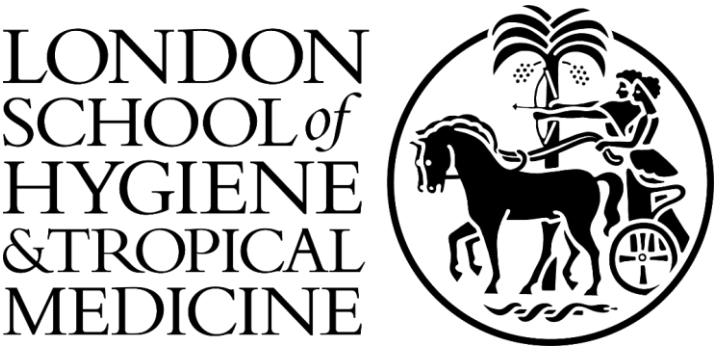
ADJUVANT ENDOCRINE THERAPY USE IN FEMALE BREAST CANCER SURVIVORS AND THE RISK OF CARDIOVASCULAR DISEASE

ANTHONY MATTHEWS



Thesis submitted in accordance with the requirements for the degree of Doctor of Philosophy

January 2019

Department of Non-Communicable Disease Epidemiology
Faculty of Epidemiology and Population Health
London School of Hygiene and Tropical Medicine
University of London

Research group affiliation: LSHTM Electronic Health Records Research Group

Funding: Anthony Matthews was employed on a Wellcome Trust/Royal Society Sir Henry Dale Fellowship held by Krishnan Bhaskaran for the duration of the PhD project

Declaration of authorship

I, Anthony Matthews, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this had been indicated in the thesis

Signed: 

Acknowledgements

Throughout the duration of this project, I have been very lucky to be supervised by Krishnan Bhaskaran and Susannah Stanway, who were always on hand to offer guidance, support, and inspiration. I would also like to thank Liam Smeeth and Alexander Lyon for being part of my advisory committee and offering their knowledge and expertise.

Thanks to Sara Thomas, Ruth Farmer, and Helen Strongman for their collaboration on the systematic review; Jennifer Lund and Sharon Peacock-Hinton, who both warmly welcomed me to the University of North Carolina at Chapel Hill for three months and offered valuable insight throughout the stay; and my friends at LSHTM who have made my time at the School thoroughly enjoyable and memorable.

Finally, I would like to thank Carolin for her continuous love and support, and my parents and brother Jordan for their endless guidance and encouragement.

ABSTRACT

BACKGROUND: Women diagnosed with oestrogen receptor positive breast cancer are prescribed endocrine therapies, tamoxifen or aromatase inhibitors (AI), to block the effects of oestrogen on the growth of tumour cells. Tamoxifen is known to be associated with an increased risk of venous thromboembolic events, but the long-term effect of both tamoxifen and AI use on the risk of a range of distinct cardiovascular diseases (CVD) remain unclear.

METHODS: Two studies were carried out using prospectively collected data from the UK Clinical Practice Datalink linked with Hospital Episode Statistics, and data from the US SEER-Medicare database, to assemble cohorts of postmenopausal women diagnosed with breast cancer. Cox proportional hazards regression models were used to examine the associations between tamoxifen and AIs and a comprehensive range of CVDs in both the UK and US study populations. The UK study directly compared AI users with tamoxifen users; the US study included a third "unexposed" group of women with oestrogen receptor positive breast cancer prescribed no endocrine therapies.

RESULTS: Results in the UK study suggested a pattern of an increased risk of non-venous CVDs in AI compared with tamoxifen users, with evidence of an increased risk of heart failure and arrhythmias in AI compared with tamoxifen users (adjusted HR: 1.70, 95% CI: 1.26-2.29; adjusted HR: 1.38, 95% CI 1.12-1.70 respectively). The US study then suggested that these associations may be driven by a decreased risk of non-venous CVD outcomes in tamoxifen users compared with those unexposed, with adjusted HRs ranging from 0.44 (95% CI: 0.30-0.63) in the myocardial infarction analysis to 0.91 (95% CI: 0.75-1.10) in the peripheral vascular disease analysis. Evidence of a modest reduced risk of several non-venous CVD outcomes among AI users compared with those unexposed was also observed, but results were suggestive of residual confounding. As expected there were more venous thromboembolic events in tamoxifen users compared with both AI users and those unexposed. There was a general consistency between comparable results in the two studies.

CONCLUSIONS: Among postmenopausal women diagnosed with breast cancer, there was convincing evidence of a higher risk of several non-venous CVDs in those prescribed an AI compared with those prescribed tamoxifen, but this appeared to be driven by protective effects of tamoxifen on these outcomes, rather than any toxic effects of AIs. The known association between tamoxifen use and an increased risk of venous thromboembolism was also confirmed. As more postmenopausal women diagnosed with oestrogen receptor positive breast cancer are prescribed AIs rather than tamoxifen, additional large-scale population based studies are needed to better understand the risk-benefit balance of endocrine therapies with respect to both cancer and cardiovascular outcomes.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 6

CHAPTER 7

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

ABBREVIATIONS

ABI - Ankle-brachial index

ACEi - ACE inhibitor

AI - Aromatase inhibitor

ARB - Angiotensin receptor blocker

BMI - Body mass index

BNP - B-type natriuretic peptide

CABG - Coronary artery bypass grafting

CCB - Calcium channel blocker

CI - Confidence interval

CKD - Chronic kidney disease

CPRD - Clinical Practice Research Datalink

CT - Computed tomography

CVD - Cardiovascular disease

DVT - Deep vein thrombosis

ECG - Electrocardiography

EHR - Electronic health records

ER – Oestrogen receptor

ER/PR – Oestrogen or progesterone receptor

GP - General practitioner

HCPCS - Healthcare Common Procedural Coding System

HER2 - Human epidermal growth factor receptor 2

HES - Hospital Episode Statistics

HF - Heart failure

HR - Hazard ratio

ICD-10 - International Classification of Diseases 10th revision

IMD – Index of Multiple Deprivation

LVEF - Left ventricular ejection fraction

LVH - Left ventricular hypertrophy

MI - Myocardial infarction

MRI - Magnetic resonance imaging

NDC - National Drug Code

NHS - National Health Service

NT-pro-BNP - N-terminal pro-brain natriuretic peptide

OECD - Organisation for Economic Cooperation and Development

OPCS-4 - Office of Population Censuses and Surveys Classification of Interventions and Procedures version 4

PCI - Percutaneous coronary intervention

PE - Pulmonary embolism

PR – Progesterone receptor

PVD - Peripheral vascular disease

RA - Rheumatoid arthritis

RCT - Randomised controlled trial

RR - Rate ratio

SACT - Systemic Anti-Cancer Therapy

SCA - Sudden cardiac arrest

SEER - Surveillance, Epidemiology and End Results Program

STEMI - ST elevation MI

THIN - The Health Improvement Network

UK - United Kingdom

UNC - University of North Carolina

US - United States

VF - Ventricular fibrillation

VHD - Valvular heart disease

VT- Ventricular tachycardia

VTE - Venous thromboembolism

WHO - World Health Organization

1 BACKGROUND

This chapter will give an overview of the definition and epidemiology of breast cancer and cardiovascular disease (CVD), and discuss the overlap between the two diseases with respect to their risk factors. This is followed by a description of the aims and objectives of the thesis.

1.1 BREAST CANCER

1.1.1 Definition and diagnosis

Breast cancer is the formation of malignant tumours in the breast tissue. The main symptoms of breast cancer are a lump in the breast, a change in the shape and size of the breast, nipple retraction, dimpling of the skin, or patches of red on the skin of the breast.[1] In patients with suspected breast cancer or in those with a possible breast cancer detected through screening, investigations include ultrasounds and magnetic resonance imaging (MRI) scans, but final diagnoses are normally through a biopsy. Breast cancer can be diagnosed in both men and women, but this thesis will cover only breast cancer among female patients.

