.8 years and 60 years. The cohort sizes varied from only 129 women to 2489 women (Table 4.1.14). In the study by O'Keefe et al[207], out of a cohort of 2436 women, the authors recruited 137 consecutive women that had taken HRT and 200 women that had not taken any HRT matching the two groups on age, left ventricle ejection fraction, number of diseased vessels and revascularization status. In most studies the attrition rate was less than 10 %. The study by Khan et al[208] was an exception with an attrition of 22% among the users.

The details of exposure ascertainment are summarized in Table 4.1.15. Details of the method and frequency of ascertainment, classification of users, route, preparations and regimen used and duration of exposure are discussed. The ascertainment of exposure to HRT was by self-reporting in two studies, by using pharmacy/medical records in three studies and by interview in one study. Some of the studies used two of these methods simultaneously for greater accuracy. The classification of HRT exposure in the six studies was as follows. Alexander et al[206] classified their users as current if women had been on HRT at the time of initial coronary event and new if HRT was initiated after the event. Heckbert et al[204]presented results for women who initiated their HRT after the index coronary event only. Grodstein et al[205] classified users as current, past and never but have not defined it further. The remaining three studies classified women who had been on HRT at the time of their initial coronary event as users. The proportion of women who continued to be users at the end of follow up period is not indicated. Three of the six studies re-assessed the HRT use during the course of the study to minimize mis-classification bias. Oral HRT was the most common route of administration. Two studies have some women that used transdermal estrogen although the proportion was very low. None of the studies have presented a separate analysis by the route of administration of HRT.

In most studies, oral CEE or EE were used alone. Less than a third of women have used a combination of estrogen and progestin. Three studies analyzed the difference

Table 4.1.16 Ascertainment of exposure in cohort studies examining the association between CAD and HRT in postmenopausal women with pre-existing heart disease.

Author year	Ascertainment of HRT	Frequency of ascertainment	Data on dose, route, regimen, duration	Definitions of HRT use
Alexander 2001	Self reported and confirmed with source documents	Every 3 months during FU	Oral estrogen 67%, E+P 28%, topical 5%. Analysis by new/current use, regimen	New: those who started HRT after index coronary event. Current/prior: women on HRT at or within 2 years of index event
Heckbert 2001	Pharmacy records	Single ascertainment at baseline	Oral E or E+P: CEE and esterified estrogen commonly used. Analysis by recency of HRT use after index event, regimen, and type of estrogen	Users and nonusers at baseline.
Grodstein 2000	Questionnaire filled by postmenopausal women	Updated every two years	Oral HRT, 50% on CEE, 19% on CEE+ progestin, rest on estradiol or transdermal HRT, analysis by duration	Current, past and never users: not defined. Status determined according to the last questionnaire
Khan 2000	Not stated	Single ascertainment at baseline	Estrogen replacement therapy (ERT), majority of estrogen alone, no analysis on dose or duration	ERT users and nonusers at initial coronary event. No definition of users
Abu-Hawala 1998	Physicians and nurses notes, medication cards, and personal contact	Single ascertainment at baseline	ERT: no details. No analysis by dose, duration or regimen	Users and nonusers at index angioplasty.
O'Keefe 1997	Chart review, mailed questionnaire or interview	Reassessed during study by questionnaire and interview: frequency not stated.	Oral CEE 72% (mean dose 0.72 mg/d), oral estradiol 11% (mean dose 1.12 mg/d), transdermal 4%, E+P 13%	Users and nonusers: no definition

in association between risk of CAD with use of estrogen and progestin and estrogen alone. First, Alexander et al[206] reported that among current users, women who used estrogen and progestin had a lower hazard ratio (0.61, 95% CI: 0.38, 0.99) than women who used estrogen alone. The result was consistent for new users and all users. Second, Grodstein et al[205] reported a lack of any strong evidence of varying effects between the two regimens but also stated that the study had insufficient power to detect this difference. Third, Heckbert et al[204] reported an age adjusted RR for use of estrogen and progestin compared with estrogen alone for the risk of recurrent coronary event as 0.96 (95% CI: 0.41, 2.25) demonstrating a lack of any association. Heckbert et al[204] was the only study to compare the effects of conjugated estrogen with esterified estrogens and the risk of recurrent CAD and reported a RR of 1.35 (95% CI: 0.55, 3.30) demonstrating a non-significant increase in the risk of recurrent CAD with CEE preparation. None of the studies reported any analysis on the dose of estrogen in the HRT.

Three studies reported the association between duration of use of HRT and risk of recurrent CAD. First, Alexander et al[206] reported women on HRT at the time of the initial coronary event had no increase or decrease in the risk of a recurrent coronary event (RR: 0.94, 95% CI: 0.75, 1.18). In contrast, women who had initiated HRT after the first coronary event had a significant increase in risk of recurrent CAD (RR: 1.44, 95% CI: 1.05, 1.99). The follow up in this study was only 15 months after the initial coronary event and long-term effects of HRT could not be studied. Grodstein et al[205] analyzed the data on duration of exposure in two ways. In their main analysis, they examined duration of HRT use since the initial coronary event and found an increase in the risk of a second coronary event (RR 1.25, 95% CI: 0.78, 2.00) in short term users (defined as women who used HRT for less than a year). There was a significant decrease in risk of recurrent CAD in women who had been on HRT for over two years (RR 0.38, 95% CI: 0.22, 0.66). In an alternate analysis, the authors included the duration of exposure before the initial coronary event to assess the effects of women's full experience with HRT. They reported a two fold increase in risk for recurrent coronary events (RR: 2.1, 95% CI: 0.88, 5.01) in the short-term users compared with never users but the increase was not significant. The longer-term users had a significantly decreased risk of recurrent CAD (RR 0.5, 95% CI: 0.32, 0.77). Heckbert et al[204] reported a similar trend of decreasing risk over increased

duration of exposure in women who initiated HRT after their initial coronary event. They found an adjusted RR of 2.16 (95% CI: 0.94, 4.95) in women within 60 days of initiation, this decreased as duration of HRT use increased and was 0.76 (95% CI: 0.4, 2.12) among women who had been on HRT for over 365 days but none of these results were statistically significant.

There are several well-established risk factors that could potentially confound the association between HRT and CAD. Five of the six studies adjusted for some of the risk factors for CAD. These are summarized in Table 4.1.14. Khan et al[208] reported that there were no differences in their exposed and non-exposed groups on several confounding variables but did not present any formal comparisons or perform a multivariate analysis.

Four out of six studies reported a difference in baseline characteristics between HRT users and non-users. All of them found that users were less often diabetic, and two reported that users were younger and leaner than their counter-parts. One study reported that there were more smokers in the user group and one study reported that users of HRT smoked less frequently. These findings emphasize that confounding is an important parameter that needs to be adjusted for in these non-randomized studies.

Two out of six studies, Khan et al[208] and Abu-hawala et al[209] did not present their results as relative risks or confidence intervals. These could not be included in quantitative synthesis.

4.1.5.2 Quality of the cohort studies on secondary prevention

The NOS was applied to these six cohort studies for objective assessment of the quality of these studies. The quality scores ranged from five to a maximum score of nine. The two studies with scores of five and six did not present their results as relative risks. The results of NOS are summarized in Table 4.1.16.

For the first component of selection of cohort, two studies did not have cohorts representative of the general population while in the remaining four studies the cohorts were somewhat representative of the general population. All studies had selected their non-exposed cohort from the same community. The studies scoring well on the ascertainment of exposure had either used pharmacy records or conducted a structured interview. This was applicable to three out of six studies. All studies had a clear

Table 4.1.17 Quality assessment of cohort studies using Newcastle Ottawa Scale

NOS criteria	Selection				Comparability		Exposure			
Study, year	Exposed group representative of community	Selection of non exposed	Secure HRT exposure records	Outcome not present at the start	Study controls for age	Study controls for any other risk factor	Blinded outcome assessment	Adequate follow up for outcome	Adequate follow up of cohort	Total score out of 9
Alexander 2001	*	*	*	*	*	*	*	-	*	8
Heckbert 2001	*	*	*	*	*	*	*	*	*	9
Grodstein 2000	-	*	-	*	*	*	*	*	*	7
Khan 2000	*	*	*	*	-	-	-	*	*	6
Abu-Hawala 1998	*	*	-	*	*	*	-	-	*	6
O'Keefe 1997	*	*	-	*	*	*	-	*	*	7

distinction between the initial coronary event on the basis of which women were eligible to be included in the study and the occurrence of a subsequent coronary event.

The second component on the comparability of the two groups provided little variation. All, except one study, controlled for age as a confounder. Similarly, all except one study, controlled for a second risk factor for CAD. Majority of studies had controlled for several confounding variables. A detailed list of the variables that were controlled is provided in Table 4.1.14.

For the third component of outcome assessment, the maximum variation was provided by the first item, namely, outcome assessment. Three studies either did not provide any data on the method of outcome assessment for the second coronary event. The average duration of follow up ranged from 15 months to 20 years. All studies except one scored for adequacy of follow up as the attrition rate was below 20%.

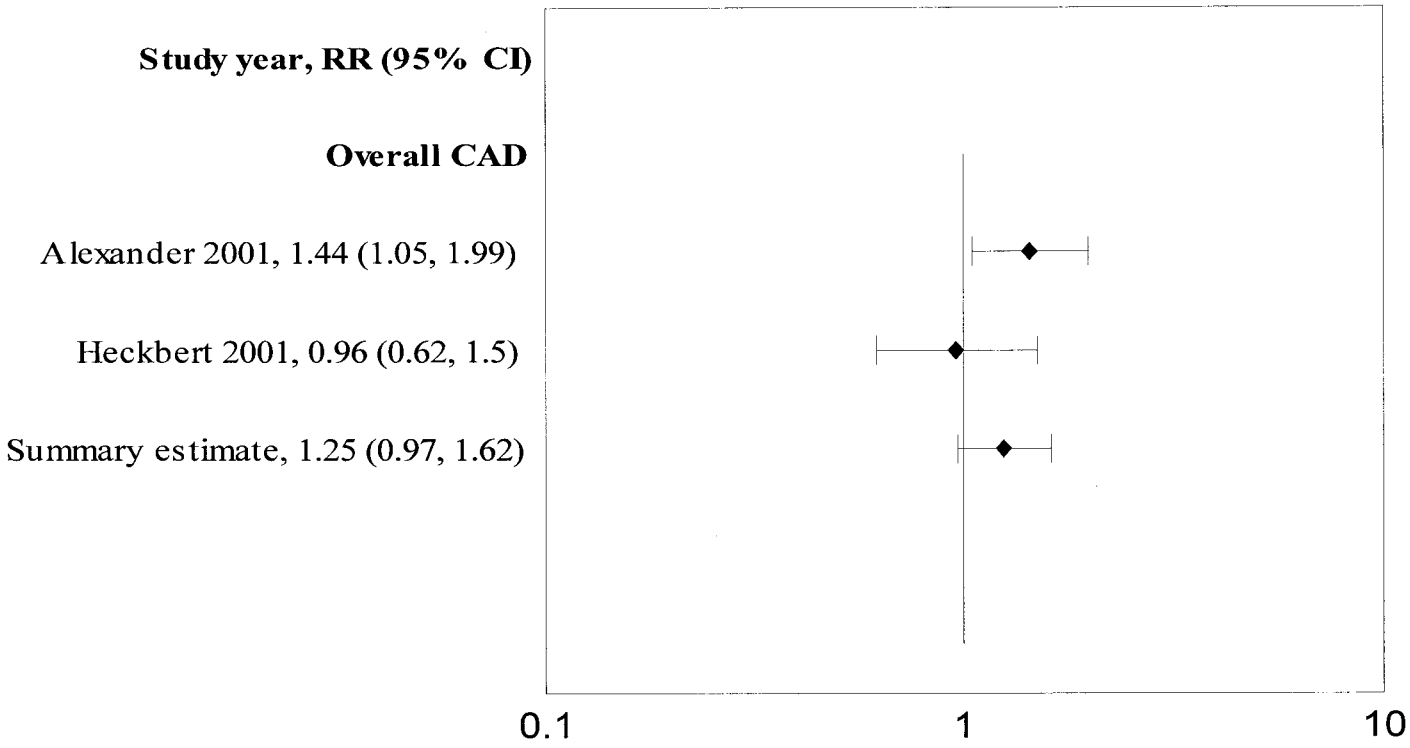
The inter-rater reliability scores for these 6 studies were calculated in three ways. The Cohen's kappa coefficient was utilized for agreement on each item and was 0.83 (0.65, 1.0) demonstrating a high degree of agreement beyond chance. The Cohen's kappa coefficients for selection, comparability and outcome ascertainment were all 1.00 showing complete agreement. The agreement on overall score was assessed using intraclass correlational coefficient and this was 0.95 also demonstrating a high degree of agreement between raters.

4.1.5.3 Quantitative synthesis of cohort studies on secondary prevention

Of the six studies on the association between HRT use and secondary prevention of CAD two did not present their results as relative risks and could not be included in quantitative synthesis. The studies were grouped according to exposure status (started after the initial MI) or had been on HRT at the time of initial MI. This approach would allow pooling of similar studies and making meaningful inferences. Since there were only four studies, separate subgroup and sensitivity analysis was not performed.

Only two studies analyzed their data for use of HRT started after the initial MI and the risk of subsequent fatal and nonfatal CAD. The results from these studies are shown in Figure 4.1.25. Alexander et al[206]reported an increase in the number of subsequent events in HRT users that was statistically significant. Heckbert et al reported

Figure 4.1.25 Risk of recurrent CAD in postmenopausal women with pre-existing CAD in cohort studies: relative risks and 95% confidence intervals



no association between HRT use and recurrent CAD. Pooled estimate of these two studies using a fixed effects model yielded a relative risk of 1.25 (95% CI: 0.97, 1.62) demonstrating a non-significant increase in number of recurrent coronary events in women starting HRT after a MI.

Three studies analyzed data for the association between women who had been on HRT at the time of their initial MI and subsequent coronary event. Of these three studies, Alexander et al was the only study to include unstable angina as one of the outcomes. The other two studies only included nonfatal and fatal MI only. Alexander et al[206] reported no association between risk of recurrent overall CAD in women who had been on HRT at the time of their initial coronary event (RR: 0.94, 95% CI: 0.75, 1.18) but a significant decrease in risk of fatal CAD (RR: 0.36, 95% CI: 0.17, 0.77). The increase in cardiac events was largely accounted for by users of HRT having significantly more UA than never users (RR: 1.11, 95% CI: 0.86, 1.43). Since the study did not report the risk of CAD excluding UA, this study was not pooled with the other two studies. Both other studies reported a statistically significant reduction in the risk of recurrent CAD (fatal and nonfatal) among women on HRT at the time of their initial MI (Figure 4.1.26). The results from these two studies were pooled using a fixed effects model and yielded a summary relative risk of 0.58 (95% CI: 0.42, 0.81) indicating a significant decrease in risk of recurrent CAD among HRT users.

Testing of summary statistics demonstrated a decreased risk of recurrent CAD which was statistically significant for users of HRT. The Cochran's Q test for homogeneity had a p-value greater than 0.10 and implying the study estimates were not different from each other.

4.1.6 Evaluating publication bias in the RCTs

A visual inspection of a funnel plot was done to assess the publication bias among the RCTs evaluating the association between HRT use and CAD in postmenopausal women. A graph was made with sample size of the studies on the vertical axis and treatment effect presented as relative risks on the horizontal axis (Figure 4.1.27). There are some pertinent points to be noted. First, none of the six trials had statistically significant results and yet all were published. Second, the two trials with small sample

Figure 4.1.26 Risk of recurrent CAD in postmenopausal women with pre-existing CAD using HRT at the time of their initial cardiac event in cohort studies: relative risks and 95% confidence intervals

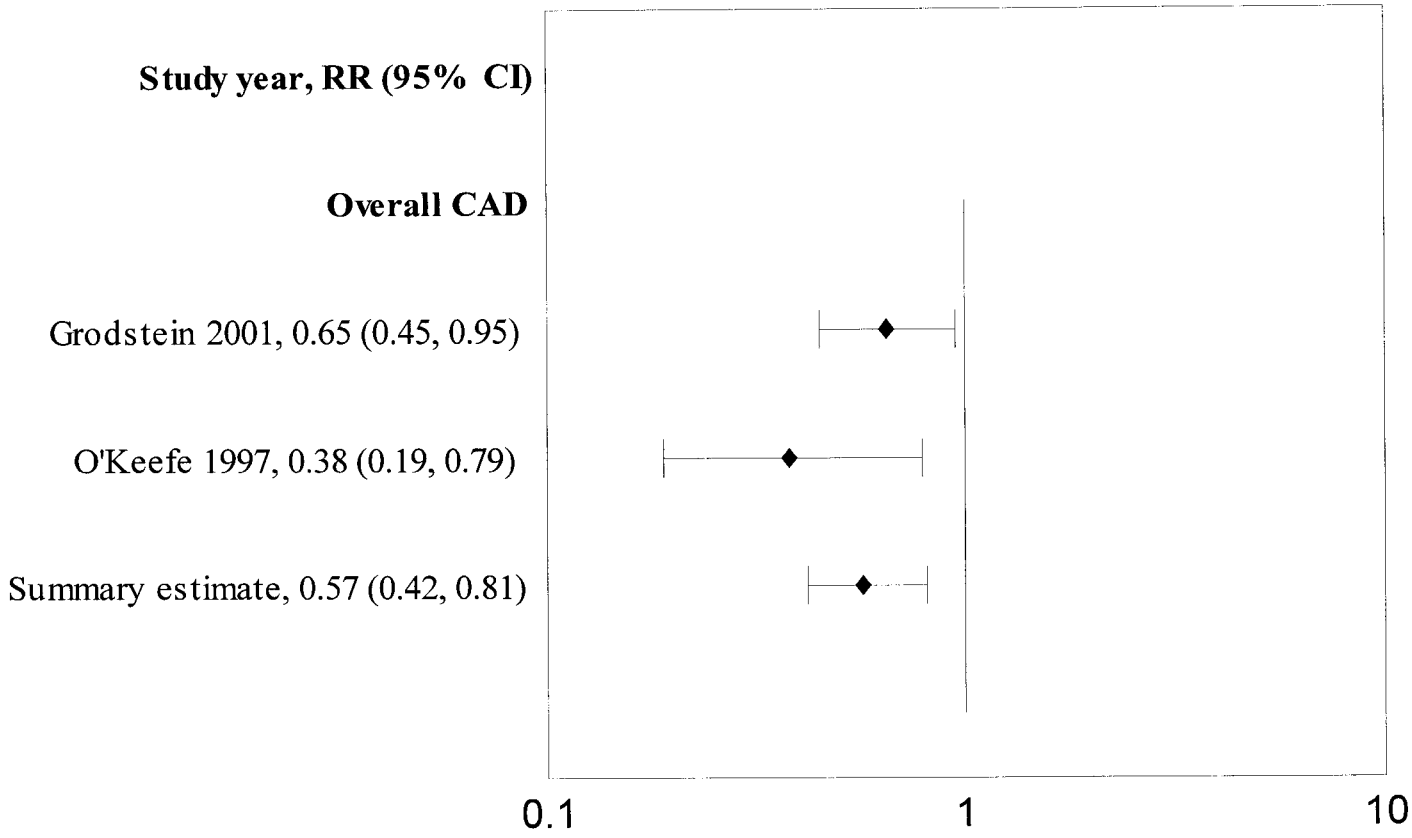
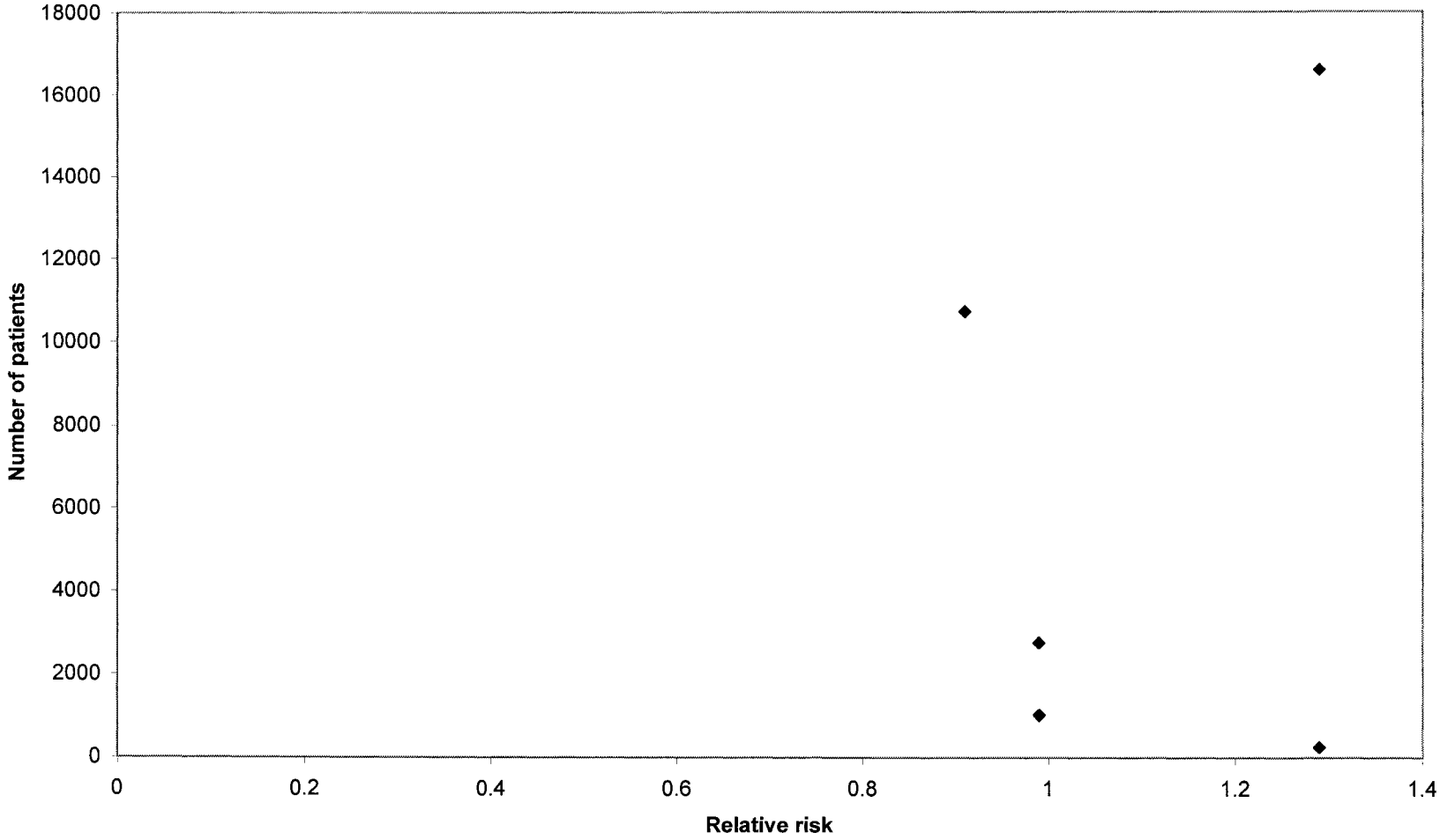


Figure 4.1.27 Evaluating publication bias in randomized trials on the association between HRT and CAD among postmenopausal women using a funnel plot



sizes did not present their results as a relative risk and could not be plotted on the funnel plot. There were too few points on the graph to visually assess for publication bias.

4.2 Evaluation of Modified Newcastle Ottawa Scale

Modifications were made only in the 'comparability' component of the original NOS. The other two components on 'selection' and 'outcome/exposure' were not altered. One of the major concerns in non-randomized studies is the bias due to differences in the groups being compared. These groups should ideally be similar on all variables other than the ones under study so that any differences found between the groups can be attributed to the effect of that treatment on the outcome of interest. The objective of modifying the comparability component of existing NOS was to improve its sensitivity in differentiating studies that achieve more comparable groups and awarding them a higher quality score. The modified items in the comparability component are summarized in Appendix 6.

The scores on the comparability component of the original and the modified NOS were compared to assess if the modified version was differentiating between the studies more efficiently (Appendix 7a and 7b). The first item on the two scales was identical and awarded a star if a study controlled for a primary confounding variable (age in this review). The second item on the original scale awarded a star if a study controlled for any secondary factor while in the modified scale, a study would get a star only if it adjusted for a group of pre-determined factors (diabetes, high cholesterol, and smoking for this review). Seven case control studies and twelve cohort studies that had scored on this item in the original NOS failed to score on the modified item. The last item on the modified scale assessed the appropriateness of the analysis in studies. Almost all studies scored a star on this item.

4.2.1 Face Validity

The modified items (Appendix 6) in the comparability component were examined to determine the relevance of these items to quality assessment of non-randomized studies as well to ensure their presentation was unambiguous.

All the three items were found to be clearly worded and addressed the question of comparability between the two groups. The first two items addressed the choice of confounding variables in a study. These would be different for different studies. The choice of these variables would be based on available literature and would be dependent on the person conducting the meta-analysis. It would be important to decide on these variables before studies are assessed for quality. In this review, age was the primary confounding variable. The secondary variables needed to be adjusted for, either in design or analysis, were diabetes, high cholesterol and smoking status. These are all well-known risk factors for CAD. The last item assesses whether the analysis conducted in a study was appropriate. This would need the input of an expert who would decide which analysis is appropriate for the question being addressed. In the present research question there are well-established risk factors for CAD which also impact upon the decision to use HRT. For the binary outcome variable of interest in this review, multivariate analysis, either logistic regression or Cox proportional hazards model adjusting for the confounding variables would be the appropriate methods of analyzing data. We found that most studies conducted appropriate analysis and this item did not provide much variability.

4.2.2 Criterion Validity

Criterion validity was assessed by studying the correlation between two quality assessment scales. The scores obtained from modified NOS were correlated with scores obtained by application of Black and Down's scale to the same set of case control and cohort studies. Spearman's rank correlation coefficient was calculated to evaluate the degree of correlation.

Black and Down scale is a scale with 27 items that can be applied to RCTs and non-randomized studies. However, there were five items that could not be applied to either case control or cohort studies (items 14, 19, 23, 24 and 27). These items assessed if study subjects were blinded to intervention, compliance to intervention, randomization, allocation concealment and calculation of power in these studies. In addition, for case control studies, two items on follow up of subjects (items 9 and 17) were not applicable. The highest score that could be achieved for case control studies and cohort studies was

22 and 24 respectively. The highest score achievable for modified NOS was 10 for both study designs.

Spearman's correlation coefficient was 0.91 ($p<0.01$) and 0.77 ($p<0.01$) for case control and cohort studies respectively demonstrating a high degree of correlation between the two scales. This implied that both scales ranked the studies on quality scores in a similar manner.

4.2.3 Inter-rater reliability

The inter-rater reliability was examined for the modified component of NOS. The score awarded for each item was compared and was found to be identical in 39 studies out of 42. In three studies, there was a discrepancy for one item, once for control of secondary variables and twice on the appropriateness of statistical analysis. In the systematic review on the effects of HRT on CAD, two raters had assessed the quality of the same set of articles using the original NOS. Therefore, the other two components were not re-assessed.

The Cohen's kappa for quantifying inter-rater agreement for all individual items was 0.94 demonstrating a high degree of agreement.

4.2.4 Measure of burden

The average time taken to evaluate the quality of a case control or cohort study by applying NOS was 8.1± 2.2 minutes. The average time to evaluate the same study by applying Black and Down's criteria was nearly twice at 15.3 ± 4.8 minutes. The time to apply Black and Down's scale was under-estimated because several items had been omitted. Thus, the measure of burden was considerably lesser when modified NOS was used for quality assessment.

Chapter 5.0
DISCUSSION

Every woman who reaches menopause will be faced with the dilemma of whether or not to take hormone replacement. The decision making will involve a risk benefit analysis for each individual woman. The association between HRT and CAD is a small part of this decision making process. We critically appraised all available literature studying this association and will now focus on questions which have been answered and questions that remain unanswered. The goal of this discussion is to analyze the relevant information provided by each of the study designs with their limitations and strengths and integrate it to enable women as well as their health care providers to make an informed choice about HRT and risk of CAD.

5.1 Study designs, their strengths and limitations

The search strategy for this review ultimately identified seven randomized trials, 20 case control studies and 24 unique cohort studies. According to the Cochrane Collaboration, the four key components of a study one must consider when attempting a meta-analyses are: population of interest, outcome, study design and interventions. The two broad categories of study designs we included were randomized trials and non-randomized studies. The results of these study designs are discussed for each of these components.

5.1.1 Randomized trials

In the hierarchy of evidence, RCTs are considered the highest level of evidence of causality. A large, well designed, properly conducted RCT is expected to yield comparison groups similar in every way except for the intervention of interest. Differences in outcome between groups are attributable to differences in the effect of the intervention. Proponents of RCTs believe if such a trial is available, results of non-randomized studies addressing the same question need not be considered[8]. Of the six trials studying the association between HRT and CAD, HERS[5] and WHI[6] fulfilled most

of the criteria of a rigorous trial, PHASE[153] and ESPRIT[154] fulfilled some of these criteria while the remaining two only met a few criteria. In the following section, these RCTs are critically analyzed. The strengths and limitations are discussed with a focus on the population of interest, intervention, outcome and study duration of these RCTs.

Population of interest

The largest proportion of women wanting to start on HRT are the young postmenopausal women usually between the ages of 45 to 55 years. In these six trials, the ages of the postmenopausal women ranged from 59 to 67 years, over a decade past menopause. The WHI arm which used unopposed estrogen was the only trial where the authors analyzed data according to three age categories (50-59, 60-69 and 70-79 years). They found that in women between 50 to 59 years of age on oral CEE, there was a decreased risk of CAD which was not statistically significant. (RR: 0.56, 95% CI: 0.30, 1.03). A third of all recruited women were in this age group. The emphasis on younger postmenopausal women is important, as primate studies have shown a delay in initiation of HRT by even 5 years after menopause fails to protect against progression of atherosclerosis but if initiated immediately after menopause, it decreased the atheroma formation. This association was also demonstrated in postmenopausal women in a cohort study by Cauley et al[191] which reported that HRT decreased the risk of CAD if initiated early after menopause but the cardio-protection was lost if HRT was given after an interval of six years post-menopause. None of the RCTs clearly address the question about the effects of hormone therapy on CAD when initiated soon after menopause.

Body mass index is a strong marker for endogenous hormones in postmenopausal women. A large cohort study[200] reported an interaction between BMI and HRT and found the coronary benefits of HRT were seen exclusively in women with lower BMIs. In both the HERS trial and the WHI trial the average BMI of women was over 27 kg/m^2. Further research is needed to evaluate the association between HRT and the risk of CAD in thin women.

Most RCTs tend to exclude patients with co-morbidities which may constitute large proportions of women being considered for HRT. These exclusions have been justified on the premise that the principal aim of most efficacy trials is to demonstrate

drug-placebo differences and patients believed to have high placebo response rates or low drug response rates (for example, presence of co-morbid conditions) should not be included. Women with conditions adversely affected by hormone therapy were excluded from these six trials as well. However, such women may be poor candidates for HRT and usually would not be recommended hormone use. This exclusion of complicated patients may decrease the generalizability of the results but it may not have a large clinical relevance in this context.

Four trials out of the six included women with pre-existing heart disease and failed to find a decrease in risk of recurrent coronary event. The two trials on relatively healthy women also failed to find any cardio-protective effect of HRT. Despite the differences in populations studied, the consistency of results emphasizes that in older postmenopausal women, with or without heart disease, HRT does not decrease the risk of CAD. This should not be extrapolated to younger postmenopausal women who would start HRT immediately after menopause.

Intervention

There are several differences in the measurement of an intervention such as HRT between randomized and non-randomized study designs. First, in an RCT, all women receive the same preparation, same route of delivery and same regimen over the duration of the trial and are all followed in a similar manner. In non-randomized studies, the physician and the participant often decide the preparation, dose and regimen of HRT prescribed. These women are all analyzed together as HRT users and are compared with those who have never used HRT. This is based on an assumption that all HRT use is equivalent, which may or may not be true. Second, non-compliance with HRT complicates the results of all studies. Among the RCTs there was a 25% to 52% drop-out rate which was often higher in the HRT group than the placebo group. In these studies, both intention to treat as well as per-protocol analysis were conducted. The non-compliant population decreases the power of the study to detect a difference but does not affect the internal validity of the study. In non-randomized studies it was assumed that women classified as HRT users at baseline remained users until the completion of the

study. This assumption may or may not be true. Thus, a non-compliance rate was neither measured nor accounted for and may affect the internal validity of the results.

RCTs are limited in their breadth by the fact they often evaluate a single preparation, dose and route of intervention. Also, most trials study a single intervention and cannot study benefit or harm from a combination of interventions. In the six trials included in this review, different oral and transdermal preparations were studied but within each trial, a single HRT preparation, dose and route was assessed.

The two parallel arms of WHI studied the two most common regimens (oral estrogen alone or in combination with progestin) of HRT administration. The results of both these arms failed to show any significant association between HRT and incidence of CAD. However, there were subtle differences: the WHI arm with the combined regimen showed an increase in coronary events in the first year and did not show any cumulative benefit on CAD. The estrogen alone arm showed an insignificant increase in CAD events in the first year but the cumulative effect over time suggested a modest benefit. Neither achieved statistical significance. The differences in the results of these two arms may be due to differences in baseline characteristics of women, their event rates, longer duration of intervention, as well as the regimen used.

Two trials evaluated the transdermal route of HRT. These two studies yielded inconsistent results with one trial showing an insignificant increase in the number of coronary events and the other showing no difference in the outcome of anginal symptoms. Neither achieved statistical significance. The PHASE[153] trial assessing transdermal patches was under-powered, unblinded and suffered from a high dropout rate, while the trial evaluating symptoms of angina[152] had only 20 women per group and was of poor quality. There is a lack of evidence of effect of transdermal HRT on CAD. An adequately powered, double blinded, placebo-controlled trial using transdermal route of HRT may be needed to study its effects on CAD and provide a definitive answer.

Despite the differences in routes, regimens and preparations, both primary and secondary prevention trials failed to show an association between HRT and CAD. This emphasizes that in older postmenopausal women with or without pre-existing heart disease, HRT does not decrease the risk of CAD.

Outcome

The studies included in this review evaluated the effect of HRT on clinical CAD, both fatal and nonfatal. Studies using surrogate endpoints for CAD were not included. Surrogate outcomes are often chosen on the surmise that they are a measure of the disease being evaluated. Studies using surrogate markers often need smaller sample sizes, and the endpoints are reached before clinical events decreasing the duration of these studies. Both of these contribute to a decreased cost of trials which is an important practical consideration. There are well known risk factors for CAD. A large number of trials have been conducted on the effect of HRT on lipid levels, coronary blood flow, blood pressure, CRP levels, pro-coagulant and prothrombotic factors. The results from several such trials concluded estrogen therapy in postmenopausal women improved the risk profile for CAD. This provided the biological plausibility of HRT decreasing CAD events. However, none of the trials evaluating HRT and clinical CAD demonstrated a decrease in coronary events.

Several other areas of research have demonstrated discordant results of an intervention on a surrogate outcome with the results of the same intervention on the disease that the surrogate outcome was supposed to measure. For example, ventricular arrhythmia correlated with sudden death and overall mortality in patients with heart disease and was used as a surrogate outcome for measuring sudden cardiac death. Treatment with anti-arrhythmic drugs successfully decreased the occurrence of ventricular arrhythmia., the surrogate outcome measure. However, the Cardiac Arrhythmia Suppression Trial (CAST)[210] showed these drugs were not effective in reducing sudden cardiac death and actually increased mortality. Using surrogate outcomes is an attractive and often a logical step before measuring clinical events but trials to measure the actual outcome of interest provide definitive answers.

Measurement of clinical coronary events was done in a standardized fashion in most of the trials. Four of the RCTs studied the combined outcome of CAD including fatal and nonfatal events. Two studies only examined nonfatal events. The consistency of results from all the trials emphasizes the results are robust.