1.1.2 Stage

Breast cancer staging refers to the size of the tumour, and if it has spread when diagnosed. The TNM classification is widely used to determine the stage of breast cancer at diagnosis. It uses information on the following:

- Tumour size (T)
- Spread of the cancer to lymph nodes (N)
- The presence of distant metastases (M)

The cancer is then given a stage of 0, I, II, III, or IV, with stage 0 being in situ, stage I being early locally invasive cancer, and stage IV the most advanced metastatic disease.[2]

1.1.3 Grade

The Nottingham grading system is a description of breast tumours used to indicate how quickly it is likely to grow and spread. The system grades tumours based on the following features:

- Tubule formation - how much of the tumour tissue has normal breast duct structures

- Nuclear grade - an evaluation of the size and shape of the nucleus in the tumour cells
- Mitotic rate - how many dividing cells are present

The tumour is then given a grade from 1 to 3, with grade 1 being a low grade or well-differentiated tumour, and grade 3 being a high grade or poorly differentiated tumour.[3] Well-differentiated cancer cells present like normal cells under a microscope and tend to grow and spread slower than poorly differentiated cancer cells.

1.1.4 Molecular subtypes

There are up to 21 distinct histological subtypes and at least 4 different molecular subtypes of breast cancer that differ in terms of risk factors, presentation, response to treatment, and outcomes.[4-6] Approximations of molecular subtypes use the biological markers oestrogen receptor (ER), progesterone receptors (PR), and human epidermal growth factor receptor 2 (HER2). The four main subtypes are:

- Luminal A (ER+/PR+ (ER+ or PR+) and HER2-) (73% of breast cancers [7]) - slow growing and relatively less aggressive breast cancers, which have the most favourable prognosis, partly because they are more responsive to endocrine therapy.[8, 9]
- Triple negative (ER-, PR- and HER2-) (12% of breast cancers [7]) - Poorer short-term prognosis than other subtypes, partly because there are currently no targeted therapies for these tumours.[10]
- Luminal B (ER+/PR+ and HER2+) (10% of breast cancers [7]) - Tend to be higher-grade breast cancers and are associated with poorer survival than luminal A cancers.[9]
- HER 2-enriched (ER-, PR- and HER2+) (5% of breast cancers [7]) - Grow and spread more aggressively than other subtypes and are associated with poorer short-term prognosis compared with ER+/PR+ breast cancers.[9]

1.1.5 Treatment

This section will cover the UK guidelines on breast cancer treatment, but any differences in guidelines will be taken into account when analysing data from the US. Following diagnosis of an early or locally advanced breast cancer, women typically undergo breast-conserving surgery such as lumpectomy, partial mastectomy, quadrantectomey, or a complete mastectomy. The extensiveness

of surgery is dependent on the stage, grade, and biology of the tumour; breast size; and other risk factors such as germline mutations in the BRCA genes. Whilst waiting for the BRCA status of the tumour to return, women will potentially receive neo-adjuvant systemic treatment in an attempt to shrink large cancer tumours prior to surgery, or instead of surgery if necessary. Following surgery, a decision will be made on the best adjuvant treatment for the woman based on the assessment of the prognostic and predictive factors of the cancer, and the potential benefits and side effects of treatment. The four main types of treatment are chemotherapy, biological therapy, radiotherapy, and endocrine therapy.

1.1.5.1 Chemotherapy

Chemotherapy agents typically work through inhibition of DNA and RNA synthesis, preventing the replication of rapidly growing cancer cells. The decision to offer chemotherapy to a women typically depends on the risk of recurrence of breast cancer, age, and personal preferences.[11] Women at low risk of recurrence are not recommended chemotherapy as the absolute survival benefit for chemotherapy is either unproven or extremely small. The factors contributing to what defines women at low or high risk are outlined in Table 1.1. The accepted chemotherapy regimens for women at high risk of recurrence are: FEC (-T) (5-fluorouracil, epirubicin, cyclophosphamide (-docetaxel)); EC (epirubicin, cyclophosphamide); AC (doxorubicin, cyclophosphamide); CMF (cyclophosphamide, methotrexate, 5-fluorouracil); TC (docetaxel, cyclophosphamide). The decision on whether to administer a regimen containing a taxane (docetaxel, paclitaxel) is dependent on the risk of recurrence, with the addition more likely in higher risk breast cancers. However, addition of a taxane is likely to induce additional risk of side effects such as neuropathy, neutropenia and hypersensitivity. Chemotherapy regimens are given over a period of 9 to 12 weeks (over 3 to 4 cycles).[12]

Women with more advanced disease may be administered neo-adjuvant chemotherapy to reduce tumour size prior to surgery, enabling some women to have breast-conserving surgery who would otherwise have a full mastectomy. However, only 17% of women convert from mastectomy to breast conservation after neo-adjuvant chemotherapy.[13] Furthermore, it has been suggested that neo-adjuvant chemotherapy does not confer any survival benefit compared with adjuvant chemotherapy (risk ratio (RR): 1.00, 95% CI: 0.90-1.12).[14]

Table 1.1: Overview of what constitutes a low- or high-risk breast cancer from the London Cancer Alliance guidelines [11]

Low risk (all of the following factors)	**High risk (any of the following factors)**
• Tumour ≤2cm in diameter	• Tumour ≥2cm in diameter
• Oestrogen receptor (ER) positive	• ER negative
• Grade 1 histology	• Grade 2 or 3 histology
• No lympho-vascular invasion	• Lympho-vascular invasion present
• Negative Lymph nodes	• Positive lymph nodes
• Human Epidermal Growth Factor 2 (HER2) negative	• HER2 positive
	• Younger than 35 years old

1.1.5.2 Biological Therapy

At diagnosis of breast cancer, the tumour tissue is tested for HER2 expression and women with overexpression should be considered for therapy with trastuzumab.[12] Trastuzumab is a monoclonal antibody that works by blocking the HER2 pathway, essentially stalling the growth of HER2 positive breast cancers. If required, adjuvant trastuzumab is given at 3-week intervals for one year in combination with surgery and other treatments. However, due to the known cardiotoxicities of trastuzumab, it is not recommended for use among women who have any of the following:

- Left ventricular ejection fraction (LVEF) of 55% or less
- History of documented congestive heart failure
- High-risk uncontrolled arrhythmias
- Angina pectoris requiring medication
- Clinically significant valvular disease
- Evidence of transmural infarction on electrocardiograph
- Poorly controlled hypertension

Cardiac functional assessments are carried out prior to treatment initiation and then repeated every three months during treatment. If LVEF drops by 10% from baseline or to below 50%, suspension of trastuzumab is recommended.[12]

1.1.5.3 Radiotherapy

Radiation uses high-energy x-ray beams, such as photon beams, to damage the DNA of cancerous cells, so they are unable to replicate, and stopping further tumour growth. Radiotherapy is normally offered to women after breast conserving surgery, however, in women with a very low absolute risk of recurrence (over the age of 65 years with tumours that are early stage, grade 1-2, HER2-, and ER+/PR+) there is no evidence that radiotherapy increases overall survival, so any therapy is potentially omitted. In women who have a mastectomy, radiotherapy is given to those in which the cancer has spread to the under arm lymph nodes. However, among those with lymph node negative breast cancer only those that are stage III or IV are given radiotherapy due to limited evidence on decreased recurrence (RR for breast cancer recurrence in those irradiated vs not: 1.06, 95% CI: 0.76-1.48), and worries of lung and cardiac morbidity being associated with radiotherapy due to the breasts being situated within close proximity of the heart and lungs.[15]

1.1.5.4 Endocrine Therapies

Oestrogen promotes the growth of ER+ breast cancers, and all patients with these tumours are indicated endocrine therapies to lower oestrogen, or block the effects of oestrogen on the growth of tumour cells. UK guidelines state that women with ER+ breast cancers (regardless of PR status) should be offered tamoxifen or aromatase inhibitors (AI) within 2-3 weeks after the completion of chemotherapy, or as soon as convenient after surgery if chemotherapy is not administered.[16] However, in routine practice endocrine therapies are prescribed to women with ER+/PR+ breast cancer as it has been shown that they are efficacious in certain sub-populations of with an ER- and PR+ diagnosis.[17] The remainder of this thesis will assume endocrine therapies are given to women with an ER+/PR+ breast cancer.

Tamoxifen

Tamoxifen is a selective oestrogen receptor modulator that works in breast tissue by inhibiting the growth of the tumour through competitive antagonism of oestrogen at its receptor. Premenopausal

women with invasive ER+/PR+ breast cancer are offered tamoxifen for an initial 5 years, which can be extended if the benefits of further therapy outweighs the risk of side effects such as thrombosis, endometrial cancer, and possible bone density loss.[12] Evidence suggests that 5 years of tamoxifen reduces the rate of breast cancer recurrence (RR: 0.53, 95% CI: 0.47-0.59) and breast cancer mortality (RR: 0.71, 95% CI: 0.61-0.80) in the decade following diagnosis, regardless of menopausal status.[18] However, there is no clear evidence of a benefit in disease recurrence (OR: 0.89, 95% CI: 0.76-1.05) or overall survival (OR: 0.99, 95% CI: 0.84-1.16) from continuing tamoxifen for over 5 years compared with stopping therapy after 5 years.[19] Postmenopausal women that are at low risk of disease recurrence, or those in whom AIs are not tolerated, are also offered tamoxifen.

AIs

Aromatase is responsible for the conversion of the adrenal androgen substrate androstenedione to oestrogen in the breast tissue, which is the main source of oestrogen in postmenopausal women. By inhibiting this process, AIs can reduce oestrogen production by over 90%.[20] As of 2006, when guidelines were changed, postmenopausal women with invasive ER+/PR+ breast cancer are offered AIs for an initial 5 years.[21] Prior to 2006, all women now eligible for AI therapy were offered tamoxifen. AIs are associated with a decrease in breast cancer recurrence in comparison with tamoxifen in postmenopausal women (RR: 0.78, 95% CI: 0.68-0.89).[22] Evidence also suggests that both extending therapy beyond 5 years instead of stopping, and switching to AIs after 2-3 years of tamoxifen therapy instead of continuing tamoxifen, decreases the risk of breast cancer recurrence (RR: 0.75, 95% CI: 0.65-0.86; and RR: 0.79, 95% CI: 0.65-0.92, respectively).[22, 23] Women with a high risk of recurrence are therefore recommended to continue treatment after the initial 5-year treatment, or switch from tamoxifen.

1.1.6 Epidemiology

Breast cancer remains the most common cancer worldwide, with the highest incidence in Europe, North America, and Australasia. Latest statistics reported an estimated 84,272 new cases per year in the UK, and 262,347 new cases per year in the US in 2018.[24] However, earlier detection through screening programs and the advent of new treatments has facilitated an increase in 5-year survival rate from 53% and 75% in the 1970's to 87% and 91% in the 2010's in the UK and US respectively.[25, 26]

It is suggested that 23% of all breast cancers are preventable through modification of risk

factors.[27] There is evidence that increasing body mass index (BMI), alcohol use, and tobacco use is associated with breast cancer risk,[28-30] and 8% of all cases in the UK can be attributed to obesity and alcohol use individually.[27] There is also evidence that not breastfeeding and oral contraceptive use is associated with a 22% (95% CI: 18-26%) and 24% (95% CI: 15-33%) increased risk of breast cancer respectively.[31, 32] Women with one first degree relative who had breast cancer are known to have an 85% (95% CI: 69-91%) higher risk of breast cancer in comparison with those without,[33] and women carrying the BRCA1/2 gene mutations have a 45-65% increased risk by age 70.[34] Finally, several reproductive factors such as older age at first giving birth, younger age at menarche, older age at menopause, and not having children are known to be associated with an increased risk of breast cancer.[35]

1.2 CVD

1.2.1 Definition and diagnosis

This thesis will focus on a range of clinically specific CVD outcomes. The World Health Organization (WHO) defines CVDs as a group of disorders of the heart and blood vessels. CVD can broadly be broken down into three main types: vascular disease, myocardial disease, and venous thromboembolism.

- Vascular disease – disease of the blood vessels, arteries and, and veins of the circulatory system.
- Myocardial disease – conditions which affect heart muscle but spare other anatomic structures within the cardiovascular system.
- Venous thromboembolism – blood clots in the veins, which can dislodge and move into the heart and lungs.

CVD is a disease of many different sub-types, and there is not one homogenous diagnosis. This thesis will explore 12 clinically specific CVD outcomes that are defined below. Throughout the thesis the outcomes will regularly be grouped into venous and non-venous CVD. Venous outcomes refer to all venous thromboembolic events such as deep vein thrombosis and pulmonary embolism, and non-venous outcomes refer to all other CVD outcomes described below. The reason for this binary distinction is due to the known increased risk of venous CVD outcomes associated with tamoxifen use, and the homogenous direction of effect when exploring the risk of endocrine therapy use on all

other non-venous CVD outcomes. Although grouping all non-venous CVDs is not standard within the cardiology literature, it is relevant for this study question and will allow ease of reporting within the thesis.

1.2.1.1 Angina

Angina is characterised by discomfort that occurs when myocardial oxygen demand exceeds the supply, and can generally by classified into stable and unstable angina. Unstable angina, an attack usually follows a precipitating event (climbing stairs, a heavy meal, emotional stress etc.) and has the same severity as previous attacks. The onset of stable angina is due to fixed narrowing or stenosis of the coronary blood vessels. Unstable angina occurs because of incomplete or temporary coronary artery occlusion from plaque rupture, and has increasing severity, duration, or frequency over time.[36, 37] Common symptoms of angina are chest pain, breathlessness, and nausea.

Patients with suspected stable angina normally undergo an initial clinical assessment that includes physical examination and an assessment of blood pressure, haemoglobin, thyroid function, cholesterol and glucose levels. Patients then undergo a series of diagnostic tests including a 12-lead electrocardiography (ECG), exercise tolerance testing, and myocardial perfusion scintigraphy,[38] and those with abnormal measures are diagnosed by clinicians. ECG, high-sensitivity troponin tests, and computed tomography (CT) angiography are used to diagnose unstable angina.[38]

1.2.1.2 Myocardial Infarction (MI)

Coronary artery occlusion, due to rupture of built up plaques in coronary blood vessels and resulting blockage, causes most cases of MI.[39]. Common symptoms of an MI include severe chest pain, pain in other parts of the body, dizziness, shortness of breath, and nausea.