Duration of trials

The average life expectancy of a woman is approximately 80 years of age and the average age of onset of menopause is approximately 50 years. Hormone replacement therapy is a unique treatment which could be potentially used for nearly a third of a woman's life. For this reason, it is of immense clinical significance to study the long-term effects of HRT. RCTs are not a suitable design for measuring long-term effects due to limitations of cost and feasibility. Non-compliance, losses to follow up, contamination between groups, and changes in 'standard of treatment' over time decrease the validity of trial results.

In the six included trials, the follow up period ranged from 1 to 6.8 years. Thus the long-term effects of HRT on CAD have not been determined from trial results. The HERS trial showed an initial increase in coronary events in the intervention group which decreased over time. Similarly, the estrogen alone arm of WHI showed an insignificant increase in the coronary events in the first year but the cumulative effect suggested a modest benefit with longer-term use. Since this trial was stopped before the planned date, the precision of the estimated effect is decreased. A longer intervention period may have provided stronger statistical evidence of the effects of CEE on CAD. In the estrogen and progestin arm of WHI, as well as the PHASE trial, there was a non-significant excess risk in the first year and no change over time. On the contrary, the authors of ESPRIT trial reported a small but statistically insignificant beneficial effect in the first three months among the users of estrogen. There is a need to study the association between HRT and CAD over a longer duration of time.

RCTs are unsuitable for studying adverse events for the following two reasons. First, they are generally under-powered to detect important differences in adverse event rates. Second, monitoring for adverse events is not done as thoroughly as outcome assessments particularly if adverse events occur after several years after the treatment. Both these lead to under-reporting of adverse events.

5.1.2 Non-randomized studies

Case control studies and cohort studies were the two non-randomized study designs included in this review. The description of individual studies in the earlier section

highlighted the differences between the studies on the patient populations, exposure ascertainment and outcome measurement. Besides these differences, the non-comparability between the two groups on factors other than HRT poses a threat to the internal validity of these studies. There are several biases operating in these studies and these are discussed in the following section. The pertinent information derived from these studies is also discussed.

"Healthy User Bias" in non-randomized studies

Blinding and allocation concealment are two components of randomized trials that ensure the two groups being compared differ only on the intervention. As a result of random allocation, the two groups have a similar profile of participants. This would imply all potential known confounders as well as unknown factors would be balanced in the two groups if the sample size was adequate. Thus any difference in the outcome is attributable to the intervention. One of the biggest differences between randomized and non-randomized study is the lack of random allocation in the latter study design. This implies any association found between the outcome and exposure may not be fully attributable to the exposure.

The non-comparability of groups in non-randomized studies has been explored and a " healthy user bias" has been described. This bias implies HRT users are healthier than their non-user counterparts and are therefore less likely to develop CAD. There are several components of being a "healthy user". Women who self select themselves as HRT users are associated with a higher social class. Higher social class has been shown to be inversely related to mortality due to various causes and cardiovascular mortality is one of them. "Healthy users" use the health care system more often, undergo regular check-ups for other risk factors such as lipid profile, blood pressure, diabetes etc and receive treatment early. A more subtle effect could be that physicians tend to prescribe hormones to women who are healthy. The differences in baseline characteristics between users and nonusers of HRT were described by several studies included in this investigation. Briefly, users were described to be younger, leaner, had lower blood pressures, had less chance of being diabetic, belonged to higher social class, and were physically more active.

There are several studies conducted with the objective of studying baseline differences in users and nonusers of HRT. Rodstorm et al[211] conducted a prospective population based study in Sweden to assess whether risk factor profile for cardiovascular diseases differed prior to starting HRT between women who subsequently started the hormone therapy and those who did not. Multivariate models indicated HRT users had significantly lower blood pressure, less obesity, and belonged to a higher social class before starting the therapy. Similar results were reported by Derby et al[212] who studied two cohorts randomly assembled from the communities in New England. They found estrogen users were less likely to have smoked, were more likely to get their cholesterol monitored and exercised regularly. The healthier profile would inflate the apparent benefit of treatment in non-randomized studies. Matthews et al[213] followed a cohort of premenopausal women for eight years and compared baseline characteristics of those women who started HRT after menopause with those who did not. Comparison showed the estrogen users were better educated, had higher HDL levels, greater physical activity, and consumed more alcohol. They also had lower blood pressure, apolipoprotein B, lesser weight and lower fasting insulin.

Posthuma et al[214]quantified this bias by calculating relative risk for total cancer and cardiovascular disease in the same set of non-randomized studies which showed a protective effect of HRT for cardiovascular diseases. The pooled summary estimate calculated by using a fixed effects model for all cancers was 0.83 (95% CI: 0.71, 0.96) and for CVD was 0.57 (95% CI: 0.50, 6.4). They concluded the benefit of estrogen on all cancers is unlikely to be real. Most estrogen dependent cancers would increase with hormone therapy and for the rest, the relation is not known. The protective effect of hormone therapy was thought to reflect a selection bias among hormone users who were healthier women to begin with and the lower incidence of cancer was a corollary of the same. This 'healthy cohort effect or bias' was thought to be responsible for at least 20% of the benefit previously attributed to hormone therapy.

The presence of this bias is widely accepted now and most of the non-randomized studies adjusted for several of the confounding variables. The extent to which this measure can decrease the bias is unclear. We examined all studies which adjusted for a measure of socioeconomic status. Pooled estimate from case control

studies adjusting for SES demonstrated no association between HRT use and CAD. However, the pooled estimate for cohort studies adjusting for SES showed a decreased risk of CAD with use of HRT.

Compliance bias

Compliance bias is a type of confounding which results from the fact that participants who adhere to a treatment also tend to adhere to other protective types of behaviour. Compliance to any treatment, whether it is active drug or placebo, has been associated with a lower overall risk of mortality. Compliance bias was only quantified in the context of RCTs. In particular, analyses of data in two randomized controlled trials of drug treatments for CAD addressed this issue[215;216]. In both these trials, separate analysis was conducted in the treatment group and the placebo group to study the effect of adherence to therapy. These trials reported participants who were compliant with the treatment showed a statistically significant decrease in mortality irrespective of the treatment they received. The magnitude of difference was 30% to 60%. Multivariate adjustment for socioeconomic status and other medical risk factors failed to decrease this magnitude. Separate analysis on the women participants also showed the same findings. Taken together, data from the Coronary Drug Project and the BHAT trial demonstrate the existence of 'compliance or adherence bias'. This bias remained even after statistical adjustment for a large number of pertinent variables, and across both genders.

In non-randomized studies users of HRT all qualify as being 'compliers' as they opted for a treatment. Pettiti et al[217] have discussed these findings in detail and put forth an argument that compliance bias may have contributed to the decreased risk of CAD demonstrated by the previous meta-analyses by Stampfer et al[3]and Grady et al[2].

Population

The age of postmenopausal women included in the different studies ranged between 39 years to over 70 years. Thus, non-randomized studies had a wider age range than the RCTs and some of these studies provided information on the association between HRT and CAD in younger postmenopausal women. In such women, HRT was initiated early after menopause. In our sensitivity analysis, the pooled estimate of studies

including women younger than 50 years demonstrated a fifty percent decrease in risk of CAD with the use of HRT and was highly significant.

The exclusion criteria in non-randomized studies were less stringent and the population studied was a better reflection of the population encountered by a treating physician. In case control studies, control populations were selected from those women who had no contra-indications to estrogen use. Selection of an unbiased control group is difficult to achieve and contributes to selection bias in such studies.

A small set of cohort studies were conducted only on women with pre-existing heart disease. Quantitative pooling of these non-randomized studies on recurrent CAD demonstrated a lack of cardioprotection among women who started HRT after developing CAD but not among those women who had been on HRT at the time of index cardiac event and continued to take their hormones. These contrasting results may be a reflection of the effect of early initiation of HRT in the latter group of women.

Intervention

In non-randomized studies ascertainment of HRT exposure poses several challenges. First, one has to frequently rely on the accuracy of a participant's recall on the details of exposure. In retrospective studies such as case control studies, information is collected after the outcome is known and participant is being asked to recall past events. The ability to recall may be influenced by the presence of the outcome and may bias an observed association. This is defined as recall bias. In our review, women who had suffered a coronary event may be more likely to remember the details of HRT used than those women who did not have CAD contributing to recall bias. Second, categorization of users as current, past, ever or never users was based on a single interview in the majority of studies. The non-compliance among the HRT users was neither measured nor adjusted for in the non-randomized studies. This may affect the validity of the results in these studies. Third, differences in dose, duration of exposure, regimen of HRT used and routes of HRT were analyzed in a single category of HRT use and are a source of heterogeneity among studies. These issues are addressed in more detail in the following section.

The various methods employed by non-randomized studies for ascertainment of exposure were pharmacy records, medical records, personal interviews, and self-reports on mailed questionnaires. Medical records should provide the most accurate and unbiased assessment of the exposure but is dependent on the completeness and accuracy of the records used. Studies have shown personal interviews and medical records are of comparable accuracy[218;219]and therefore, using personal interviews is a valid method of ascertainment of exposure.

The exposures were classified as current, past or ever use of HRT. There are two important factors which could affect the results of these studies. First, current users were uniformly defined as women on HRT at the time of ascertainment but the duration for which women were on HRT varied greatly. Similarly, ever users of HRT had used hormone replacement for varying lengths of time. A few studies showed discontinuation of hormones for more than three years may diminish the effect of HRT on CAD[184;191]. Thus, users of HRT were not a uniform group in different studies. Second, Criqui et al[192] found many more users discontinued HRT while the non-users remained non-users. In this article the authors state " such misclassification likely biased the relative risks toward unity". This would adversely affect the internal validity of studies and pooling data from these studies may provide inaccurate summary estimates.

Duration of use of HRT is an important consideration in studying its long-term effects on cardiovascular system. Although several non-randomized studies evaluated the association between duration of HRT and CAD, the results were inconsistent with three studies reporting a significant decrease in the risk of CAD with longer duration of exposure and six studies failing to show such an association. The duration of exposure ranged from 3 to 10 years in these studies. Three recent cohort studies[204-206] on secondary prevention of CAD found an excess risk in the first year of initiation of hormones corroborating the HERS findings. One of these studies showed a significant decrease in risk of CAD after two years of HRT use. Non-randomized studies are limited in their ability to identify clinical events occurring early after the initiation of HRT because most studies have a single contact at baseline and then at the end of the study.

The difference in study results on regimen, route of administration, and dose of estrogen was explored. Difference in associations between estrogen alone and a

combination of estrogen and progesterone is a pertinent one, as progesterone is known to attenuate some of the cardiovascular benefits of estrogen. There was no difference in the pooled summary estimates between the two regimens in the non-randomized studies. The few case control studies examining the association between transdermal route of delivery and CAD showed a decrease in the risk of CAD which was statistically insignificant. The number of women in transdermal group was small. There is a paucity of literature on transdermal HRT and its effects on CAD and requires further research. The dose of estrogen in HRT was different among studies. The pooled analysis from the cohort studies found no difference between the three potencies while a single case control study concluded only medium potency dose (0.625 mg CEE) of estrogen was associated with decreased risk of CAD. None of the systematic reviews in literature have analyzed data for dose of estrogen used. A report on the website by the US Preventive Tasks Force (USPSTF) failed to find a consistent association of CAD with dose of HRT[220].

To summarize, accurate measurement of HRT exposure is difficult in non-randomized studies and is a source of heterogeneity among them.

Outcome

Clinical coronary artery disease (including fatal and nonfatal events) was the outcome of interest. The choice of outcome had an impact on magnitude and direction of association between HRT and nonfatal and fatal CAD. Pooled estimates from both non-randomized designs (case control and cohort) showed a greater decrease in the risk of fatal CAD when compared to nonfatal CAD among HRT users. This was particularly apparent in two cohort studies[180;201] where a separate analysis was reported on fatal and nonfatal CAD within each study. Both studies reported that inclusion of nonfatal events either increased the risk of CAD or demonstrated no association between CAD and HRT use. When nonfatal events were excluded, all the three studies found a decreased risk of fatal CAD. This decrease achieved statistical significance in two out of three studies. This would imply that use of HRT is more effective against fatal CAD in comparison with nonfatal CAD. Since the patho-physiology of the disease is same in both these outcomes, it is difficult to explain these facts biologically. The sample size in each group and their exposure characteristics could be responsible for the differences in addition to choice of

outcome. One case control study[172]examined the two outcomes separately and found no association between either fatal or nonfatal CAD and use of HRT.

In this review separate analysis was conducted for studies evaluating fatal CAD and those evaluating nonfatal CAD.

5.2 Discrepancy between results of RCTs and non-randomized studies

The pooled results of RCTs and non-randomized studies were discordant in this review. All the six RCTs, whether on primary or secondary prevention, failed to show any association between HRT and CAD. In the HERS trial, there was an increase in the number of events in users in the first year which decreased with longer use. A small excess risk was also noticed in the WHI trial. All trials failed to find a decreased risk of CAD in HRT users as had been expected based on previous epidemiological data.

Among non-randomized studies (both case control and cohort), this meta-analysis found current users of HRT were at a decreased risk of CAD, the decrease being more marked for fatal CAD. The risk was not decreased consistently in ever users of HRT. Humphrey et al[118] is the only other meta-analysis which assessed fatal CAD and overall CAD separately in association with their exposure status. Their meta-analysis did not include studies on secondary prevention or any of the recent RCTs and presented combined results for the three different study designs (case control, cohorts and one RCT) together. They also found decreased risk of CAD among current users of HRT and the magnitude of decrease in risk was more for fatal CAD than overall CAD. Three of the other meta-analyses [2;3;115] addressing the association between HRT and CAD did not categorize their outcomes or their users by exposure status. All of these meta-analyses showed a statistically significant decrease in the risk of CAD in association with HRT ranging from 30% to 45%.

There are several examples of intervention evaluation where RCTs and non-randomized studies have yielded discordant results [221-227]. For example, a systematic review evaluated the effect of pelvic denervation in addition to conservative surgery on dysmenorrhoea and deep dysparonia. The pooled analyses from non-randomized studies showed a significant treatment benefit of pre-sacral neurectomy but both the RCTs failed to show any benefit[221]. In two other examples[223;224], one on the effectiveness of

pneumococcal polysaccharide vaccine in the elderly and two, on a comparison of efficacy of heparin among patients with ulcerative colitis, results of non-randomized studies were inconsistent with those of RCTs. Generally, non-randomized data showed a greater efficacy that was not supported by RCTs. Ioannidis et al[228] examined reviews which had compared RCTs with non-randomized studies. Findings from 240 RCTs and 168 non-randomized studies across 45 topics were pooled using meta-analytical techniques. Non-randomized studies often showed larger treatment effects, with a twofold variation in odds ratios in 15 topics (33%). There was a statistically significant difference between the results of the two study designs in 16% cases. On testing for heterogeneity of the results, significant heterogeneity was found in 41% of the reviews of non-randomized studies and 23% of reviews of RCTs.

There are also several examples where results from non-randomized studies and RCTs were consistent. Benson et al[229] published their results on comparison of observational studies with a control group and randomized trials. They identified 136 articles on 19 treatment areas that had been assessed by both study designs. They found that the results from these study designs produced similar results in most areas. In their study, they found only two areas out of 19 for which the pooled summary estimate of observational studies was outside the 95% confidence interval for the pooled summary estimate for RCTs. Since they had investigated diverse areas of treatment, they concluded their results could be applied to any area of treatment. Another investigation comparing the two study designs was conducted by the United Kingdom Health Technology Assessment Group[230]. They evaluated eight treatments assessed in an RCT setting as well as non-randomized studies with a control group. They found no difference in the results between non-randomized studies and RCTs in seven out of eight treatments. In one, the effect was greater in non-randomized studies. Both these studies concluded that the results of RCTs and non-randomized studies do not differ systematically from each other. Britton et al[230] suggested 'a well-designed non-randomized study is preferable to a small, poorly designed and exclusive RCT'. These results were different from those published earlier[231;232] that had found greater treatment benefits in observational studies than those in RCTs.

Discordant results between non-randomized studies and RCTs are difficult to interpret. In this review, we discussed several variables which could contribute to this discrepancy. In view of the differences between individual studies, and presence of bias that cannot be measured and adjusted for, should non-randomized studies be pooled and if so, how should their results be applied to clinical practice. These issues are discussed in the following section.

5.3 Meta-analyses of non-randomized studies

Meta-analyses or quantitative synthesis of results of individual studies has been accepted as a tool to provide answers to research questions for randomized trials. There is controversy about conducting a meta-analysis for non-randomized data. Meta-analysis of randomized trials is based on the assumption that each of the included trials are providing an unbiased estimate of the effect of an intervention. The apparent variability in the results of different studies is attributable to random variation alone. This is fundamentally different for non-randomized studies as several biases and confounding factors also contribute to the variation beyond chance. In their chapter titled " Systematic Reviews of Observational Studies", Egger et al[233] presented several examples of how quantitative syntheses of observational studies can lead to completely erroneous conclusions. The authors state there is an imminent danger that such meta-analyses could produce very precise but spurious results and recommend statistical combination of data should not be the prominent feature of systematic reviews of non-randomized studies. The focus should be on careful examination of sources of heterogeneity to gain greater insight on the possible components of the variability between studies. Shapiro et al[125] recommended quantitative synthesis of observational data should be altogether abandoned. Greenland et al[126], on the other hand, stated qualitative synthesis is not the most reliable method of summarizing data and even for non-randomized data, qualitative and quantitative pooling should be attempted.

A recent Health Technology Assessment (HTA) report by Deeks et al[8], examined whether results of non-randomized studies were more variable than results of RCTs and whether methods of correcting for selection bias introduced by lack of random allocation achieve unbiased comparison groups. They empirically generated non-randomized

studies from two large multicentre RCTs by selectively resampling trial participants according to allocated treatment, center and period. They found that results from non-randomized studies often differ from those of RCTs. Despite apparent similarity on key prognostic factors, non-randomized studies may provide unreliable results as none of the methods used currently for adjustment achieve unbiased comparison groups. They found residual confounding remains even when reliable data on confounding variables is present and in some instances, adjusted results appeared to be more biased than unadjusted results. They recommended healthcare policies should not be based on results of non-randomized studies or their meta-analyses. Since they failed to identify non-randomized studies which were free of selection bias, they recommended such studies should be undertaken only when it is unethical or unfeasible to conduct RCTs. Some of these thoughts were reiterated in "Research Design Policy Brief" designed by the Campbell Corporation[234] which addressed several methodology questions on the conduction of systematic reviews of different study designs.

The few RCTs included in this review have answered some of the questions on the effects of HRT on CAD in postmenopausal women but some key issues remain unanswered. Results from non-randomized studies were reviewed to address some of these issues such as inclusion of younger women, effect of transdermal route, effect of duration of HRT use and dose of estrogen in HRT. The results of this review were looked at with a healthy skepticism and with a thorough knowledge of the operating biases and confounding factors. We were careful not to over-interpret the findings and were cognizant that the uncertainty in the results of non-randomized studies was not adequately summarized by the confidence intervals. In this review, the main focus was a qualitative synthesis of all available data and quantitative synthesis in those groups of studies that were similar in their study population, outcome, exposure measurement and design.

5.4 Quality of Studies

Assessment of quality of studies is an important component of a systematic review. Quality of a study assesses the methodological vigor and is a measure of the internal validity of a study. Inclusion of poor quality studies tends to over-estimate the

treatment effects[7]. Most of the research on quality assessment and its relevance on treatment effect has been done for randomized trials. According to Cochrane Collaboration[139] as well as other experts[235;236], reviewers should formally assess the quality of primary trials. Juni et al[131] stated inclusion of poor quality trials can distort the results of a meta-analyses and the influence of quality should be assessed in a sensitivity analysis. They also suggested instead of using a total score, it was preferable to examine the influence of individual components of quality assessment.

Studies have shown allocation concealment, blinding of outcome and proper handling of patient attrition are some of the important components that need to be assessed for judging the quality of RCTs. A scale developed by Jadad[141] covered the latter two components and Schulz's[143] component on adequacy of allocation concealment provided a quality assessment tool for the RCTs. In this review, we found the two trials scoring poorly on quality assessment for RCTs also failed to present their results as relative risks and could not be pooled quantitatively in a formal meta-analysis. Trials published recently were generally methodologically superior; except the PHASE trial which lacked blinding and had an unclear description of allocation concealment. We identified several sources of heterogeneity among the six trials and the results of primary and secondary prevention trials were not pooled together. Individually all trials showed a lack of a protective effect between HRT and CAD. Since there were only three trials in the quantitative synthesis, separate sensitivity analysis by quality scores was not performed.

A health technology assessment report on the evaluation of non-randomized intervention studies[8] conducted a review of the content of all quality assessment tools for non-randomized studies and identified 194 tools. About half of these tools were checklists and the other half were scales. Sixty tools included at least five of the six pre-specified internal validity domains (creation of groups, blinding, soundness of information, follow up, analysis of comparability, and analysis of outcome). This report identified four core items of increased importance for assessment of quality of non-randomized studies: method of allocation, comparability among groups, identification of prognostic factors and adjusting for those factors. Only 14 of the 60 instruments covered at least three out of four core items and of these, six tools were deemed suitable for use in

systematic reviews [132;237-241]. The Newcastle Ottawa Scale was one of those six tools and was used in the current review.

In this review, results of non-randomized studies and their quality assessment scores did not show any consistent pattern. Studies were grouped according to their NOS score to differentiate high quality studies from low quality studies. Summary estimates were calculated when clinical heterogeneity was minimal. However, a pattern could not be identified, for example, high quality studies with the highest achievable score of nine showed a decreased risk of CAD with HRT, while those with scores of eight and seven showed an absence of any protective effect. Studies with a score of six again showed a protective effect. This inconsistency was present both in case controls and cohort studies. We could not find a clear association between study quality and the study results. However, we did find that summary estimates of studies with low scores on exposure ascertainment component in case control design and the selection component in cohort design were associated with wide confidence intervals and a lack of protective effect of HRT on CAD. This emphasizes that the scores of the three components of the NOS need to be considered separately.

We are not the only authors to find a lack of association between treatment effect and study quality. The HTA report by Deeks et al[8] identified 511 systematic reviews examining non-randomized studies. Of these, 69 reviews investigated the impact of quality on study results in a quantitative manner. The results of these analyses failed to show a consistent pattern in the relationship between study quality and treatment effects. Some of these reviews were confounded by a variety of study designs. We did not pool results from different study designs together to avoid confounding by design. However, lack of an association between quality and study effect may reflect other sources of variability that are unmeasured. Among the meta-analyses published on this association, two have assessed the effect of quality assessment of included studies. The one by Grady et al[2] was published in 1992 and since then, a large number of non-randomized studies have been conducted. The more recent meta-analyses by Humphrey et al[118] did quality assessment for the identified studies. This study used the criteria recommended by US Preventive Tasks Force (USPTFC)[220] for quality assessment. Instead of conducting a sensitivity analysis, these authors excluded all studies judged as poor quality by their

criteria. The stringent criteria for quality excluded all but four of the 16 unique cohort studies and nine of the 20 unique case controls studies. The authors grouped the studies according to the exposure status as current, past or ever as described in the original study. They also examined the outcomes of CAD incidence and CAD mortality separately. Despite selecting only fair or good quality studies, the authors found the studies had inconsistent results. A few studies showed a decreased risk of CAD in current users and others showed no cardioprotective effect. Research is ongoing to determine how quality assessment of non-randomized studies be incorporated into reviews and the implications of individual study quality features for interpretation of review results.

5.5 Adequacy of control of confounding variables

All quality assessment tools include the evaluation of comparability of the two groups in non-randomized studies on all confounding variables other than the one being studied. The Deeks report [8] on evaluation of non-randomized studies stated the four common techniques of dealing with variations between two groups in a non-randomized study are: first, omitting comparisons between groups that differ at baseline, second, doing stratified analyses, third, modelling by logistic regression and fourth, by using propensity score methods. In a unique experiment, these authors created a "resampling" method to generate multiple unconfounded comparisons between RCTs and historically controlled and concurrently controlled studies. They found none of the four common methods were successful in removing bias in non-randomized studies. In fact, logistic regression tended to increase the bias in some of the results. The authors found adequate adjustment for selection bias could only be made when selection depended on a single prognostic factor which can be measured accurately and included in the model.

In the present review, studies showed the decision of a woman to be a HRT user was a complex one based on several variables. Often, it was affected by the physician's recommendation and a woman's own preference. Presence of risk factors for heart disease, general health, presence of other co-morbid conditions and other subtle factors that are impossible to measure often affect this decision. Under-adjustment can be explained by omission of some confounding variables, however, the observation that adjustment may result in an increase in bias needs to be investigated. The authors stated

adjustment of those confounders in the model that are more likely to be measured inaccurately or misclassified could contribute to the increase in bias. As discussed before, HRT ascertainment is one such variable prone to inaccurate measurement and misclassification.

5.6 Publication Bias

Publication bias is said to occur when more studies with statistically significant results are published than those that fail to achieve statistical significance. In a systematic review and meta-analysis it would be easier to locate and include these published studies. Thus, only a subset of studies would be included and biased summary estimates would be obtained.

In the present investigation, it is unlikely the results were affected by publication bias. There are several reasons for this. First, questions on HRT and its risks and benefits are of great interest not only to researchers but also to consumers. It is likely that any research on this topic would be published. Second, in our own review, we found the majority of published studies did not achieve statistical significance. This was particularly applicable to the trial results, all of which failed to show a significant effect.

The few approaches described to assess publication bias pertain to RCTs. The popular method called 'trim and fill' is intuitively appealing and involves the construction of funnel plots. In this technique, a measure of sample size is plotted against the relative measure of exposure or treatment effects. The resulting graph should resemble an inverted funnel. The method is subjective and the sensitivity of this method in detecting publication bias is not well studied. A recent study by Terrin et al[242]evaluated this technique for studies with different effects and clinical heterogeneity. They found this method might inappropriately adjust for publication bias where none exists. Their results suggest funnel plot is an inappropriate method for heterogeneous meta-analyses. The funnel plot for the six RCTs in this review had too few points to visually assess the presence of a publication bias but for reasons discussed before, we do not think such a bias existed. We did not find any suitable statistical approaches for assessing the presence of publication bias in non-randomized studies.

Several steps were taken in this review to minimize publication bias. The search strategy was done without any language barriers and Embase database was searched to ensure European publications were not omitted. We identified 20 articles published in languages other than English and all had an English abstract except one. Fifteen articles out of the 19 studies were review articles, two were studying surrogate outcomes and the remaining two were cross-sectional studies. Egger et al[148] argue international papers with significant findings may choose to publish them in an English language journal while the non-significant findings might be relegated to local journals. Since we searched other languages, this bias should not affect the results of this review. Moher et al[243] examined the effects of restricting their meta-analysis to English only articles. They did not find any difference in the summary estimates on inclusion or exclusion on non-English articles. This observation would probably change according to the results of individual studies.

There are contrary opinions on the inclusion or exclusion of unpublished literature in a meta-analysis. We looked at the effects of inclusion and exclusion of unpublished literature. We only found two unpublished studies, and both of these were dissertations. Some authors believe inclusion of gray literature would contribute to bias as many of the unpublished studies have exaggerated treatment effects [244]. We found data from the two unpublished studies was similar to published studies and did not change the overall results.

5.7 Non-randomized studies and causality

Non-randomized epidemiological studies are usually the first type of studies performed for generating hypotheses. They provide an opportunity to explore associations between exposure and outcome. These studies also provide an estimate of effect sizes, a rationale to design trials, and contribute in sample size calculations. However, these type of studies do not provide direct evidence of causality. In an elegant paper on HRT and CAD by Barrett-Connor[245] the evidence from the observational data for causality has been explored. She applied some of the standard criteria accepted for causality to understand the discordant results obtained between the trials and the non-randomized studies. In an effort similar to hers, we applied the Bradford Hill criteria of

causality[246] in our work to understand which of the criteria were fulfilled and which aspects of the data were inconclusive and inadequate.

a) *Strength of the association*: This implies the magnitude of effect, for example, the odds ratio or relative risks of individual studies and 95% confidence intervals. In our review, over 50% of studies showed a cardio-protective effect that was not statistically significant. The only subset of women in whom most studies showed a significant protective effect was among current users of HRT for a combined outcome of fatal and nonfatal coronary disease or fatal disease alone. Among past and ever users, the studies had inconsistent results. The statistical significance among current users became non-significant on adjusting for socioeconomic status in the case control studies but not in the cohort studies. Two meta-analyses by Stampfer et al in 1991[3] and Grady et al[2] in 1992 who concluded " there is extensive and consistent observational evidence that estrogen use reduces risks for coronary heart disease by 35%" led to a real change in practice guidelines. The meta-analysis provided a credible strength of association which had been lacking from the results of individual studies. Although the meta-analyses stated that a fixed effects model was used despite significant heterogeneity, the enthusiasm for the HRT-CAD hypothesis was so great that these results were accepted without much question.

b) *Consistency*: Consistency refers to results of individual studies being similar. In the current review, there were several studies showing decreased risk of CAD in HRT users while others demonstrated an increased risk. Most studies attaining statistical significance showed a decreased risk in users of HRT except two studies which reported an increased risk and were statistically significant. There was a lack of consistency between the results of non-randomized studies.

c) *Specificity of the association*: This has been deemed the least useful criteria for establishing causality. Since estrogen receptors are present in almost every organ of the body, studies on estrogen and progestin would not be expected to fulfil the specificity criteria. In truth, most endogenous and exogenous exposures have different effects on various tissues and organs and this exposure was no exception.

d) *Temporality*: This criterion aims to establish that the exposure preceded outcome and was responsible for it. Atherosclerosis, which is the main underlying pathology

leading to CAD is known to start years before the clinical manifestations occur. It is extremely difficult to ascertain whether the changes were present before the start of exposure in either study design. Such studies can generate hypothesis about an exposure and outcome being correlated which needs further testing. Among most cohort studies, exposure was measured at baseline alone. The discontinuation of HRT or addition of other medications may affect the outcome and was not measured leading to misclassification bias. The two studies that repeatedly measured exposure, the Framingham study[190] and Grodstein et al[185] came up with contrary results with one demonstrating increased risk in CAD in HRT users in one and other showed a decreased risk. Both studies had statistically significant results.

e) *Biological gradient/Dose response relation*: Biological gradient or dose response relation implies a change in dose or duration of exposure would result in a corresponding change in treatment effect. This would depend on the association between exposure and outcome being harmful or protective. If there is a stepwise graded response between an exposure and an outcome, it is more likely to be causal. The strongest evidence for a graded response for HRT has been found for breast cancer. Our subgroup analysis on duration of HRT demonstrated prolonged duration decreased the risk in case control studies. Cohort studies demonstrated a decrease in risk irrespective of duration of exposure. Results of studies assessing the dose of estrogen and effect on CAD showed no difference between low, medium and high potency doses of estrogen.

f) *Plausibility*: The strongest advocate for estrogen therapy and its protective effect against CAD came from biological plausibility. There were several studies done in vitro, in vivo, and on animal models demonstrating potentially beneficial effects of estradiol on cardiac risk factors. The mechanism of action has been at the genomic and nongenomic levels. Studies have been done at molecular and cellular levels and on several intermediate surrogate markers for CAD. These were reviewed in the background section. These studies contributed greatly to the wide acceptance of the HRT-CAD hypotheses. Some of the most convincing evidence came from the PEPI[4] trial demonstrating the beneficial effects of estradiol on HDL levels and the attenuation of this benefit by addition of androgenic progestins. Intermediate markers

do not necessarily predict clinical disease and the relation between HRT and CAD seems to be an example of this fact.

g) *Coherence*: This implies whether all the observations on a research question fit the hypothesized model to form an overall coherent picture. Results of individual studies were varied with some studies reporting a cardio-protective effect while others reporting an increased risk of CAD with use of HRT. Even within the same study designs the results were incoherent. Meta-analyses of these studies demonstrated a decreased risk of CAD among current users of HRT. This was most consistent for fatal CAD. This protective effect was not present for nonfatal CAD or in past or ever users of HRT. Data from the six RCTs was more coherent. The trials were different from each other on the population studied and on HRT formulations used but all failed to show any decrease in risk of CAD among HRT users.

h) *Experiment*: Conduction of a study under controlled conditions and demonstrating a change in exposure results in a change in outcome is considered indispensable for inferring causality. Non-randomized studies do not provide such controlled conditions. In this review, the two well conducted trials failed to show any decrease in risk of CAD in women on HRT. The results have limited generalizability but for that set of women and for the preparation studied, the results provide reliable evidence of lack of any association between HRT and CAD.

i) *Analogy*: Analogy implies we are more ready to accept new arguments which resemble the previously accepted ones. This is not a strong predictor of causality as much as a reflection of whether the scientific world would believe the results of research. The relation between smoking and lung cancer was derived principally from non-randomized studies. The association was consistent, over-whelming and statistically significant. Similarly, there were several non-randomized studies that showed a decreased risk of CAD in HRT users and these were accepted as sufficient evidence. The fact that majority of these were not statistically significant was attributed to small sample sizes rather than a lack of effect.

5.8 Questions remaining unanswered

HRT has been a victim of media hype and hysteria over the last few years. From being one of the most popular drugs prescribed worldwide it has suffered a sharp decline in its popularity with the publication of the HERS and the WHI trials. In this review, an attempt was made to critically evaluate all available data on HRT and CAD. The three study designs included for the review were randomized trials, cohorts and case control studies. The main component of analysis was a quantitative synthesis of data, exploration of clinical heterogeneity and identification of questions that remain unanswered. Unanswered questions include:

1) The effect of initiating HRT immediately after menopause among women who are between 45 years to 60 years has not been studied in a randomized trial setting. The two trials qualifying as good quality trials were both conducted on older postmenopausal women in whom the HRT was initiated several years after menopause.
2) The one proper trial on primary prevention failed to provide a definitive answer about the association between CAD and HRT. Women's Health Initiative was the largest trial on the role of HRT in primary prevention of CAD. It was halted early due to the increased risk of invasive breast cancer according to a pre-established statistical criterion. Although it showed that the much-anticipated decrease in coronary risk did not occur in 5.2 years, it failed to demonstrate any association between CAD and oral HRT.
3) There is a need to study estrogen preparations other than CEE as well as the transdermal route for HRT administration and risk of CAD. A single estrogen and progestin formulation (CEE alone or with MPA) was evaluated in WHI and HERS. These findings cannot be extrapolated to any other formulation. The one trial on estradiol valerate lacked sufficient power to provide definitive clinical evidence. Similarly, both the large trials used only the oral route of delivery. The small trial on transdermal route had an inadequate sample size to begin with, had poor compliance to intervention and was halted early as instead of decreased risk, the study found an increased risk of CAD in HRT users. This study lacked sufficient power and was not helpful for decision making on this contentious issue.

4) The early harm shown by HERS trials, and then by WHI needs to be explored further. A few of the prospective cohort studies on secondary prevention also demonstrated this excess risk in the first few months of starting therapy. Several of these studies demonstrated that the excess risk does not persist after the first year of use of HRT. Many explanations have been put forth to explain this phenomenon but need further documentation.
5) None of the trials provide any data on long-term use. In women who have been on HRT for over five years with no untoward effects, there is no evidence that continued use would add new or greater risk to their wellbeing. Data from a few non-randomized studies showed that long term use of HRT was associated with a decrease in risk of CAD while some studies failed to show any additional advantage. With the present dichotomy among non-randomized and randomized data, the credibility of results from non-randomized studies alone would not be strong.

The description of non-randomized studies amply demonstrates the large differences between individual studies. Although quantitative pooling was done cautiously and there was no evidence of statistical heterogeneity, these summary estimates should be used primarily as hypothesis generating. Using these results for policy decisions is not warranted.

After reviewing data from all the present studies, one major conclusion is that oral HRT should not be initiated in older postmenopausal woman for prevention of CAD. The lack of protective effect demonstrated by the trial data would make it extremely difficult to address the unanswered questions in another randomized controlled setting. Obtaining ethical approval and recruiting adequate number of women would be a challenge. Conducting more non-randomized studies would provide data which may be biased. The decision making for recommending HRT needs to be tailored to each woman's profile of benefits and risks.