Investigation of MI includes ECG and measurements of high sensitivity cardiac troponin, which are biomarkers indicative of tissue death. A diagnosis requires at least abnormal ECG readings, and rise or fall of cardiac troponin above the 99th percentile.[40]

1.2.1.3 Revascularisation

Following occlusion of coronary blood vessels, coronary revascularisation is the process of restoring the flow of blood to the heart. The most common procedure is coronary artery bypass grafting (CABG), where vessels from other parts of the body (usually chest, leg, or arm) are used to surgically divert blood around the occlusion. Another method of revascularisation is Percutaneous Coronary Intervention (PCI), which is a non-surgical procedure that uses a catheter to place a stent to open up blood vessels in the heart that have been narrowed by plaque build-up. In CABG, bypass grafts are placed to the mid-coronary vessel beyond the "culprit" lesion, providing extra sources of blood flow to the myocardium and offering protection against the consequences of further obstructive disease. In contrast, coronary stents aim to restore the normal flow of the coronary vessel without offering protection against new disease proximal to the stent.[41]

Revascularisation procedures can be offered to patients presenting with stable multi-vessel and/or left main coronary artery disease. Although revascularisation is essentially an intervention, it can help to capture coronary artery diseases that are otherwise not captured by general practitioner (GP) or hospital systems when identifying outcomes in routinely collected data.

1.2.1.4 Sudden Cardiac Arrest (SCA)

SCA is absent or inadequate contraction of the ventricles of the heart that immediately causes circulatory failure. If not treated immediately, SCA usually results in death. Most commonly, the sequence of events leading to arrhythmic SCA is the degeneration of ventricular tachycardia (VT) into ventricular fibrillation (VF), followed by asystole or pulseless electric activity.[42] Pre-existing coronary artery disease (acute myocardial ischemia, scarring from previous myocardial infarction, heart failure) are present in 80% of SCAs. Dilated non-ischemic and hypertrophic cardiomyopathies account for the second largest number of SCAs, whereas other cardiac disorders such as congenital heart disease and underlying genetically determined ion channel anomalies account for 5% to 10% of SCAs.[43] Signs and symptoms include loss of consciousness, rapid shallow breathing progressing to apnea, and profoundly low blood pressure.

1.2.1.5 Peripheral Vascular Disease (PVD)

PVD is facilitated through atherosclerotic plaque causing arterial stenosis or occlusion and reducing blood flow and oxygen supply to the body. Critical limb ischemia happens when reduction of blood

loss is so severe that there is pain during rest, ulceration, or gangrene.[44] Although asymptomatic in many people, some may experience intermittent claudication (muscle pain on mild exertion).

Initial exploration of PVD is through the use of the ankle-brachial index (ABI) that compares the blood pressure in the feet to the blood pressure in the arms and determines how well blood is flowing. If a patient records an ABI score of below 0.9 or above 1 .4, which is considered abnormal, they are referred for further testing through Doppler and Ultrasound imaging, CT Angiography, Magnetic Resonance Angiography, or Angiography.[45]

1.2.1.6 Stroke

Stroke is characterised as a neurological deficit attributed to an acute injury of the central nervous system by a vascular cause lasting more than 24 hours or leading to death, and can be due to cerebral infarction, intracerebral haemorrhage, and subarachnoid haemorrhage. Ischemic strokes are an episode of neurological dysfunction caused by focal cerebral, spinal, or retinal infarction, where blood flow is to the brain is restricted. Strokes caused by intracerebral haemorrhage show rapidly developing clinical signs of neurological dysfunction attributable to a focal collection of blood within the brain parenchyma or ventricular system that cannot be attributed to trauma. Stroke caused by subarachnoid haemorrhage show rapidly developing signs of neurological dysfunction and/or headache because of bleeding into the subarachnoid space that also cannot be attributed to trauma. Transient ischemic attacks are those episodes of temporary and focal dysfunction of vascular origin that leave no persistent neurological deficit and are variable in duration, commonly lasting from 2 to 15 minutes, but occasionally lasting as long as a day.[46] Symptoms of a stroke include a drooped face on one side, weak or numb arms, slurred speech, paralysis of one side of the body, dizziness, and balance or coordination problems.

CT scanning and magnetic resonance imaging of the brain are the main methods used to diagnose a stroke and identify if the source is ischemic or haemorrhagic. Although commonly measured markers include S100 calcium binding protein B or S100B, glial fibrillary acidic protein, brain natriuretic peptide, and matrix metalloproteinase-9, none of these biomarkers are routinely measured to inform diagnoses.[46]

1.2.1.7 Arrhythmia

Arrhythmias occur due to the heart rate being too fast (tachycardia, >90-100 beats per minute (bpm)), too slow (bradycardia, <50-60 bpm), or beating irregularly. The three main mechanisms of arrhythmias are: abnormal automaticity due to creation of electrical impulses by non-pacemaker cells of the myocardium because they acquire automaticity and spontaneously depolarize; triggered electrical activity which is the abnormal propagation of electrical activity in individual heart cells and are often due to problems in the ion channels in the heart muscle cells; and re-entry that happens when the electrical impulse travels backwards from the ventricles to the atria, initiating another heartbeat while the first heartbeat is still descending into the ventricles.[47] Common symptoms of arrhythmia include palpitations, dizziness, fainting, and breathlessness.

The most common method used to diagnose arrhythmias is through electrocardiograms, allowing clinical evaluation of the electrical activity of the heart within each contraction.

1.2.1.8 Heart failure (HF)

HF is the pathophysiological process in which the heart is unable to meet the metabolic requirements of the tissue for oxygen and substrates. Broadly, HF is due to abnormalities in the myocardium leading to inabilities to fulfil its function in pumping blood around the body. Systolic HF is characterised by a reduced ejection fraction (a measurement, expressed as a percentage, of how much blood the left ventricle pumps out with each contraction) and an enlarged left ventricular chamber, and is clinically associated with left ventricular failure in the presence of cardiomegaly. Diastolic HF is characterised by an increased resistance to filling with increased filling pressure, which is accompanied by pulmonary congestion with a normal or slightly enlarged ventricle.[48] Common symptoms of HF include breathlessness, fatigue, and swollen ankles and legs.

Diagnosis is through echocardiogram coupled with Doppler flow studies and measurement of B-type natriuretic peptide. Based on left ventricular ejection fraction, HF is defined as follows:[49]

1. HF with reduced ejection fraction: symptoms and signs with LVEF <40%.
2. HF with mid-range ejection fraction: symptoms and signs with LVEF 40% to 49%. Other features include elevated natriuretic peptides (B-type natriuretic peptide (BNP) >35 nanograms/L (>35 picograms/mL) or N-terminal pro-brain natriuretic peptide (NT-pro-BNP) >125 nanograms/L (125 picograms/mL)) and at least one additional criterion: (a) relevant

structural heart disease (e.g., left ventricular hypertrophy (LVH) or left atrial enlargement), (b) diastolic dysfunction.

3. HF with preserved ejection fraction: symptoms and signs with LVEF >50%. Other features include elevated natriuretic peptides (BNP >35 nanograms/L (>35 picograms/mL) or NT-pro-BNP >125 nanograms/L (>125 picograms/mL)) and at least one additional criteria: (a) relevant structural heart disease (e.g., LVH or left atrial enlargement), (b) diastolic dysfunction.

1.2.1.9 Pericarditis

The pericardium is a fibroelastic sac made up of visceral and parietal layers separated by the pericardial cavity. In healthy individuals, the pericardial cavity contains 15-50 ml of an ultrafiltrate of plasma. Pericarditis is the inflammation of the pericardial sac. Common symptoms include sharp chest pain, breathlessness, palpitations, fever, and fatigue.