5.9 Newcastle Ottawa Scale and its modification

The Newcastle Ottawa Scale is a tool developed for quality assessment of non-randomized studies. Cochrane collaboration recommends that quality assessment of a

study should focus on how well a study controls for bias. Four types of biases were described for non-randomized studies: selection bias, performance bias, attrition bias and detection bias. The NOS addresses these biases in its three components on selection, comparability and outcome/exposure depending on the type of non-randomized study being assessed.

On application of the NOS, we found the comparability component of the scale with its two items on the choice of confounding variables was not differentiating between the studies which achieved better comparable groups from those studies which did not achieve similar groups. Other users of NOS have also stated the comparability component did not perform as well as the other two components. Since non-randomized studies do not have the advantage of random allocation, creating comparison groups similar on factors other than the one being measured is important. Studies achieving comparable groups (cases and controls or exposed and unexposed) should score higher on quality assessment than studies failing to achieve comparable comparison groups. The HTA report[130] evaluated NOS and recommended it as a useful tool for quality assessment of non-randomized studies. It also suggested an item on appropriateness of analysis would make the current tool more comprehensive. The rationale behind modifying the comparability component of NOS was to improve its sensitivity. The other two components on selection and outcome/exposure were kept unchanged.

By consulting existing quality assessment tools and literature on the types of biases and methods of minimizing them, about 15 items were generated. All items measured some aspect of comparability between the two groups under study. These focused on choice of confounding variables, on the methods used for achieving comparable groups, or on the appropriateness of analysis used for adjusting for confounding variables. A team of experts defined previously debated each item for its merits and limitations and selected three items to replace the two items in the comparability section of the original NOS. The first item in the modified scale was kept the same as in the original NOS. This item awards a star to studies adjusting for a primary confounding variable. The second item requires reviewers to agree on a group of confounding variables that all need to be adjusted. A study would receive a star only if it adjusts for all of these variables. This is a major difference from the original NOS where

a star was awarded if any confounding variable was adjusted. The last item in the comparability component was on appropriateness of analysis. The appropriate analysis for a review is decided a priori.

The modified component of the NOS was evaluated for validity and reliability. Assessment of face and content validity of the modified component of NOS showed the modifications were worded clearly and without ambiguity. The first two items addressed the choice of confounding variables. These variables would differ according to the question being addressed in a systematic review. Thorough research must be done before undertaking the quality assessment to identify appropriate confounding variables for the systematic review. For the second item, a group of relevant confounding variables need to be decided on by the reviewers. The number of confounding variables to be included is at the reviewers' discretion offering a fair degree of flexibility. The performance of the scale would be dependent on the choice of these confounding variables. This would make the NOS user-dependent. The last item on appropriateness of analysis may need input from someone well versed in biostatistics, adding complexity to this otherwise simple instrument. Overall, the face and content validity of the modified component was deemed acceptable.

The criterion validity was measured as correlational validity with another validated scale for quality assessment of studies called the Black and Down's scale[132]. This was also one of the six scales recommended by the HTA report[8] as appropriate for evaluating quality in non-randomized studies. Black and Down's scale can be applied to RCTs as well and therefore some of the items were not pertinent to non-randomized studies. The Spearman's rank correlation coefficient was high for both case control and cohort study designs demonstrating both scales appear to rank the studies similarly on quality assessment. On comparing the application of the two scales, we found the modified NOS was simpler to understand and easier to apply as there were fewer redundant items. There are other ways for evaluating criterion validity. For example, a panel of experts could give an overall assessment on a set of articles as being high or poor quality studies and this could be correlated with the scores awarded by the modified NOS. The scale needs to be validated in other areas of research as well by other methods of validation.

The inter-rater reliability of the modified component of comparability was very high as demonstrated by a large kappa coefficient. Most meta-analytic guidelines recommend two raters should perform quality assessment to minimize bias. There is a need for greater standardization of meta-analytical methods. If different meta-analyses used similar tools for quality assessment, there would be greater consistency in their comparisons. A tool with high inter-rater reliability would be highly desirable and modified NOS qualifies as one such tool.

The modified NOS took significantly less time to apply than the Black and Down's scale. On comparison between the two scales, modified NOS had fewer items which were easier to understand. A comparison of the two scales showed the rankings of studies was very similar. Since both scales are providing similar information, modified NOS is a much more practical, economical and faster instrument to use. This is of importance especially when the number of studies to be evaluated is large as in the present review.

We found a few difficulties with application of the original as well as the modified NOS. For example, in case control studies, the selection component worked well and differentiated studies with better selection of cases and controls from those studies with a selection bias (non-valid case definitions, unrepresentative cases or hospital controls). Some of the items in the exposure component did not differentiate between the studies as well. Most studies use the same methods of ascertainment and almost all studies scored a star. Similarly, many case control studies excluded women on whom exposure ascertainment data was incomplete. These studies lost a star on representativeness of cases but all scored for equal non-response rate. In the cohort studies, the component on outcome performed well differentiating those studies with blinded outcome assessment from those using self-reports. Some of the items on selection component did not perform as well, for example, selection of non-exposed cohort was almost always from the same community as exposed cohort. Exposure ascertainment at a single point was not difficult in either study design but whether HRT users remained users till the end of the study was not known. The modified component on comparability performed well for the first two items awarding a higher score to studies adjusting for pertinent confounding variables. The item on appropriateness of analysis did not perform

as well because most non-randomized studies had applied appropriate analysis and there was little variability added to the quality score. Whether this is true for other areas of research needs to be evaluated.

In summary modified NOS is a comprehensive quality assessment tool and addresses the pertinent biases in non-randomized studies. The generic version needs to be tailored appropriately to the area being researched without distorting the scale. The performance of the modified component needs further assessment in other areas of research.

References

1. The Coronary Drug Project. Findings leading to discontinuation of the 2.5-mg day estrogen group. The coronary Drug Project Research Group. JAMA 1973;226:652-57.

2. Grady D, Rubin SM, Petitti DB, Fox CS, Black D, Ettinger B *et al.* Hormone therapy to prevent disease and prolong life in postmenopausal women. Ann.Intern.Med. 1992;117:1016-37.

3. Stampfer MJ, Colditz GA. Estrogen replacement therapy and coronary heart disease: a quantitative assessment of the epidemiologic evidence. Prev.Med. 1991;20:47-63.

4. Writing Group for the PEPI trial. Effects of estrogen or estrogen/progestin regimens on heart disease risk factors in postmenopausal women. The Postmenopausal Estrogen/Progestin Interventions (PEPI) Trial. The Writing Group for the PEPI Trial. [see comments.] [erratum appears in JAMA 1995 Dec 6;274(21):1676.]. JAMA 1995;273:199-208.

5. Hulley S, Grady D, Bush T, Furberg C, Herrington D, Riggs B *et al.* Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. Heart and Estrogen/progestin Replacement Study (HERS) Research Group. [see comments.]. JAMA 1998;280:605-13.

6. Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. [see comments.]. JAMA 2002;288:321-33.

7. Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M *et al.* Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet 1998;352:609-13.

8. Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F *et al.* Evaluating non-randomised intervention studies. Health Technol.Assess. 2003;7:iii-173.

9. Notelovitz M. Estrogen replacement therapy: indications, contraindications, and agent selection. [see comments.]. [Review] [107 refs]. Am.J Obstet Gynecol 1989;161:1832-41.

10. Lobo RA. Cardiovascular disease, menopause, and the influence of hormone replacement therapy. [Review] [50 refs]. Prog.Clin Biol Res. 1989;320:313-32.

11. McKinlay SM, Bifano NL, McKinlay JB. Smoking and age at menopause in women. Ann.Intern.Med. 1985;103:350-56.

12. Pearson MBE. Coping with menopause:A guide to hormone therapy. The Canadian Journal of CME 1994;27-45.

13. The International Classification of Diseases. 9th Revision, vol. DHHS Publication No (PHS) 89-1260, Third Edition: Diseases Tabular List ed. US Department of Health and Human Services. Public Health Service, Health Care Financing Administration; 1989.

14. Evans L. Epidemiology of ischemic heart disease in women, Canadian Cardiovascular Society 2000 Consensus Conference: Women and Ischemic Heart Disease. 2000.

15. Robertson RM. Women and cardiovascular disease: the risks of misperception and the need for action. Circulation 2001;103:2318-20.

16. Mosca L, Ferris A, Fabunmi R, Robertson RM. Tracking women's awareness of heart disease: an American Heart Association national study. Circulation 2004;109:573-79.

17. Kannel WB, Abbott RD. Incidence and prognosis of myocardial infarction in women: The Framingham Study, In: Eaker ED, Packard B, Wenger NK, Clarkson TB, Tyroler HA, editors. Coronary Heart Disease in Women. Bethesda: National Institute of Health; 1987.

18. Fodor JG, Tzerovska R, Dorner T, Rieder A. Do we diagnose and treat coronary heart disease differently in men and women? Wien.Med.Wochenschr. 2004;154:423-25.

19. Lerner DJ, Kannel WB. Patterns of coronary heart disease morbidity and mortality in the sexes: a 26-year follow-up of the Framingham population. Am.Heart J. 1986;111:383-90.

20. Greenland P, Reicher-Reiss H, Goldbourt U, Behar S. In-hospital and 1-year mortality in 1,524 women after myocardial infarction. Comparison with 4,315 men. Circulation 1991;83:484-91.

21. Charney P. Presenting symptoms and diagnosis of coronary heart disease in women. J.Cardiovasc.Risk 2002;9:303-07.

22. Burke AP, Farb A, Malcom GT, Liang Y, Smialek J, Virmani R. Effect of risk factors on the mechanism of acute thrombosis and sudden coronary death in women. Circulation 1998;97:2110-16.

23. Johansen H, Nargundkar M, Nair C, Neutel I, Weilgosz A. Women and cardiovascular disease. Chronic Dis.Can. 1990;11:41-47.

24. Marmor A, Geltman EM, Schechtman K, Sobel BE, Roberts R. Recurrent myocardial infarction: clinical predictors and prognostic implications. Circulation 1982;66:415-21.

25. Loop FD, Golding LR, MacMillan JP, Cosgrove DM, Lytle BW, Sheldon WC. Coronary artery surgery in women compared with men: analyses of risks and long-term results. J.Am.Coll.Cardiol. 1983;1:383-90.

26. Khan SS, Nessim S, Gray R, Czer LS, Chaux A, Matloff J. Increased mortality of women in coronary artery bypass surgery: evidence for referral bias. Ann.Intern.Med. 1990;112:561-67.

27. Wenger NK. Gender differences in coronary risk and risk factors, In: Manson JE ea, editor. Prevention of myocardial infarction. New York: Oxford University Press; 1996. p. 387-412.

28. Yusuf S, Hawken S, Ounpuu S, Dans T, Avezum A, Lanas F *et al.* Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. Lancet 2004;364:937-52.

29. LaCroix AZ, Lang J, Scherr P, Wallace RB, Cornoni-Huntley J, Berkman L *et al.* Smoking and mortality among older men and women in three communities. N.Engl.J.Med. 1991;324:1619-25.

30. Women and smoking: A report of the surgeon general. Executive summary. MMRW 51, 1-30. 2002.
Ref Type: Report

31. Sauer WH, Berlin JA, Strom BL, Miles C, Carson JL, Kimmel SE. Cigarette yield and the risk of myocardial infarction in smokers. Arch.Intern.Med. 2002;162:300-06.

32. Craig WY, Palomaki GE, Haddow JE. Cigarette smoking and serum lipid and lipoprotein concentrations: an analysis of published data. BMJ 1989;298:784-88.

33. Willett WC, Green A, Stampfer MJ, Speizer FE, Colditz GA, Rosner B *et al.* Relative and absolute excess risks of coronary heart disease among women who smoke cigarettes. N.Engl.J.Med. 1987;317:1303-09.

34. Prescott E, Hippe M, Schnohr P, Hein HO, Vestbo J. Smoking and risk of myocardial infarction in women and men: longitudinal population study. BMJ 1998;316:1043-47.

35. Gorodeski GI. Update on cardiovascular disease in post-menopausal women [Review]. Best Pract. 2002;16:329-55.

36. Bush TL. The epidemiology of cardiovascular disease in postmenopausal women. [Review] [23 refs]. Ann.N.Y.Acad.Sci. 1990;592:263-71.

37. Manson JE, Tosteson H, Ridker PM, Satterfield S, Hebert P, O'Connor GT *et al.* The primary prevention of myocardial infarction. N.Engl.J.Med. 1992;326:1406-16.

38. Joffres MR, Hamet P, Rabkin SW, Gelskey D, Hogan K, Fodor G. Prevalence, control and awareness of high blood pressure among Canadian adults. Canadian Heart Health Surveys Research Group. CMAJ. 1992;146:1997-2005.

39. MacMahon S, Peto R, Cutler J, Collins R, Sorlie P, Neaton J *et al.* Blood pressure, stroke, and coronary heart disease. Part 1, Prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. Lancet 1990;335:765-74.

40. Chen Z, Peto R, Collins R, MacMahon S, Lu J, Li W. Serum cholesterol concentration and coronary heart disease in population with low cholesterol concentrations. BMJ 1991;303:276-82.

41. Stevenson JC, Crook D, Godsland IF. Influence of age and menopause on serum lipids and lipoproteins in healthy women. Atherosclerosis 1993;98:83-90.

42. Welty FK. Preventing clinically evident coronary heart disease in the postmenopausal woman. Menopause. 2004;11:484-94.

43. Connelly PW, MacLean DR, Horlick L, O'Connor B, Petrasovits A, Little JA. Plasma lipids and lipoproteins and the prevalence of risk for coronary heart disease in Canadian adults. Canadian Heart Health Surveys Research Group. CMAJ. 1992;146:1977-87.

44. Sacks FM, Pfeffer MA, Moye LA, Rouleau JL, Rutherford JD, Cole TG *et al.* The effect of pravastatin on coronary events after myocardial infarction in patients with average cholesterol levels. Cholesterol and Recurrent Events Trial investigators. N.Engl.J.Med. 1996;335:1001-09.

45. Baseline serum cholesterol and treatment effect in the Scandinavian Simvastatin Survival Study (4S). Lancet 1995;345:1274-75.

46. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high-risk individuals: a randomised placebo-controlled trial. Lancet 2002;360:7-22.

47. Lipsy RJ. The National Cholesterol Education Program Adult Treatment Panel III guidelines. J.Manag.Care Pharm. 2003;9:2-5.

48. Sharma AM. Obesity and cardiovascular risk. Growth Horm.IGF.Res. 2003;13 Suppl A:S10-S17.

49. Flegal KM, Carroll MD, Ogden CL, Johnson CL. Prevalence and trends in obesity among US adults, 1999-2000. JAMA 2002;288:1723-27.

50. Anderson JW, Konz EC. Obesity and disease management: effects of weight loss on comorbid conditions. Obes.Res. 2001;9 Suppl 4:326S-34S.

51. Fodor JG, Tzerovska R. Coronary heart disease: is gender important? The journal of men's health and gender 2004;1:32-37.

52. Bello N, Mosca L. Epidemiology of coronary heart disease in women. Prog.Cardiovasc.Dis. 2004;46:287-95.

53. Hu FB, Stampfer MJ, Solomon CG, Liu S, Willett WC, Speizer FE *et al.* The impact of diabetes mellitus on mortality from all causes and coronary heart disease in women: 20 years of follow-up. Arch.Intern.Med. 2001;161:1717-23.

54. Khaw KT, Wareham N, Bingham S, Luben R, Welch A, Day N. Association of hemoglobin A1c with cardiovascular disease and mortality in adults: the European prospective investigation into cancer in Norfolk. Ann.Intern.Med. 2004;141:413-20.

55. Mokdad AH, Ford ES, Bowman BA, Dietz WH, Vinicor F, Bales VS *et al.* Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. JAMA 2003;289:76-79.

56. Gotto AM. Lipid management in diabetic patients: lessons from prevention trials. Am.J.Med. 2002;112 Suppl 8A:19S-26S.

57. Tonkin AM. The metabolic syndrome(s)? Curr.Atheroscler.Rep. 2004;6:165-66.

58. Meigs JB. The metabolic syndrome. BMJ 2003;327:61-62.

59. Sallis JF, Patterson TL, Buono MJ, Nader PR. Relation of cardiovascular fitness and physical activity to cardiovascular disease risk factors in children and adults. Am.J.Epidemiol. 1988;127:933-41.

60. Wessel TR, Arant CB, Olson MB, Johnson BD, Reis SE, Sharaf BL *et al.* Relationship of physical fitness vs body mass index with coronary artery disease and cardiovascular events in women. JAMA 2004;292:1179-87.

61. Barrett-Connor E. Sex differences in coronary heart disease. Why are women so superior? The 1995 Ancel Keys Lecture. Circulation 1997;95:252-64.

62. Jacobsen BK, Knutsen SF, Fraser GE. Age at natural menopause and total mortality and mortality from ischemic heart disease: the Adventist Health Study. J.Clin.Epidemiol. 1999;52:303-07.

63. Colditz GA, Willett WC, Stampfer MJ, Rosner B, Speizer FE, Hennekens CH. A prospective study of age at menarche, parity, age at first birth, and coronary heart disease in women. Am.J.Epidemiol. 1987;126:861-70.

64. Ridker PM, Rifai N, Rose L, Buring JE, Cook NR. Comparison of C-reactive protein and low-density lipoprotein cholesterol levels in the prediction of first cardiovascular events.[comment]. N.Engl.J.Med. 2002;Online. 347:1557-65.

65. Ridker PM. High-sensitivity C-reactive protein: potential adjunct for global risk assessment in the primary prevention of cardiovascular disease. [see comments.]. Circulation 2001;103:1813-18.

66. Mendelsohn ME, Karas RH. The protective effects of estrogen on the cardiovascular system. [Review] [128 refs]. N.Engl.J.Med. 1999;340:1801-11.

67. Gilligan DM, Quyyumi AA, Cannon III RO, Johnson GB, Schenke WH. Effects of physiological levels of estrogen on coronary vasomotor function in postmenopausal women. Circulation 1994;89:2545-51.

68. Gorodeski GI, Yang T, Levy MN, Goldfarb J, Utian WH. Effects of estrogen in vivo on coronary vascular resistance in perfused rabbit hearts. Am.J.Physiol 1995;269:R1333-R1338.

69. Fischer GM, Swain ML. Effects of estradiol and progesterone on the increased synthesis of collagen in atherosclerotic rabbit aortas. Atherosclerosis 1985;54:177-85.

70. Altura BM, Altura BT. Influence of sex hormones, oral contraceptives and pregnancy on vascular muscle and its reactivity, In: Carrier O, Shibata S, editors. Factors Influencing Vascular Reactivity. Tokyo: Ygaku-Shoin; 1977. p. 221-54.

71. Lobo RA. Cardiovascular implications of estrogen replacement therapy. Obstet.Gynecol. 1990;75:18S-25S.

72. Proudler AJ, Ahmed AI, Crook D, Fogelman I, Rymer JM, Stevenson JC. Hormone replacement therapy and serum angiotensin-converting-enzyme activity in postmenopausal women. Lancet 1995;346:89-90.

73. Mendelsohn ME. Genomic and nongenomic effects of estrogen in the vasculature. [Review] [17 refs]. Am.J.Cardiol. 2002;90:3F-6F.

74. Delyani JA, Murohara T, Nossuli TO, Lefer AM. Protection from myocardial reperfusion injury by acute administration of 17 beta-estradiol. J.Mol.Cell Cardiol. 1996;28:1001-08.

75. Hale SL, Birnbaum Y, Kloner RA. beta-Estradiol, but not alpha-estradiol, reduced myocardial necrosis in rabbits after ischemia and reperfusion. Am.Heart J. 1996;132:258-62.

76. Clarkson TB, Anthony MS, Wagner JD. A comparison of tibolone and conjugated equine estrogens effects on coronary artery atherosclerosis and bone density of postmenopausal monkeys. J Clin Endocrinol.Metab 2001;86:5396-404.

77. Mikkola TS, Clarkson TB. Estrogen replacement therapy, atherosclerosis, and vascular function. [Review] [143 refs]. Cardiovasc.Res. 2002;53:605-19.

78. Rosano GM, Sarrel PM, Poole-Wilson PA, Collins P. Beneficial effect of oestrogen on exercise-induced myocardial ischaemia in women with coronary artery disease. [see comments.]. Lancet 1993;342:133-36.

79. O'Connell MB. Pharmacokinetic and pharmacologic variation between different estrogen products. J.Clin.Pharmacol. 1995;35:18S-24S.

80. Stevenson JC, Oladipo A, Manassiev N, Whitehead MI, Guilford S, Proudler AJ. Randomized trial of effect of transdermal continuous combined hormone replacement therapy on cardiovascular risk markers. Br.J.Haematol. 2004;124:802-08.

81. Lindgren R, Berg G, Hammar M, Larsson-Cohn U, Olsson AG. Plasma lipid and lipoprotein effects of transdermal administration of estradiol and estradiol/norethisterone acetate. Eur.J.Obstet.Gynecol.Reprod.Biol. 1992;47:213-21.

82. Campagnoli C, Colombo P, De Aloysio D, Gambacciani M, Grazioli I, Nappi C *et al.* Positive effects on cardiovascular and breast metabolic markers of oral estradiol and dydrogesterone in comparison with transdermal estradiol and norethisterone acetate. Maturitas 2002;41:299-311.

83. Taskinen MR, Puolakka J, Pyorala T, Luotola H, Bjaorn M, Kaarianen J *et al.* Hormone replacement therapy lowers plasma Lp(a) concentrations. Comparison of cyclic transdermal and continuous estrogen-progestin regimens. Arterioscler.Thromb.Vasc.Biol. 1996;16:1215-21.

84. Hokanson JE, Austin MA. Plasma triglyceride level is a risk factor for cardiovascular disease independent of high-density lipoprotein cholesterol level: a meta-analysis of population-based prospective studies. J.Cardiovasc.Risk 1996;3:213-19.

85. Vehkavaara S, Silveira A, Hakala-Ala-Pietila T, Virkamaki A, Hovatta O, Hamsten A *et al.* Effects of oral and transdermal estrogen replacement therapy on markers of coagulation, fibrinolysis, inflammation and serum lipids and lipoproteins in postmenopausal women. Thromb.Haemost. 2001;85:619-25.

86. Gangar KF, Vyas S, Whitehead M, Crook D, Meire H, Campbell S. Pulsatility index in internal carotid artery in relation to transdermal oestradiol and time since menopause. Lancet 1991;338:839-42.

87. Holdright DR, Sullivan AK, Wright CA, Sparrow JL, Cunningham D, Fox KM. Acute effect of oestrogen replacement therapy on treadmill performance in postmenopausal women with coronary artery disease. Eur.Heart J. 1995;16:1566-70.

88. Mueck AO, Seeger H. Effect of hormone therapy on BP in normotensive and hypertensive postmenopausal women. Maturitas 2004;49:189-203.

89. Lindheim SR, Duffy DM, Kojima T, Vijod MA, Stanczyk FZ, Lobo RA. The route of administration influences the effect of estrogen on insulin sensitivity in postmenopausal women. Fertil.Steril. 1994;62:1176-80.

90. Padwick ML, Pryse-Davies J, Whitehead MI. A simple method for determining the optimal dosage of progestin in postmenopausal women receiving estrogens. N.Engl.J.Med. 1986;315:930-34.

91. Godsland IF. Effects of postmenopausal hormone replacement therapy on lipid, lipoprotein, and apolipoprotein (a) concentrations: analysis of studies published from 1974-2000 [Review]. Fertil Steril. 2001;75:898-915.

92. Ylikorkala O, Lim P, Caubel P, olavi.ylikorkala@huch.fi, beta e, Norgestimate *et al.* Effects on serum lipid profiles of continuous 17 beta-estradiol, intermittent norgestimate regimens versus continuous combined 17 beta-estradiol/norethisterone acetate hormone replacement therapy. Clin.Ther. 2000;22:622-36.

93. Lobo RA, Zacur HZ, Caubel P, Lane R. A novel intermittent regimen of norgestimate to preserve the beneficial effects of 17beta-estradiol on lipid and lipoprotein profiles. Am.J Obstet Gynecol 2000;182:41-49.

94. Lee WS, Harder JA, Yoshizumi M, Lee ME, Haber E. Progesterone inhibits arterial smooth muscle cell proliferation. Nat.Med. 1997;3:1005-08.

95. Sarrel PM. The differential effects of oestrogens and progestins on vascular tone. Hum.Reprod.Update 1999;5:205-09.

96. Azen SP, Mack WJ, Cashin-Hemphill L, LaBree L, Shircore AM, Selzer RH *et al.* Progression of coronary artery disease predicts clinical coronary events. Long-term follow-up from the Cholesterol Lowering Atherosclerosis Study. Circulation 1996;93:34-41.

97. Pletcher MJ, Tice JA, Pignone M, Browner WS. Using the coronary artery calcium score to predict coronary heart disease events: a systematic review and meta-analysis. Arch.Intern.Med. 2004;164:1285-92.

98. Hodis HN, Mack WJ, Barth J. Carotid intima-media thickness as a surrogate end point for coronary artery disease. Circulation 1996;94:2311-12.

99. Gnasso A, Irace C, Mattioli PL, Pujia A. Carotid intima-media thickness and coronary heart disease risk factors. Atherosclerosis 1996;119:7-15.

100. Blankenhorn DH, Hodis HN. George Lyman Duff Memorial Lecture. Arterial imaging and atherosclerosis reversal. [Review] [106 refs]. Arterioscler.Thromb. 1994;14:177-92.

101. Geroulakos G, O'Gorman DJ, Kalodiki E, Sheridan DJ, Nicolaides AN. The carotid intima-media thickness as a marker of the presence of severe symptomatic coronary artery disease. Eur.Heart J. 1994;15:781-85.

102. Holaj R, Spacil J, Petrasek J, Malik J, Haas T, Aschermann M. Intima-media thickness of the common carotid artery is the significant predictor of angiographically proven coronary artery disease. Can.J.Cardiol. 2003;19:670-76.

103. Byington RP, Furberg CD, Herrington DM, Herd JA, Hunninghake D, Lowery M *et al.* Effect of estrogen plus progestin on progression of carotid atherosclerosis in postmenopausal women with heart disease: HERS B-mode substudy. Arterioscler.Thromb.Vasc.Biol 2002;22:1692-97.

104. Herrington DM, Reboussin DM, Klein KP, Sharp PC, Shumaker SA, Snyder TE *et al.* The estrogen replacement and atherosclerosis (ERA) study: study design and baseline characteristics of the cohort. [erratum appears in Control Clin Trials 2000 Aug;21(4):414.]. Control.Clin.Trials 2000;21:257-85.

105. Angerer P, Stork S, Kothny W, Schmitt P, Von Schacky C. Effect of oral postmenopausal hormone replacement on progression of atherosclerosis : a randomized, controlled trial. Arterioscler.Thromb.Vasc.Biol 2001;21:262-68.

106. Hodis HN, Mack WJ, Azen SP, Lobo RA, Shoupe D, Mahrer PR *et al.* Hormone therapy and the progression of coronary-artery atherosclerosis in postmenopausal women. N.Engl.J.Med. 2003;349:535-45.

107. Waters DD, Alderman EL, Hsia J, Howard BV, Cobb FR, Rogers WJ *et al.* Effects of hormone replacement therapy and antioxidant vitamin supplements on coronary atherosclerosis in postmenopausal women: a randomized controlled trial. JAMA 2002;288:2432-40.

108. Husak L, Vaccarino V, Veledar E, Murrah N, Wenger NK. Comparison of angiographic findings among postmenopausal women using unopposed estrogen, estrogen/progestin combinations, and nonusers. Am.J.Cardiol. 2004;93:563-68.

109. Hodis HN, Mack WJ, Lobo RA, Shoupe D, Sevanian A, Mahrer PR *et al.* Estrogen in the prevention of atherosclerosis. A randomized, double-blind, placebo-controlled trial. Ann.Intern.Med. 2001;135:939-53.

110. de Kleijn MJ, Bots ML, Bak AA, Westendorp IC, Planellas J, Coelingh Bennink HJ *et al.* Hormone replacement therapy in perimenopausal women and 2-year change of carotid intima-media thickness. Maturitas 1999;32:195-204.

111. Akhrass F, Evans AT, Wang Y, Rich S, Kannan CR, Fogelfeld L *et al.* Hormone replacement therapy is associated with less coronary atherosclerosis in postmenopausal women. J.Clin.Endocrinol.Metab 2003;88:5611-14.

112. Lobo, R. A. Menopause management for the millenium. http://www.medscape.com/Medscape/WomensHealth/ClinicalMgmt/CM.v01/pnt-CM.v01html. 2000. Ref Type: Electronic Citation

113. Mattox JH, Shulman LP. Combined oral hormone replacement therapy formulations. Am.J.Obstet.Gynecol. 2001;185:S38-S46.

114. Murphy JL, Burke J. Nurse Practitioner's Prescribing Reference, vol. Spring 2000. New York: Prescribing Reference Inc; 2000.

115. Barrett-Connor E, Grady D. Hormone replacement therapy, heart disease, and other considerations. Annu.Rev.Public Health 1998;19:55-72.

116. Grodstein F, Stampfer M. The epidemiology of coronary heart disease and estrogen replacement in postmenopausal women. [Review] [86 refs]. Prog.Cardiovasc.Dis. 1995;38:199-210.

117. Hemminki E, McPherson K. Impact of postmenopausal hormone therapy on cardiovascular events and cancer: pooled data from clinical trials. [see comments.]. BMJ 1997;315:149-53.

118. Humphrey LL, Chan BK, Sox HC. Postmenopausal hormone replacement therapy and the primary prevention of cardiovascular disease. [see comments.]. [Review] [82 refs]. Ann.Intern.Med. 2002;137:273-84.

119. Potocki J. [Effects of estrogen treatment on coronary insufficiency in women after menopause]. Pol.Tyg.Lek. 1971;26:1812-15.

120. Wolf PH, Madans JH, Finucane FF, Higgins M, Kleinman JC. Reduction of cardiovascular disease-related mortality among postmenopausal women who use hormones: evidence from a national cohort. Am.J Obstet Gynecol 1991;164:489-94.

121. U.S.Preventive Services Task Force. Postmenopausal hormone replacement therapy for primary prevention of chronic conditions: recommendations and rationale.[summary for patients in Ann Intern Med. 2002 Nov 19;137(10):I48; PMID: 12435233]. Ann.Intern.Med. 2002;Online. 137:834-39.

122. Last JM. The Dictionary of Epidemiology, Fourth ed. Oxford University Press; 2001.

123. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D *et al.* Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000;283:2008-12.

124. Hardy RJ, Thompson SG. Detecting and describing heterogeneity in meta-analysis. Stat.Med. 1998;17:841-56.

125. Shapiro S. Meta-analysis/Shmeta-analysis. Am.J.Epidemiol. 1994;140:771-78.

126. Greenland S. Can meta-analysis be salvaged? Am J Epidemiol. 1994;140:783-87.

127. Pettiti DB. Meta-analysis, decision analysis and cost effectiveness analysis. London: Oxford University Press; 2000.

128. Clemens JD, Chuong JJ, Feinstein AR. The BCG controversy. A methodological and statistical reappraisal. JAMA 1983;249:2362-69.

129. Detsky AS, Naylor CD, O'Rourke K, McGeer AJ, L'Abbe KA. Incorporating variations in the quality of individual randomized trials into meta-analysis. J.Clin.Epidemiol. 1992;45:255-65.

130. Egger M, Smith GD, Altman DG. Systematic Reviews in Health Care, Second edition ed. London, UK: Blackwell BMJ Books; 2001.

131. Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. JAMA 1999;282:1054-60.

132. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol.Community Health 1998;52:377-84.

133. Heneghan AM, Horwitz SM, Leventhal JM. Evaluating intensive family preservation programs: a methodological review. Pediatrics 1996;97:535-42.

134. Zola P, Volpe T, Castelli G, Sismondi P, Nicolucci A, Parazzini F *et al.* Is the published literature a reliable guide for deciding between alternative treatments for patients with early cervical cancer? Int J Radiat.Oncol Biol.Phys. 1989;16:785-97.

135. Bass JL, Christoffel KK, Widome M, Boyle W, Scheidt P, Stanwick R *et al.* Childhood injury prevention counseling in primary care settings: a critical review of the literature. Pediatrics 1993;92:544-50.

136. Hadorn DC, Baker D, Hodges JS, Hicks N. Rating the quality of evidence for clinical practice guidelines. J Clin Epidemiol. 1996;49:749-54.

137. Stieb DM, Frayha HH, Oxman AD, Shannon HS, Hutchison BG, Crombie FS. Effectiveness of Haemophilus influenzae type b vaccines. CMAJ 1990;142:719-33.

138. Brodsky, L. M. The association between estrogen-progestin replacement therapy and the risk of breast cancer among post-menopausal women: a systematic review and meta-analysis. 2002. University of Ottawa.
Ref Type: Thesis/Dissertation

139. Clarke M, Oxman AD. Cochrane reviewers handbook 4.1.2. The Cochrane Library. (2). Oxford, Update Software; 2001.

140. Berlin JA. Does blinding of readers affect the results of meta-analyses? University of Pennsylvania Meta-analysis Blinding Study Group. Lancet 1997;350:185-86.

141. 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;17:1-12.

142. 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;273:408-12.

143. Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. Lancet 2002;359:614-18.

144. Fleiss JL. Statistical Methods for Rates and Proportions. Toronto, Canada: John Wiley and Sons; 1981.

145. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557-60.

146. Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. J Clin Epidemiol. 2000;53:207-16.

147. Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. BMJ 1997;315:640-45.

148. Egger M, Zellweger-Zahner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. Lancet 1997;350:326-29.

149. Ravnskov U. Cholesterol lowering trials in coronary heart disease: frequency of citation and outcome. BMJ 1992;305:15-19.

150. Tramer MR, Reynolds DJ, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. BMJ 1997;315:635-40.

151. Sterne JA, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. J Clin Epidemiol. 2001;54:1046-55.

152. Hall G, Pripp U, Schenck-Gustafsson K, Landgren BM. Long-term effects of hormone replacement therapy on symptoms of angina pectoris, quality of life and compliance in women with coronary artery disease. Maturitas 1998;28:235-42.

153. Clarke SC, Kelleher J, Lloyd-Jones H, Slack M, Schofiel PM. A study of hormone replacement therapy in postmenopausal women with ischaemic heart disease: the Papworth HRT atherosclerosis study. BJOG 2002;109:1056-62.

154. Cherry N, Gilmour K, Hannaford P, Heagerty A, Khan MA, Kitchener H *et al.* Oestrogen therapy for prevention of reinfarction in postmenopausal women: a randomised placebo controlled trial.[comment]. Lancet 2002;360:2001-08.

155. Nachtigall LE, Nachtigall RH, Nachtigall RD, Beckman EM. Estrogen replacement therapy II: a prospective study in the relationship to carcinoma and cardiovascular and metabolic problems. Obstet Gynecol 1979;54:74-79.

156. Anderson GL, Limacher M, Assaf AR, Bassford T, Beresford SA, Black H *et al.* Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: the Women's Health Initiative randomized controlled trial. JAMA 2004;291:1701-12.

157. Petitti DB, Sidney S, Quesenberry CP, Jr. Hormone replacement therapy and the risk of myocardial infarction in women with coronary risk factors. Epidemiology 2000;11:603-06.

158. Sidney S, Petitti DB, Quesenberry CP, Jr. Myocardial infarction and the use of estrogen and estrogen-progestogen in postmenopausal women. Ann.Intern.Med. 1997;127:501-08.

159. Psaty BM, Heckbert SR, Atkins D, Lemaitre R, Koepsell TD, Wahl PW *et al.* The risk of myocardial infarction associated with the combined use of estrogens and progestins in postmenopausal women. Arch.Intern.Med. 1994;154:1333-39.

160. Heckbert SR, Weiss NS, Koepsell TD, Lemaitre RN, Smith NL, Siscovick DS *et al.* Duration of estrogen replacement therapy in relation to the risk of incident myocardial infarction in postmenopausal women. Arch.Intern.Med. 1997;157:1330-36.