Diagnoses are normally made through electrocardiogram, chest radiography, biomarker count (troponin level, erythrocyte sedimentation rate, and C-reactive protein level), and echocardiogram. Pericarditis can be diagnosed with at least two of: pericarditic chest pain; pericardial rubs; new widespread ST-elevation or PR depression on electrocardiogram; or pericardial effusion. Additional supporting findings include elevation of the biomarkers outlined above. Pericarditis can be acute (<4 weeks), incessant (>4 weeks but <3 months, without remission), recurrent (recurrence of an episode with a symptom-free interval of >4 weeks), or chronic (lasting longer than 3 months).[50]

1.2.1.10 Valvular Heart Disease (VHD)

VHD is caused by either damage or defect in one of the four heart valves (aortic, mitral, tricuspid or pulmonary), which can be either congenital or acquired. The dominant functional and anatomic consequences of VHD are stenosis and insufficiency. Stenosis is a mild thickening of the valve, which results in severe impairment of the valve motion, and insufficiency describes the inability of the valve to close properly and prevent back flow of blood.[51] Common symptoms of VHD are breathlessness, swelling of feet and ankles, and tiredness.

Echocardiography is the key technique used to confirm a diagnosis of all types of VHD,[52] allowing assessment of the appearance and mobility of valves. Other techniques consist of Doppler

parameters including peak velocity, mean gradient, and effective orifice area; aortic size measurement; left ventricular geometry and function; pulmonary artery size and pressure; and colour mapping of the valves.

1.2.1.11 Deep Vein Thrombosis (DVT)

DVT refers to one or more thrombi in one of the body's large veins, most commonly in the lower limbs, that cause partial or complete blocking of the circulation of the vein. Virchow's triad described the three broad categories that contribute to thrombosis: venous stasis, vascular injury, and hypercoagulability. The presence of venous stasis and either of the other two factors greatly increases the risk of clot formation. VTE regularly occurs in areas with decreased or altered blood flow, such as pockets adjacent to valves in the deep veins of the leg.[53] Common symptoms of DVT include pain in legs, and warm or red skin around affected area.

Diagnosis of DVT is initiated though a d-dimer test, that allows people who are unlikely to have DVT to be ruled out. Doppler venous ultrasound is then the first line DVT imaging modality and allows detection of thrombi. If a diagnosis cannot be confirmed, patients then undergo venous CT scans.[54]

1.2.1.12 Pulmonary Embolism (PE)

Pulmonary embolism occurs when a deep vein thrombosis breaks free, passes through the right side of the heart, and lodges in the pulmonary arteries, resulting in partial or complete occlusion.[55] Evidence of DVT in the legs is found in around 70% of patients with PE. Patients with PE most regularly present with dyspnoea with or without pleuritic pain and haemoptysis (acute minor pulmonary embolism). The second most common presentation is haemodynamic instability, which is associated with acute massive pulmonary embolism. The third and least common presentation mimics heart failure or indolent pneumonia, especially in the elderly.[56] Common symptoms of PE include breathlessness and chest pain.

D-dimer tests are again used to rule out a PE diagnosis, and CT scans are the primary diagnostic test for PE, allowing thrombi to be identified in the pulmonary arteries.[57]

1.2.2 Treatment

Drug treatment for CVD has two main aims: relieve symptoms of the diagnosis, and reduce the risk of recurrence or worsening. Treatments can also be used for CVD prevention measures. The main drug treatments are:

- Angiotensin-converting enzyme (ACE) inhibitors – decrease blood pressure by reducing the activity of the renin–angiotensin–aldosterone system through blocking the conversion of Angiotensin I to Angiotensin II.[58]
- Statins – reduce cholesterol by competitively inhibiting HMG-CoA reductase.[59]
- Angiotensin-II antagonists – decrease blood pressure by blocking the activation angiotensin II receptors, causing vasodilation.[60]
- Beta blockers – decrease blood pressure by reducing the heart rate through inhibition of β-adrenergic receptors.[61]
- Calcium channel blockers – decrease blood pressure by disrupting the movement of calcium though calcium channels.[62]

1.2.3 Epidemiology

CVD causes 17.9 million deaths worldwide every year, which accounts for 31% of all global deaths.[63] There are around 7 and 92 million people living with CVD in the UK and US respectively.[64, 65] Although CVD continues to be large burden on the health of society, age standardised mortality from CVD has fallen 70% over the last 30 years in the UK, from 1,152 per 100,000 in 1979 to 333 per 100,000 in 2013.[66]

High blood pressure, increasing BMI, high cholesterol levels, unhealthy diet, physical inactivity, and tobacco and alcohol use are known to be associated with an increased risk of CVD, with 49% of all coronary artery diseases attributable to high blood pressure, 56% to high cholesterol, 21% to high BMI, 31% to unhealthy diet, 22% to physical inactivity, 12% to tobacco, and 2% to alcohol.[67] Overall, it is estimated that 83–89% of all coronary heart diseases and 70-76% of all strokes could be prevented and are attributable to the joint effects of these modifiable risk factors.[67] There is an increasing prevalence of CVD as age increases,[66] and evidence also suggests that type 2 diabetes is associated with an increased risk of all CVD outcomes other than arrhythmia and sudden cardiac

arrest.[68] A CVD diagnosis in at least one parent is also known to be associated with an increased risk of the disease.[69]

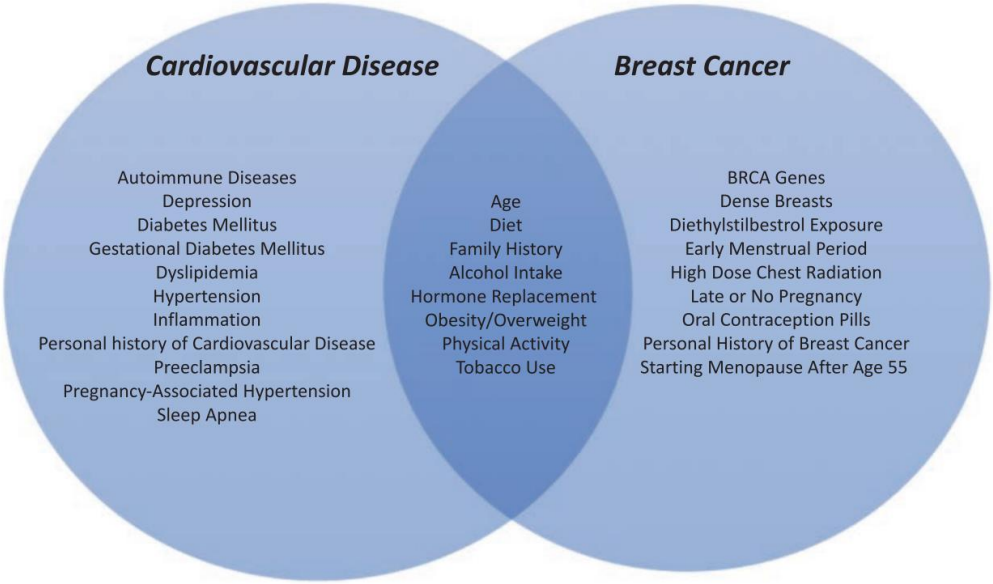
1.3 INTERSECTION OF BREAST CANCER AND CVD

In older women diagnosed with breast cancer, there is a greater likelihood of dying of diseases other than cancer itself, with CVD being the most frequent other cause.[70, 71] The risk of CVD is also higher in women with a breast cancer diagnosis compared with those without.[72] This overlap between breast cancer and risk of CVD is likely to be due to either similar risk factors, cardiotoxicity of treatments, or a combination of the two.

1.3.1 Risk factors

CVD may be a particular problem for those breast cancer survivors with existing CVD risk factors as many of these overlap with the risk factors of breast cancer (Figure 1.1). Uptake of a number of behaviours aimed at increased CVD health, such as diet, activity, and smoking, have also been shown to be associated with a lower incidence of breast cancer.[73] Lifestyle changes can therefore be a means by which women diagnosed with breast cancer are able to reduce the risk of breast cancer recurrence, as well as the risk of CVD.