161. Jick H, Dinan B, Rothman KJ. Noncontraceptive estrogens and nonfatal myocardial infarction. JAMA 1978;239:1407-09.

162. Jick H, Dinan B, Herman R, Rothman KJ. Myocardial infarction and other vascular diseases in young women. Role of estrogens and other factors. JAMA 1978;240:2548-52.

163. Adam S, Williams V, Vessey MP. Cardiovascular disease and hormone replacement treatment: a pilot case-control study. Br Med J (Clin Res Ed) 1981;282:1277-78.

164. Avila MH, Walker AM, Jick H. Use of replacement estrogens and the risk of myocardial infarction. Epidemiology 1990;1:128-33.

165. Bain C, Willett W, Hennekens CH. Use of postmenopausal hormones and risk of myocardial infarction. Circulation 1981;64:42-46.

166. Beard CM, Kottke TE, Annegers JF, Ballard DJ. The Rochester Coronary Heart Disease Project: effect of cigarette smoking, hypertension, diabetes, and steroidal estrogen use on coronary heart disease among 40- to 59-year-old women, 1960 through 1982. [see comments.]. Mayo Clin.Proc. 1989;64:1471-80.

167. Chilvers CE, Knibb RC, Armstrong SJ, Woods KL, Logan RF. Post menopausal hormone replacement therapy and risk of acute myocardial infarction--a case control study of women in the East Midlands, UK. Eur.Heart J. 2003;24:2197-205.

168. Croft P, Hannaford PC. Risk factors for acute myocardial infarction in women: evidence from the Royal College of General Practitioners' oral contraception study. BMJ 1989;298:165-68.

169. Grodstein F, Stampfer MJ, Colditz GA, Willett WC, Manson JE, Joffe M *et al.* Post menopausal hormone therapy and mortality. N.Engl.J.Med. 1997;336:1769-75.

170. La Vecchia C, Franceschi S, Decarli A, Pampallona S, Tognoni G. Risk factors for myocardial infarction in young women. Am J Epidemiol 1987;125:832-43.

171. Mann RD, Lis Y, Chukwujindu J, Chanter DO. A study of the association between hormone replacement therapy, smoking and the occurrence of myocardial infarction in women. [see comments.]. J.Clin.Epidemiol. 1994;47:307-12.

172. Pfeffer RI, Whipple GH, Kurosaki TT, Chapman JM. Coronary risk and estrogen use in postmenopausal women. Am J Epidemiol 1978;107:479-97.

173. Rosenberg L, Armstrong B, Jick H. Myocardial infarction and estrogen therapy in post-menopausal women. N.Engl.J.Med. 1976;294:1256-59.

174. Rosenberg L, Slone D, Shapiro S. Noncontraceptive estrogens and myocardial infarction in young women. Journal of the American Medical Association 1980;244:339-42.

175. Rosenberg L, Palmer JR, Shapiro S. A case-control study of myocardial infarction in relation to use of estrogen supplements. Am.J.Epidemiol. 1993;137:54-63.

176. Ross RK, Paganini-Hill A, Mack TM, Arthur M, Henderson BE. Menopausal oestrogen therapy and protection from death from ischaemic heart disease. Lancet 1981;1:858-60.

177. Szklo M, Tonascia J, Gordis L, Bloom I. Estrogen use and myocardial infarction risk: A case-control study. Prev.Med. 1984;13:510-16.

178. Varas-Lorenzo C, Garcia-Rodriguez LA, Perez-Gutthann S, Duque-Oliart A. Hormone replacement therapy and incidence of acute myocardial infarction. A population-based nested case-control study. Circulation 2000;101:2572-78.

179. Karagas MR. Reproductive history, menopausal factors and noncontraceptive hormone use in relation to the risk of coronary heart disease in women. Dissertation Abstracts Internationalume:51-12.

180. Beard CM, Crowson CS, Malkasian GD, O'Fallon WM, Melton III LJ. Cardiovascular disease and cancer risk following bilateral oophorectomy: A population-based study in Rochester, Minnesota. J Womens Health 1995;4:133-41.

181. Hammond CB, Jelovsek FR, Lee KL, Creasman WT, Parker RT. Effects of long-term estrogen replacement therapy. I. Metabolic effects. Am.J Obstet Gynecol 1979;133:525-36.

182. Hernandez AM, Walker AM, Jick H. Use of replacement estrogens and the risk of myocardial infarction. Epidemiology 1990;1:128-33.

183. Stampfer MJ, Willett WC, Colditz GA, Rosner B, Speizer FE, Hennekens CH. A prospective study of postmenopausal estrogen therapy and coronary heart disease. N.Engl.J.Med. 1985;313:1044-49.

184. Grodstein F, Stampfer MJ, Manson JE, Colditz GA, Willett WC, Rosner B *et al.* Postmenopausal estrogen and progestin use and the risk of cardiovascular disease. N.Engl.J.Med. 1996;335:453-61.

185. Grodstein F, Manson JE, Colditz GA, Willett WC, Speizer FE, Stampfer MJ. A prospective, observational study of postmenopausal hormone therapy and primary prevention of cardiovascular disease. [see comments.]. Ann.Intern.Med. 2000;133:933-41.

186. Henderson BE, Ross RK, Paganini-Hill A, Mack TM. Estrogen use and cardiovascular disease. Am.J Obstet Gynecol 1986;154:1181-86.

187. Henderson BE, Paganini-Hill A, Ross RK. Estrogen replacement therapy and protection from acute myocardial infarction. Am.J Obstet Gynecol 1988;159:312-17.

188. Henderson BE, Paganini-Hill A, Ross RK. Decreased mortality in users of estrogen replacement therapy. Arch Intern Med 1991;151:75-78.

189. Eaker ED, Chesebro JH, Sacks FM, Wenger NK, Whisnant JP, Winston M. Cardiovascular disease in women. [see comments.]. [Review] [139 refs]. Circulation 1993;88:1999-2009.

190. Wilson PW, Garrison RJ, Castelli WP. Postmenopausal estrogen use, cigarette smoking, and cardiovascular morbidity in women over 50. The Framingham Study. N.Engl.J.Med. 1985;313:1038-43.

191. Cauley JA, Seeley DG, Browner WS, Ensrud K, Kuller LH, Lipschutz RC *et al.* Estrogen replacement therapy and mortality among older women. The study of osteoporotic fractures. Arch Intern Med 1997;157:2181-87.

192. Criqui MH, Suarez L, Barrett-Connor E, McPhillips J, Wingard DL, Garland C. Postmenopausal estrogen use and mortality. Results from a prospective study in a defined, homogeneous community. Am.J.Epidemiol. 1988;128:606-14.

193. Ettinger B, Friedman GD, Bush T, Quesenberry CP, Jr. Reduced mortality associated with long-term postmenopausal estrogen therapy. Obstet Gynecol 1996;87:6-12.

194. Folsom AR, Mink PJ, Sellers TA, Hong CP, Zheng W, Potter JD. Hormonal replacement therapy and morbidity and mortality in a prospective study of postmenopausal women. Am.J.Public Health 1995;85:1128-32.

195. Hedblad B, Merlo J, Manjer J, Engstrom G, Berglund G, Janzon L. Incidence of cardiovascular disease, cancer and death in postmenopausal women affirming use of hormone replacement therapy. Scand.J.Public Health 2002;30:12-19.

196. Lafferty FW, Fiske ME. Postmenopausal estrogen replacement: a long-term cohort study. Am.J.Med. 1994;97:66-77.

197. Lauritzen C, Meier F. Risks of endometrial and mammary cancer morbidity and mortality in long term oestrogen treatment, The Climacteric: An update. 1979. p. 207-17.

198. Lokkegaard E, Pedersen AT, Heitmann BL, Jovanovic Z, Keiding N, Hundrup YA *et al.* Relation between hormone replacement therapy and ischaemic heart disease in women: prospective observational study. BMJ 2003;326:426.

199. Petitti DB, Perlman JA, Sidney S. Noncontraceptive estrogens and mortality: long-term follow-up of women in the Walnut Creek Study. Obstet Gynecol 1987;70:289-93.

200. Rodriguez C, Calle EE, Patel AV, Tatham LM, Jacobs EJ, Thun MJ. Effect of body mass on the association between estrogen replacement therapy and mortality among elderly US women. Am.J.Epidemiol. 2001;153:145-52.

201. Sourander L, Rajala T, Raiha I, Makinen J, Erkkola R, Helenius H. Cardiovascular and cancer morbidity and mortality and sudden cardiac death in postmenopausal women on oestrogen replacement therapy (ERT). Lancet 1998;352:1965-69.

202. Kirkland SA. Postemnopausal hormone use and ischemic heart disease: A retrospective cohort study (estrogen). Dissertation Abstracts Internationalume:58-06.

203. Newton KM, LaCroix AZ, McKnight B, Knopp RH, Siscovick DS, Heckbert SR *et al.* Estrogen replacement therapy and prognosis after first myocardial infarction. Am.J.Epidemiol. 1997;145:269-77.

204. Heckbert SR, Kaplan RC, Weiss NS, Psaty BM, Lin D, Furberg CD *et al.* Risk of recurrent coronary events in relation to use and recent initiation of postmenopausal hormone therapy. Arch.Intern.Med. 2001;161:1709-13.

205. Grodstein F, Manson JE, Stampfer MJ. Postmenopausal hormone use and secondary prevention of coronary events in the nurses' health study. a prospective, observational study. Ann.Intern.Med. 2001;135:1-8.

206. Alexander KP, Newby LK, Hellkamp AS, Harrington RA, Peterson ED, Kopecky S *et al.* Initiation of hormone replacement therapy after acute myocardial infarction is associated with more cardiac events during follow-up. [see comments.]. J.Am.Coll.Cardiol. 2001;38:1-7.

207. O'Keefe JH, Jr., Kim SC, Hall RR, Cochran VC, Lawhorn SL, McCallister BD. Estrogen replacement therapy after coronary angioplasty in women. J.Am.Coll.Cardiol. 1997;29:1-5.

208. Khan MA, Liu MW, Singh D, Pal A, Chio FL, Lawson D *et al.* Long-term (three years) effect of estrogen replacement therapy on major adverse cardiac events in postmenopausal women after intracoronary stenting. Am.J.Cardiol. 2000;86:330-33.

209. Abu-Halawa SA, Thompson K, Kirkeeide RL, Vaughn WK, Rosales O, Fujisi K *et al.* Estrogen replacement therapy and outcome of coronary balloon angioplasty in postmenopausal women. Am.J.Cardiol. 1998;82:409-13.

210. Preliminary report: effect of encainide and flecainide on mortality in a randomized trial of arrhythmia suppression after myocardial infarction. The Cardiac Arrhythmia Suppression Trial (CAST) Investigators. N.Engl.J.Med. 1989;321:406-12.

211. Rodstrom K, Bengtsson C, Lissner L, Bjorkelund C, Kerstin.Rodstrom@allmed.gu.se. Pre-existing risk factor profiles in users and non-users of hormone replacement therapy: prospective cohort study in Gothenburg, Sweden. Br.Med.J. 1999;319:890-93.

212. Derby CA, Hume AL, Mcphillips JB, Barbour MM, Carleton RA, Estrogen rt *et al.* Prior and current health charcteristics of postmenopausal estrogen replacement therapy users compared with nonusers. Am.J Obstet Gynecol 1995;173:544-50.

213. Matthews KA. Prior to use of estrogen replacement therapy, are users healthier than nonusers? Am.J.Epidemiol. 1996;143.

214. Posthuma WFM, Westendorp RGJ, Vandenbroucke JP. Cardioprotective effect of hormone replacement therapy in postmenopausal women: Is the evidence biased? BMJ 1994;308:1268-69.

215. Gallagher EJ, Viscoli CM, Horwitz RI. The relationship of treatment adherence to the risk of death after myocardial infarction in women. JAMA 1993;270:742-44.

216. Horwitz RI, Viscoli CM, Berkman L, Donaldson RM, Horwitz SM, Murray CJ *et al.* Treatment adherence and risk of death after a myocardial infarction. Lancet 1990;336:542-45.

217. Petitti DB. Coronary heart disease and estrogen replacement therapy. Can compliance bias explain the results of observational studies? Ann.Epidemiol. 1994;4:115-18.

218. Goodman MT, Cologne JB, Moriwaki H, Vaeth M, Mabuchi K. Risk factors for primary breast cancer in Japan: 8-year follow-up of atomic bomb survivors. Prev.Med. 1997;26:144-53.

219. Paganini-Hill A, Ross RK. Reliability of recall of drug usage and other health-related information. Am.J.Epidemiol. 1982;116:114-22.

220. Chemoprevention: hormone replacement therapy. US Preventive Services Task Force.Guide to clinical preventive services . 2003.
Ref Type: Internet Communication

221. Vercellini P, De Giorgi O, Pisacreta A, Pesole AP, Vicentini S, Crosignani PG. Surgical management of endometriosis. Baillieres Best.Pract.Res.Clin.Obstet.Gynaecol. 2000;14:501-23.

222. SanGiovanni JP, Berkey CS, Dwyer JT, Colditz GA. Dietary essential fatty acids, long-chain polyunsaturated fatty acids, and visual resolution acuity in healthy fullterm infants: a systematic review. Early Hum.Dev. 2000;57:165-88.

223. Malhotra S, Kondal A, Shafiq N, Sidhu S, Bhasin DK, Pandhi P. A comparison of observational studies and controlled trials of heparin in ulcerative colitis. Int.J.Clin.Pharmacol.Ther. 2004;42:690-94.

224. Conaty S, Watson L, Dinnes J, Waugh N. The effectiveness of pneumococcal polysaccharide vaccines in adults: a systematic review of observational studies and comparison with results from randomised controlled trials. Vaccine 2004;22:3214-24.

225. Schwammenthal Y, Tanne D. Homocysteine, B-vitamin supplementation, and stroke prevention: from observational to interventional trials. Lancet Neurol. 2004;3:493-95.

226. Bauer DC, Mundy GR, Jamal SA, Black DM, Cauley JA, Ensrud KE *et al.* Use of statins and fracture: results of 4 prospective studies and cumulative meta-analysis of observational studies and controlled trials. Arch.Intern.Med. 2004;164:146-52.

227. Jadad AR, Cook DJ, Browman GP. A guide to interpreting discordant systematic reviews. CMAJ. 1997;156:1411-16.

228. Ioannidis JP, Haidich AB, Pappa M, Pantazis N, Kokori SI, Tektonidou MG *et al.* Comparison of evidence of treatment effects in randomized and nonrandomized studies. JAMA 2001;286:821-30.

229. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. N.Engl.J.Med. 2000;342:1878-86.

230. Britton A, McKee M, Black N, McPherson K, Sanderson C, Bain C. Choosing between randomised and non-randomised studies: a systematic review. Health Technol.Assess. 1998;2:i-124.

231. Chalmers TC, Matta RJ, Smith H, Jr., Kunzler AM. Evidence favoring the use of anticoagulants in the hospital phase of acute myocardial infarction. N.Engl.J.Med. 1977;297:1091-96.

232. Chalmers TC, Celano P, Sacks HS, Smith H, Jr. Bias in treatment assignment in controlled clinical trials. N.Engl.J.Med. 1983;309:1358-61.

233. Egger M, Smith GD, Schneider M. Systematic reviews of observational studies, In: Egger M, Smith GD, Altman DG, editors. Systematic Reviews in Health Care, Second ed. London: Blackwell BMJ Books; 2001. p. 211-26.

234. Shadish, W. and Myers, D. Campbell Collaboration:Research design Policy Brief. Campbell Collaboration . 2004.
Ref Type: Internet Communication

235. Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M *et al.* Methodology and reports of systematic reviews and meta-analyses: a comparison of Cochrane reviews with articles published in paper-based journals. JAMA 1998;280:278-80.

236. Cook DJ, Sackett DL, Spitzer WO. Methodologic guidelines for systematic reviews of randomized control trials in health care from the Potsdam Consultation on Meta-Analysis. J.Clin.Epidemiol. 1995;48:167-71.

237. Cowley DE. Prostheses for primary total hip replacement. A critical appraisal of the literature. Int.J.Technol.Assess.Health Care 1995;11:770-78.

238. Reisch JS, Tyson JE, Mize SG. Aid to the evaluation of therapeutic studies. Pediatrics 1989;84:815-27.

239. Thomas, H. Quality assessment tool for quantitative studies. Effective Public Health Practice Project. 1998. Unpublished Work

240. Wells, G and Shay, B. Data extraction for non-randomized systematic review. 2000. Unpublished Work

241. Zaza S, Wright-De Aguero LK, Briss PA, Truman BI, Hopkins DP, Hennessy MH *et al.* Data collection instrument and procedure for systematic reviews in the Guide to Community Preventive Services. Task Force on Community Preventive Services. Am.J.Prev.Med. 2000;18:44-74.

242. Terrin N, Schmid CH, Lau J, Olkin I. Adjusting for publication bias in the presence of heterogeneity. Stat.Med. 2003;22:2113-26.

243. Moher D, Pham B, Klassen TP, Schulz KF, Berlin JA, Jadad AR *et al.* What contributions do languages other than English make on the results of meta-analyses? J Clin Epidemiol. 2000;53:964-72.

244. McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? Lancet 2000;356:1228-31.

245. Barrett-Connor E. Clinical review 162: cardiovascular endocrinology 3: an epidemiologist looks at hormones and heart disease in women. J.Clin.Endocrinol.Metab 2003;88:4031-42.