Figure 1.1: Risk factors for CVD and breast cancer [74]



1.3.2 Cardiotoxicity of treatments

As women are living longer following a diagnosis of breast cancer, there are more opportunities for any medium- to long-term adverse effects of treatments to manifest, be those effects of individual therapies or regimens incorporating multiple agents. Evidence suggests both early and delayed toxic effects of several breast cancer therapies. For example, there is known to be no safe dose anthracycline (doxorubicin, epirubicin), as even women administered the lower doses of the drug are at risk of experiencing cardiotoxicity such as heart failure, arrhythmias, atrial fibrillation, and left ventricular dysfunction.[75-77] Administration of the taxane paclitaxel has been shown to be associated with bradycardia, with patients in a clinical trial experiencing heart rates of <40 bpm.[78] MI, HF, and arrhythmias have been reported as side effects of 5-fluorouracil, but evidence is not conclusive,[79-81] and any effects are thought to be acute, with long-term cardiotoxicity uncommon.[82] There is also a known association between trastuzumab, used in women with HER2+ breast cancer, and heart failure. However, this toxic effect is mostly reversible.[83] Finally, there was an excess of non–breast cancer deaths after 5 years among patients receiving radiotherapy, with CVD being one of the main causes of death.[84] Left-sided breast cancer patients treated with radiotherapy have a 90% (95% CI: 52-137%) increased risk of CVD mortality over 20 years compared with those treated with radiotherapy for right-sided breast cancer.[85]

This thesis will focus on the cardiotoxicities associated with endocrine therapy use. Several systematic reviews and meta-analyses have attempted to map out the risk of CVD associated with use of tamoxifen, AIs, and the comparative risk between the two drugs.[86-91] The systematic review in Chapter 2 fully reviews all currently available evidence, but broadly, these reviews have suggested that tamoxifen use is associated with an increased risk of venous thromboembolic events such as DVT, and potentially a decreased risk of coronary heart disease outcomes in comparison with both AI use and no tamoxifen or placebo. Little conclusive evidence is available on the associations between AI use and CVD. Many of the randomised controlled trials (RCT) that have explored the cardiotoxicity associated with AI use have included CVD events as a secondary consideration, with breast cancer recurrence being the primary outcome. Several observational studies have also explored this association, but the majority of these studies have used composite CVD outcome definitions, which do not take into account the clinical nuances between specific CVDs.

It is known that tamoxifen use lowers total serum cholesterol by 10-20% and low-density lipoprotein

levels by 15-22%, but there is no known effect on high-density lipoprotein cholesterol.[92-95]. The reduction in cholesterol may therefore explain any protective effects of tamoxifen on CVD. However, the oestrogen agonistic effects of tamoxifen and resulting increase in thrombogenicity, through a reduction in antithrombin and protein C levels,[96, 97] could also explain a toxic effect on the risk of venous thromboembolic events. Furthermore, AIs work by inhibiting the aromatase enzyme and depleting oestrogen levels, which are known to be protective of CVD. There is therefore a biological explanation for any increased risk of CVD associated with AI use.

Fully understanding the cardiotoxicities of AIs is increasingly important in post-menopausal women as guideline changes in 2006 suggested that these women are primarily given AIs rather than tamoxifen. A full understanding of the risk of clinically specific CVDs associated with AI in comparison with tamoxifen use, as well as the risks associated with both treatments alone, is crucial in understanding the full safety profile of endocrine therapies in post-menopausal women.

1.4 AIMS AND OBJECTIVES

The overall aim of this thesis is to assess the effect of endocrine therapy use on the risk of a range of CVD outcomes in female breast cancer survivors. The specific objectives are:

1. Review RCT and observational evidence regarding the association between endocrine therapy use and a range of clinically specific CVD outcomes in female breast cancer survivors.
2. Explore codes and algorithms to define and detect CVD outcomes in primary and secondary care in routinely collected UK data.
3. Assess the effect of AI in comparison with tamoxifen use on the risk of a range of clinically specific CVD outcomes in a cohort of post-menopausal female breast cancer survivors in the UK.
4. Assess the effect of AI use, tamoxifen use, and the comparative effect between use of the two endocrine therapies on the risk of a range of clinically specific CVD outcomes in a cohort of post-menopausal female breast cancer survivors in the US.
5. Compare the methodology, results, and conclusions between the studies in the UK and US, then modify the methodology of both studies until they are as similar as possible, and compare modified results.

1.5 OUTLINE OF THESIS

This thesis follows the research paper style format, with articles incorporated into chapters, and other chapters written in a more traditional style. Three articles have been written, one of which has been published, and two that are formatted as pre-submission manuscript drafts. The thesis will be organised as follows:

- Chapter 2 – a published systematic review of long-term adjuvant endocrine therapy use and the risk of cardiovascular disease in female breast cancer survivors.
- Chapter 3 – an overview of the data sources used for the original analyses in the thesis.
- Chapter 4 – investigation into how CVD outcomes are captured in UK primary and secondary care data, followed by development and assessment of CVD outcome definitions.
- Chapter 5 – final draft of a paper assessing the effect of endocrine therapy use on the risk of CVD in the UK.
- Chapter 6 - final draft of a paper assessing the effect of endocrine therapy use on the risk of CVD in the US.
- Chapter 7 – comparison of the UK and US studies, and analysis in which both studies are modified to make them as similar as possible
- Chapter 8 – summary of the main results in the context of what is already known, discussion of the strengths and limitations of the studies, and implications for clinical practice and future work.

1.6 SUMMARY

- Breast cancer is the most common cancer worldwide, but advances in detection and treatment has meant there is now a relatively high five-year survival rate.
- Breast cancer is the formation malignant tumours in the breast tissue and has a variety of subtypes based on stage, grade, and biology of the tumour.
- The main treatment modalities are chemotherapy, biological therapy, radiotherapy, and endocrine therapy.
- Over 80% of breast cancers are defined as ER+/PR+
- Endocrine therapies (tamoxifen and AIs) are used to reduce the recurrence of ER+/PR+ breast cancer for at least five years following surgery.
- Breast cancer survivors may be at a higher risk of CVD, a group of disorders of the heart and blood vessels, due to shared risk factors, cardiotoxicity of treatments, or a combination of the two.
- Previous evidence has suggested that tamoxifen use is associated with an increased risk of venous thromboembolic events, and potentially a decreased risk of coronary heart disease outcomes in comparison with both AI use and no tamoxifen or placebo. Little conclusive evidence is available on the associations between AI use and CVD.
- This thesis will investigate the effects of tamoxifen and AI use on a range of clinically specific CVD outcomes, first through a review of existing evidence, then through analysis of data from the UK and US.

2 LONG TERM ADJUVANT ENDOCRINE THERAPY AND RISK OF CARDIOVASCULAR DISEASE IN FEMALE BREAST CANCER SURVIVORS: SYSTEMATIC REVIEW

This chapter includes a systematic review that collated all available evidence on the risk of cardiovascular disease associated with endocrine therapy used in female breast cancer survivors. This review was published in the British Medical Journal in October 2018. Online appendices that were published alongside the review are available at the end of the thesis.

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	493367	**Title**	Mr
First Name(s)	Anthony		
Surname/Family Name	Matthews		
Thesis Title	Adjuvant endocrine therapy use in female breast cancer survivors and the risk of cardiovascular disease		
Primary Supervisor	Krishnan Bhaskaran		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	British Medical Journal (evidence of copyright retention - http://bit.ly/bmjcopyright)		
When was the work published?	October 2018		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I did the initial literature search, as well as the paper selection and data extraction. This process was replicated by two coauthors. I wrote the first draft. All authors contributed to further drafts and approved the final manuscript.

SECTION E

Student Signature	[redacted]
Date	12/01/2019

Supervisor Signature	[redacted]
Date	12/01/2019

Long term adjuvant endocrine therapy and risk of cardiovascular disease in female breast cancer survivors: systematic review

Anthony Matthews,[1] Susannah Stanway,[2] Ruth E Farmer,[1] Helen Strongman,[1] Sara Thomas,[1] Alexander R Lyon,[3,4] Liam Smeeth,[1] Krishnan Bhaskaran[1]

[1]Department of Non-Communicable Diseases Epidemiology, London School of Hygiene and Tropical Medicine, London, UK
[2]Royal Marsden Hospital, London, UK
[3]Faculty of Medicine, Imperial College London, London, UK
[4]Royal Brompton Hospital, London, UK

Correspondence to: A Matthews
anthony.matthews@lshtm.ac.uk

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2018;363:k3845
http://dx.doi.org/10.1136/bmj.k3845

Accepted: 23 August 2018

ABSTRACT

OBJECTIVE

To investigate the effect of endocrine therapies on a wide range of specific clinical cardiovascular disease outcomes in women with a history of non-metastatic breast cancer.

DESIGN

Systematic review and meta-analysis of randomised controlled trials and observational studies.

DATA SOURCES

Medline and Embase up until June 2018.

ELIGIBILITY CRITERIA FOR SELECTING STUDIES

Studies were included if they investigated the risk of a specific cardiovascular disease outcome associated with use of either tamoxifen or an aromatase inhibitor, or compared the two treatments, in women with a history of non-metastatic breast cancer.

APPRAISAL AND DATA EXTRACTION

Relevant studies were originally identified and results extracted by one researcher, with a full replication of the study identification process by a combination of two other researchers. The Cochrane Collaboration's tool for assessing risk of bias was used to assess risk of bias in randomised controlled trials, and this tool was adapted to assess risk of bias in observational studies.

RESULTS

26 studies were identified, with results for seven specific cardiovascular disease outcomes (venous thromboembolism, myocardial infarction, stroke, angina, heart failure, arrhythmia, and peripheral vascular disease). Results suggested an increased risk of venous thromboembolism in tamoxifen users compared with both non-users and aromatase inhibitor users. Results were also consistent with a higher risk of the vascular diseases myocardial infarction and angina in aromatase inhibitor users compared with tamoxifen users, but there was also a suggestion that this may be partly driven by a protective effect of tamoxifen on these outcomes. Data were limited, and evidence was generally inconsistent for all other cardiovascular disease outcomes.

CONCLUSION

This review has collated substantial randomised controlled trial and observational evidence on the effect of endocrine therapies on several specific cardiovascular disease outcomes including venous thromboembolism and myocardial infarction, progressing knowledge. Although the choice of aromatase inhibitor or tamoxifen will primarily be based on the effectiveness against the recurrence of breast cancer, this review shows that the individual patient's risk of venous or arterial vascular disease should be an important secondary consideration.

SYSTEMATIC REVIEW REGISTRATION

Prospero CRD42017065944.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Several meta-analyses of randomised controlled trials have reported the effect of endocrine therapies used in adjuvant treatment of breast cancer on the risk of composite cardiovascular disease outcomes

However, these reviews have not reported the effect of endocrine therapies on a range of clinically specific cardiovascular diseases

They have also omitted the growing body of evidence from observational studies, which often include large study populations in real world settings, as well as longer follow-up

WHAT THIS STUDY ADDS

Observational evidence is generally consistent with trial evidence reporting an increased risk of venous thromboembolism in tamoxifen users compared with both non-users and aromatase inhibitor users

Evidence also exists of a higher risk of vascular disease in aromatase inhibitor users compared with tamoxifen users, which may be driven by a protective effect of tamoxifen

Introduction

Endocrine therapies—namely, tamoxifen and aromatase inhibitors—reduce the risk of reoccurrence of breast cancer in patients diagnosed as having oestrogen receptor and/or progesterone receptor positive breast cancer following surgery (adjuvant treatment). The efficacy of tamoxifen, irrespective of menopausal status, has been confirmed in several randomised controlled trials,[1] but UK guidelines were changed in 2006 to reflect the evidence that aromatase inhibitors are more efficacious in postmenopausal women.[2] Concerns exist that endocrine therapies could increase the risk of cardiovascular disease—for example, through suppression of the cardiovascular protective effects of oestrogens.[3] With improved survival after breast cancer, cardiovascular disease has become an increasingly important source of long term morbidity and mortality among breast cancer survivors.[4] Understanding any associations between treatment of cancer and risk of cardiovascular disease is critical to inform prevention and management of adverse cardiovascular effects.

Several systematic reviews and meta-analyses of randomised controlled trials,[2 5-9] and some non-systematic reviews,[10-13] have compared cardiotoxicities of endocrine therapies in breast cancer survivors (systematic reviews summarised in appendix 1). Several reported a higher incidence of adverse cardiovascular disease outcomes in users of aromatase

inhibitors compared with tamoxifen, but results were not universally in agreement. The most recent meta-analysis suggested a 19% higher risk of a composite of cardiovascular disease outcomes, excluding venous thromboembolism, in users of aromatase inhibitors compared with tamoxifen but hypothesised that this may reflect the cardioprotective effects of tamoxifen.[5] Important limitations of the randomised controlled trial evidence included in these reviews may have contributed to the mixed picture, including high degrees of trial heterogeneity, limited power of individual trials, and inconsistent reporting of cardiovascular disease outcomes in trials focusing on anticancer effects. Previous reviews have also mainly reported results for composite cardiovascular disease outcomes rather than clinically specific cardiovascular diseases and omitted the growing body of evidence from observational studies on this topic, which often include large study populations in real world settings and longer follow-up.

The aims of this systematic review were to identify and summarise both randomised controlled trial and observational evidence on associations between endocrine therapies and a wide range of specific clinical cardiovascular disease outcomes in women with a history of early breast cancer, to describe the differences between findings from randomised controlled trials and real world observational studies, and to assess the quality and potential for bias in studies investigating this topic.

Methods

Inclusion criteria

We included randomised controlled trials and observational studies if they carried out at least one analysis assessing the risk of a specific cardiovascular disease outcome associated with tamoxifen, aromatase inhibitors, or a comparison of the two treatments after the diagnosis of non-metastatic breast cancer in women. The outcomes of interest were vascular disease—angina, myocardial infarction, revascularisation procedures, sudden cardiac arrest, stroke (haemorrhagic and ischaemic), and peripheral vascular disease; myocardial disease—cardiomyopathy, heart failure, and arrhythmia; venous thromboembolism; pericarditis; and valvular heart disease.

We excluded studies if only a composite cardiovascular disease outcome or mortality from cardiovascular disease was assessed, only women with metastatic breast cancer were included in the study population, or the study exclusively analysed temporal differences for the same treatment on the risk of cardiovascular disease. We also excluded previous systematic reviews and meta-analyses exploring the cardiotoxicities of systemic breast cancer therapies (specifically endocrine therapies), but we included relevant randomised controlled trials captured in these reviews that were not captured in the main search, along with any more recent or previously unidentified trials. We also manually searched all randomised controlled trials of endocrine therapy for breast cancer published since the most recent systematic review to ensure that more recent trial papers were not missed.

Search strategy and data extraction

We used the health and medical literature databases Medline and Embase to search for relevant publications. The searches were performed in June 2018. Conference abstracts, grey literature, and unpublished studies were not included. To identify all relevant literature, the search strategy for each database included a comprehensive list of both index and free text terms for breast cancer, endocrine therapies, and cardiovascular disease. The full search terms used are outlined in appendix 2. We manually searched the reference lists of all studies identified in the search to further identify relevant studies that were originally missed.