246. Hill AB. The environment and disease: Association or causation? Proc.R.Soc.Med. 1965;58:295-300.

APPENDICES

APPENDIX 1
Search strategies for electronic databases

1a Strategy used for Medline/Premedline

1 exp Hormone Replacement Therapy
2 ((oestrogen or estrogen) and replacement$).tw.
3 ((postmenopaus$ adj hormon$) and replacement$).tw.
4 (hormon$ and replacement$).tw.
5 exp estrogens/
6 (oestrogen$ or estrogen$).tw.
7 4 and (5 or 6)
8 1 or 2 or 3 or 5 or 7
9 exp Myocardial Ischemia/
10 (m yocardial ischemia or ischemic heart disease or (coronary adj2 disease)).tw.
11 (M yocardial Infarction or myocardial stunning).tw.
12 (Coronar y adj (Aneurysm or Arteriosclerosis or Stenosis or Thrombosis or Vasospasm)).tw.
13 or/9 -12
14 8 and 13
15 h ypertension.hw,tw.
16 8 and 15
17 exp lipids /bl
18 lipid level$.tw.
19 17 or 18
20 8 and 19
21 meta -analysis.pt,sh.
22 (meta -anal: or metaanal:).tw.
23 (quantitativ: review: or q uantitativ: overview:).tw.
24 (s ystematic: review: or systematic: overviewl).tw.
25 (methodologic: r eview: or methodologic: overview:).tw.
26 review.pt,sh. or r eview:.tw. or overview:.tw.
27 clinical t rial.pt.
28 randomized controlled trial.pt.
29 random$.tw.
30 (double adj blind$).tw.
31 placebo$.tw.
32 exp Cohort Stud ies/
33 exp case -control studies/
34 or/21 -33
35 14 and 34
36 16 and 34
37 20 and 34

1b Search strategy for Embase

1. exp Hormone Substitution
2. ((oestrogen or estrogen) and replacement$).tw.
3. ((postmenopaus$ adj hormon$) and replacement$).tw.
4. (hormon$ and replacement$).tw.
5. exp estrogen/
6. (oestrogen$ or estrogen$).tw.
7. 4 and (5 or 6)
8. 1 or 2 or 3 or 5 or 7
9. exp Heart Disease
10. (m yocardial ischemia or ischemic heart disease or (coronary adj2 disease)).tw.
11. (M yocardial Infarction or myocardial stunning).tw.
12. (Coronar y adj (Aneurysm or Arteriosclerosis or stenosis or Thrombosis or Vasospasm)).tw.
13. or/9 -12
14. 8 and 13
15. h ypertension.hw,tw.
16. 8 and 15
17. exp lipid
18. lipid level$.tw.
19. 17 or 18
20. 8 and 19 (2886)
21. Cohort Anal ysis
22. Case Control Stud y
23. meta anal ys#s.tw.
24. meta anal ysis.de.
25. s ystematic review$.tw.
26. metaanal$.tw.
27. meta anal$.tw.
28. s ystematic overview$.tw.
29. controlled stud y.de.
30. clinical trial.de.
31. major clinical stud y.de
32. random$.tw.
33. randomized c ontrolled trial.de.
34. trial$.tw.
35. compar$.tw.
36. clinical trial
37. double blind procedur e.de.
38. or /21 -38
39. 14 and 38
40. 16 and 38
41. 19 and 38

APPENDIX 2a Data abstraction for Randomized controlled trials

HRT and CAD in postmenopausal women - RCT Data Extraction Form

Identification:

Article #	
Date	
Reviewer #	
Title	
Investigator	
Journal	Year Vol Starting page
Study Duration	
Country	
End points 1. overall CAD 2. Myocardial infarction 3. Angina 4. Fatal CAD alone	
Extraction from graphs: Y/N	
Comments	

Study Design:

Experimental Design: Parallel Group

Intervention:

HRT specify dose, preparation, regimen and route_____

_____

_____

Treatment Groups:

1.

2.

Concurrent treatment

Methodology:

Concealment of allocation	Blind randomization e.g. sealed envelopes Unclear methods of randomization Quasi-randomization e.g. even and odd date of birth		
Blinding	Double Single Open	Assessor Patient Investigator	
Analysis		Efficacy	ITT

Participants:

Inclusion Criteria Exclusion Criteria:

Characteristics:

	Group 1	Group 2
Number		
Withdrawals		
Age		
Females		
Smokers		
Diabetes		
High Cholesterol		
BMI		
Concurrent medication		
Non-compliance		

Pretreatment group differences:

Compliance to Treatment:

Adverse Events: Described Yes No If Yes , describe:

Results:

Outcome	Intervention RR (95% CI)	Placebo RR(95%CI)

Comments:

APPENDIX 2b Data abstraction for Case control studies

Identification	*(1)* Article #	*(2)* Reviewer
(3) **Title**		
(4) **Authors**		*(5)* Year
Setting	*(6)* Location *(8)* Study Name	*(7)* Language
(9) **Study Dates**		
Outcome(s)	*(10)* Outcome CHD: Fatal MI Nonfatal MI Angina Blood Pressure Lipid levels	*(11)* Criteria
(12) **Exposure**	HRT: Estrogen and Progesterone Estrogen alone Oral/transdermal	

(13)
Comments: _____

_____

_____

Methodology

	Cases	Control
(14) # Identified		
(15) # Excluded		
(16) Group N		
Group Selection	*(17)* Case Definition Excluded cases with Dx prior to exposure ❑ Y ❑ N ❑ NA	*(18)* Control Definition ❑ Randomly Selected ❑ Excluded Hist of Disease ❑ Community Controls ❑ Hospital Controls ❑ Matched _____ ❑ Unmatched ❑ Single Control ❑ Multiple Controls_____
(19) Group Determination	❑ Direct Measurement ❑ Medical Records ❑ Self Report	❑ Direct Measurement ❑ Medical Records ❑ Self Report
(20) Ascertainment of Exposure	❑ Direct Measurement ❑ Medical Records ❑ Self Report Blinding ❑ Y ❑ N ❑ NA	❑ Direct Measurement ❑ Medical Records ❑ Self Report Blinding ❑ Y ❑ N ❑ NA

(21)
Subject Characteristics

Characteristic	Cases	Controls
Age		
Exclusion criteria		
Risk factors: family history of CHD		
Previous h/o CHD		
High BP		
Hyperlipidemia		
Diabetes		

(22)
Group Differences:

(23)
Data Analysis: ❑ All subjects ❑ Excludes ascertainment exposure unknown

(24) **Exposure**	*(25)* **N**			*(26)* **Crude OR (CI)**	*(27)* **AdjustedOR (CI)**	*(28)* **Factors Controlled**	*(29)* **Translation**
		Case	Control			❑ _____ ❑ _____ ❑ _____ ❑ _____	
	E						
	–						
	T						
		Case	Control			❑ _____ ❑ _____ ❑ _____ ❑ _____	
	E						
	–						
	T						
		Case	Control			❑ _____ ❑ _____ ❑ _____ ❑ _____	
	E						
	–						
	T						

APPENDIX 2c Data abstraction for cohort studies

Identification	*(1)* Article #	*(2)* Reviewer
(3) **Title**		
(4) **Investigator**		*(5)* Year
Setting	*(6)* Location *(8)* Study Name	*(7)* Language
(9) **Dates of Enrolment/ Follow up**		
Endpoint	*(10)* Endpoint	*(11)* Criteria
(12) **Exposure**		

(13)
Comments: _____

Methodology

(14)

Design: ❑ Prospective Cohort ❑ Retrospective Cohort
❑ Internal Cohort ❑ External Cohort

	Exposed (Cohort) Unexposed	
(15) # Identified (Describe)		
(16) # Excluded/ Lost to Follow Up (Describe)		
(17) Group N		
(18) All S's Accounted ?		
Group Selection	*(19)* Exposed	*(20)* Unexposed
(21) Group Determination	❑ Direct Measurement ❑ Medical Records ❑ Self Report Blinding ❑ Yes ❑ No	❑ Direct Measurement ❑ Medical Records ❑ Self Report Blinding ❑ Yes ❑ No
(22) Ascertainment of Outcome	❑ Direct Measurement ❑ Medical Records ❑ Self Report Blinding ❑ Yes ❑ No	❑ Direct Measurement ❑ Medical Records ❑ Self Report Blinding ❑ Yes ❑ No

(23)
Subject Characteristics

Characteristic	Cohort	Exposed	Unexposed

(24)
Group Differences:

_____

_____

_____

_____

(25)
Data Analysis: ❑ All subjects ❑ Loss to follow up excluded

(26) Outcome and Exposure	*(27)* N				*(28)* Crude RR/OR (CI)	*(29)* Adjusted RR/OR (CI)	*(30)* Factors Adjusted	*(31)* Translation
Outcome		D	_	T/PY			❑ ____ ❑ ____ ❑ ____ ❑ ____ ❑ ____ ❑ ____ ❑ ____ ❑ ____	
	E							
Exposure	–							
Outcome		D	_	T/PY			❑ ____ ❑ ____ ❑ ____ ❑ ____ ❑ ____ ❑ ____ ❑ ____ ❑ ____	
	E							
Exposure	–							
Outcome		D	_	T/PY			❑ ____ ❑ ____ ❑ ____ ❑ ____ ❑ ____ ❑ ____ ❑ ____ ❑ ____	
	E							
Exposure	–							

APPENDIX 3a NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE
CASE CONTROL STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Exposure categories. A maximum of two stars can be given for Comparability.

Selection

1) Is the case definition adequate?
   a) yes, with independent validation *
   b) yes, eg record linkage or based on self reports
   c) no description

2) Representativeness of the cases
   a) consecutive or obviously representative series of cases *
   b) potential for selection biases or not stated

3) Selection of Controls
   a) community controls *
   b) hospital controls
   c) no description

4) Definition of Controls
   a) no history of disease (endpoint) *
   b) no description of source

Comparability

1) Comparability of cases and controls on the basis of the design or analysis
   a) study controls for _____ (Select the most important factor.) *
   b) study controls for any additional factor * (This criteria could be modified to indicate specific control for a second important factor.)

Exposure

1) Ascertainment of exposure
   a) secure record (eg surgical records) *
   b) structured interview where blind to case/control status *
   c) interview not blinded to case/control status
   d) written self report or medical record only
   e) no description

2) Same method of ascertainment for cases and controls
   a) yes *
   b) no

3) Non-Response rate
   a) same rate for both groups *
   b) non respondents described
   c) rate different and no designation

APPENDIX 3b NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE
COHORT STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

Selection

1) Representativeness of the exposed cohort
   a) truly representative of the average _____ (describe) in the community *
   b) somewhat representative of the average _____ in the community *
   c) selected group of users eg nurses, volunteers
   d) no description of the derivation of the cohort

2) Selection of the non exposed cohort
   a) drawn from the same community as the exposed cohort *
   b) drawn from a different source
   c) no description of the derivation of the non exposed cohort

3) Ascertainment of exposure
   a) secure record (eg surgical records) *
   b) structured interview *
   c) written self report
   d) no description

4) Demonstration that outcome of interest was not present at start of study
   a) yes *
   b) no

Comparability

1) Comparability of cohorts on the basis of the design or analysis
   a) study controls for _____ (select the most important factor) *
   b) study controls for any additional factor * (This criteria could be modified to indicate specific control for a second important factor.)

Outcome

1) Assessment of outcome
   a) independent blind assessment *
   b) record linkage *
   c) self report
   d) no description

2) Was follow-up long enough for outcomes to occur
   a) yes (select an adequate follow up period for outcome of interest) *
   b) no

3) Adequacy of follow up of cohorts
   a) complete follow up - all subjects accounted for *
   b) subjects lost to follow up unlikely to introduce bias - small number lost - > ____ % (select an adequate %) follow up, or description provided of those lost) *
   c) follow up rate < ____% (select an adequate %) and no description of those lost
   d) no statement

APPENDIX 4 Jadad's scale for quality assessment of RCTs

Study ID Date Rater's initials

	Score
Randomization Was the study described as randomized (this includes the use of words such as randomly, random and randomization)? Yes=1 No=0 Give one additional point if: the method of randomisation appropriate (e.g. table of random numbers computer generated, coin tossing, etc.)? Appropriate=1 Not appropriate=0 **Deduct 1 point if:** the method of randomisation was described and it was inappropriate (patients were allocated alternatively, or according to date of birth, hospital number, etc.)	
SCORE	
Double-blinding Was the study described as double-blind? Yes=1 No=0 **Give 1 additional point if:** the method of double-blinding was described and it was appropriate (identical placebo, active placebo, dummy, etc.) Appropriate=1 Not appropriate=0 **Deduct 1 point if:** the study was described as double-blind but the method of blinding was inappropriate (e.g. comparison of tablet vs. injection with no double dummy)	
SCORE	
Withdrawals and dropouts Was there a description of withdrawals and dropouts?	
SCORE	
TOTAL SCORE	

ALLOCATION CONCEALMENT

Adequacy of allocation concealment (circle one)

- Central randomization; numbered or coded bottles or containers; drug prepared by a pharmacy, serially numbered, opaque, sealed envelopes, etc.....ADEQUATE
Alternation; reference to case record# or date of birth, etc.....INADEQUATE
Allocation concealment is not reported, or fits neither category.....UNCLEAR

APPENDIX 5 Excel worksheet for calculation of summary estimates for non-randomized studies

Author,year	OR/RR	Lower CI	Upper CI	ln of OR	ln of LCI	ln of UCI	UCI-LCI	Std Error	SE*SE	Weight	Sum WY	(Y) x (Y)	Chi Sq T	Chi Sq Homo
Overall CAD Current users														
Varas Lorenzo 2000	0.72	0.59	0.89	-0.3285	-0.5276	-0.11653	0.4111	0.104872	0.010998	90.92419	-29.869	0.107915	9.812077	0.1284
Heckbert 1997	0.7	0.55	0.89	-0.3567	-0.5978	-0.11653	0.4813	0.122781	0.015075	66.33377	-23.6596	0.127217	8.438784	0.2867
Mann 1994	0.83	0.66	1.03	-0.1863	-0.4155	0.029559	0.4451	0.113539	0.012891	77.5724	-14.454	0.034719	2.693214	0.8487
Avila 1990	0.7	0.40	1.40	-0.3567	-0.9163	0.336472	1.2528	0.319582	0.102133	9.791164	-3.49226	0.127217	1.245603	0.0423
Pfeffer 1978	0.78	0.38	1.63	-0.2485	-0.9676	0.48858	1.4562	0.37147	0.13799	7.246888	-1.80057	0.061733	0.447372	0.0131
									Total	251.8684	-73.2754		22.63705	1.3192
Summary														
Summary estimate	0.75	0.66	0.84											
Chi sq (overall)	22.6370													
Chi sq Association	21.3178													
Chi Sq homogeneity	23.3871													

APPENDIX 6 Modifications in the 'Comparability' component

Item 1: The study controls for the primary potential confounder (s) that were identified beforehand. *

Item 2: The study controls for the secondary potential confounder (s) that were identified.*

Item 3: The methods used for controlling the potential confounders were adequately rigorous. *

Explanation of the above items:

1.& 2. The reviewers will decide on the primary and secondary confounders before starting the quality assessment. A study will get credit if it controlled all the primary confounders. It will get a second star if it controlled for all the secondary factors. If the study fails to control for all the factors that were decided upon as primary or secondary, it will not get the credit in those categories respectively. For e.g. in the review on HRT and CAD, the primary confounder is age and the secondary confounders are smoking, diabetes and high cholesterol. A study gets a star if it controls for age. It gets another star if it controls for smoking, diabetes and high cholesterol also. If it controlled for smoking but not diabetes, it doesn't get a second star.

3. For assessing the rigor of methods used, the reviewers may need an input from an expert well versed in methodology to understand what analysis is appropriate for their review. It is an overall assessment of the appropriateness and thoroughness of the design or analysis used for achieving comparability among the groups on factors other than those being studied. For e.g. Close matching on age groups, paired analysis for matched groups, conditional logistic regression for matched data, multivariate analysis for multiple confounders, etc are all appropriate ways of achieving comparability and would get a star.

APPENDIX 7a
Comparison of scores on the 'comparability' component of the original and modified NOS for case control studies

Study, year	Original NOS Study controlled for factors		Modified NOS Study controlled for factors		
	Age	Any other	Age	DM+smoking+ HC	Analysis Appropriate
Chilvers, 2003	*	*	*	*	*
Varas Lorenzo, 2000	*	*	*	*	*
Grodstein, 2000	*	*	*	*	*
Heckbert, 1997	*	*	*	-	*
Sidney, 1997	*	*	*	*	*
Mann, 1994	*	*	*	*	*
Rosenberg, 1993	*	*	*	*	*
Avila, 1990	*	*	*	-	*
Beard, 1989	*	*	*	-	*
Croft, 1989	*	*	*	-	*
La Vecchia, 1987	*	*	*	*	*
Szklo, 1984	*	*	*	-	*
Bain, 1981	*	*	*	*	*
Ross, 1981	*	*	*	*	*
Adam, 1980	*	-	*	-	-
Rosenberg, 1980	*	*	*	*	*
Jick, 1978	-	-	-	-	-
Pfeffer, 1978	*	*	*	-	*
Rosenberg, 1976	*	*	*	-	*

Appendix 7b
Comparison of scores on the 'comparability' component of the original and modified NOS for cohort studies

Study, year	Original NOS Study controlled for factor		Modified NOS Study controlled for factors		
	Age	Any other	Age	DM+smoking+ HC	Analysis Appropriate
Lokkegard, 2003	*	*	*	*	*
Hedblad, 2002	*	*	*	*	*
Rodriquez, 2001	*	*	*	-	*
Grodstein, 2000	*	*	*	*	*
Sourander, 1998	*	*	*	-	*
Cauley, 1997	*	*	*	-	*
Ettinger, 1996	*	*	*	-	*
Folsom, 1995	*	*	*	-	*
Beard, 1995	*	*	*	-	*
Lafferty, 1994	*	-	*	-	*
Avila, 1990	*	-	*	-	*
Henderson, 1988	*	*	*	-	*
Criqui, 1988	*	*	*	*	*
Pettiti, 1987	*	*	*	-	*
Wilson, 1985	*	*	*	-	*
Lauritzen 1979	*	*	*	-	-
Hammond, 1979	-	-	-	-	-

Comparison of scores on the 'comparability' component of the original and modified NOS for cohort studies on women with pre-existing heart disease

Study, year	Original NOS Study controlled for factor		Modified NOS Study controlled for factors		
	Age	Any other	Age	DM+smoking+ HC	Analysis Appropriate
Alexander, 2001	*	*	*	*	*
Heckbert, 2001	*	*	*	-	*
Grodstein, 2001	*	*	*	*	*
Khan, 2000	-	-	-	-	-
Abu-Hawala, 1998	*	*	*	-	*
O'Keefe, 1997	*	*	*	*	*