We extracted relative risks, odds ratios, or hazard ratios if they were calculated in the paper. We calculated the relative risk and 95% confidence interval if effect estimates were not presented but data on the number of outcome events in follow-up allowed their calculation. We also collated information on the country in which the study was based, study type (randomised controlled trial or observational), data source (if an observational study), study design (if an observational study), age of included patients, inclusion criteria, exclusion criteria, intervention arm (and number of patients in the arm), reference arm (and number of patients in the arm), primary endpoint, cardiovascular disease outcome(s) assessed, mean/median follow-up time, statistical methods, and covariates adjusted for.

One researcher (AM) identified all relevant studies from the original literature search and extracted the results of these studies. The study identification process was repeated by a combination of two other authors (RF and HS). The extraction table was also piloted on two studies by two researchers (AM and RF) to check reproducibility of key information.

Risk of bias assessment

We used the Cochrane Collaboration's tool for assessing risk of bias to assess risk of bias in randomised controlled trials.[14] We then adapted this tool to produce separate risk of bias assessments for cohort and case-control studies, with domains for each type of bias that could be encountered within both observational study designs (appendices 3 and 4).

Statistical analysis

We organised seven possible comparisons between study arms/exposures during follow-up into three groups defined a priori to aid presentation. Group 1 included the direct comparison between aromatase inhibitor use and tamoxifen use during follow-up. Group 2 included three comparisons, all characterised as addition of tamoxifen in the intervention arm during follow-up (tamoxifen versus placebo, tamoxifen versus no tamoxifen, sequenced therapy (tamoxifen followed

doi: 10.1136/bmj.k3845 | *BMJ* 2018;363:k3845 | the bmj

by aromatase inhibitor or vice versa) versus aromatase inhibitor). Group 3 also included three comparisons, with addition of aromatase inhibitor in the intervention arm during follow-up (aromatase inhibitor versus placebo, aromatase inhibitor versus no aromatase inhibitor, sequenced therapy versus tamoxifen).

To investigate differences in study findings for the same cardiovascular disease outcome, study type (randomised controlled trial or observational), and comparison (for the seven possible comparisons outlined above), we used I^2 tests and P values for Cochrane Q tests to assess heterogeneity.[15] We considered an I^2 value of above 25% to be evidence of between study heterogeneity.[16] We used a fixed effect meta-analysis to combine individual study effects estimates if there was more than one study and no evidence of heterogeneity within cardiovascular disease outcome, study type, and comparison strata. If we found evidence of between study heterogeneity, we assessed studies within the same strata for differences in study population, statistical analysis methods, and covariate adjustments, but they were not meta-analysed. For the purposes of exploring heterogeneity and meta-analysing results, we considered randomised controlled trials and observational studies separately. We used a test for funnel plot asymmetry to examine publication bias if there were more than 10 studies within the same cardiovascular disease outcome, study type, and comparison strata.[17]

Patient involvement

No patients were involved in setting the research question or the outcome choices, nor were they involved in developing plans for design or implementation of the study. No patients were asked to advise on interpretation or writing up of results.

Results

Figure 1 outlines the screening process. We included 26 studies after applying the inclusion and exclusion criteria.[18-44] Six previous meta-analyses or systematic reviews of randomised trials were also identified.[2 5-9] We identified 12 individual randomised controlled trials that met the inclusion criteria from within these meta-analyses and included them in our review. One further study was identified from scanning the reference lists of the other papers. The final 26 included studies consisted of 15 randomised controlled trials and 11 observational studies. Table 1 summarises the included studies, with a more detailed breakdown in appendix 5. There were minimal discrepancies between authors in the duplication of the search strategy.

The most commonly investigated outcomes were venous thromboembolism (n=15), myocardial infarction (n=14), and stroke (n=12). Arrhythmia and peripheral vascular disease were each investigated in a single study. Studied outcomes did not include revascularisation, sudden cardiac arrest, cardiomyopathy, pericarditis, or valvular heart disease.

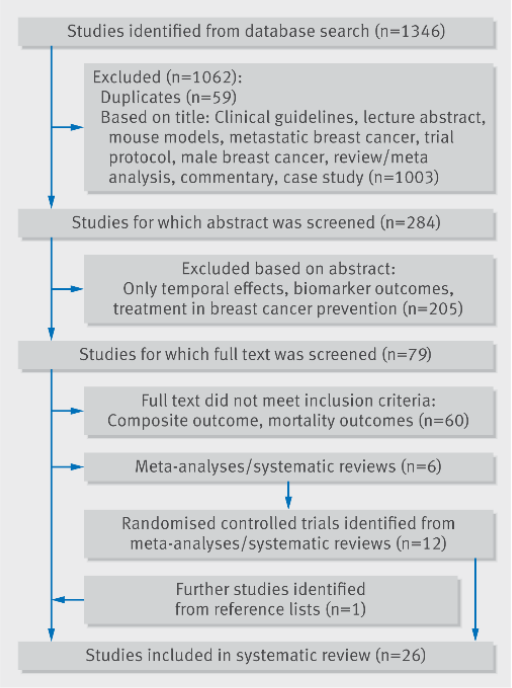


Fig 1 | Flow diagram of screening process of studies included in systematic review. RCT=randomised controlled trial

Bias assessment

Table 2 and table 3 show an overview of the risk of bias assessment of all randomised controlled trials and observational studies, with more detailed information in appendices 6 and 7. The main problem when assessing bias in randomised controlled trials was the incomplete reporting of methods, which in many cases made fully judging whether studies were prone to certain biases impossible. Three of the 15 randomised controlled trials were open label trials, and so were at higher risk of performance bias. Only one randomised controlled trial reported sufficient information to assess potential selective reporting of cardiovascular disease outcomes.

All observational studies had at least one domain categorised as being at high risk of bias. Six out of the 11 studies had a high risk of bias for the methods used to define exposure, which was mostly owing to not requiring women to have a minimum exposure period or several prescriptions before being categorised as exposed, raising the possibility of exposure misclassification. Risk of bias due to residual confounding was also present across both cohort and case-control observational studies. Seven studies adjusted only for basic risk factors and did not consider cardiovascular disease related treatment, cancer severity at diagnosis, or other cancer treatments such as chemotherapy or biological therapy.

We did not assess publication bias because no cardiovascular disease outcome, study type, or comparison strata included more than 10 studies.

Table 1 | Overview of characteristics of studies included in systematic review. Values are numbers (percentages)

	Value (n=26)
Study type	
Randomised controlled trial	15 (58)
Observational	11 (42)
Case-control	4 (15)
Cohort	7 (27)
Country/region	
North America	8 (31)
Canada	2 (8)
USA	5 (19)
USA and Canada	1 (4)
Europe	11 (42)
Denmark	2 (8)
Germany	1 (4)
Italy	1 (4)
Scotland	1 (4)
Sweden	1 (4)
UK	3 (12)
Europe-wide	2 (8)
Rest of world	3 (12)
Taiwan	2 (8)
Egypt	1 (3)
International	4 (15)
Study population	
<80 years old	1 (4)
<70 years old	1 (4)
35-70 years old	1 (4)
45-69 years old	2 (8)
All women	7 (27)
Postmenopausal	13 (50)
Premenopausal	1 (4)
Year of study	
Before 2000	4 (15)
2000-10	13 (50)
After 2010	9 (35)
Outcomes*	
Vascular disease	
Myocardial infarction	14 (54)
Stroke	12 (46)
Angina	4 (15)
Peripheral vascular disease	1 (4)
Myocardial disease	
Heart failure	4 (15)
Arrhythmia	1 (4)
Thromboembolic events	15 (58)

*Individual studies often included more than one outcome.

Vascular disease

Figure 2 shows relative risks and 95% confidence intervals for all vascular disease outcomes. Three of the four randomised controlled trials and one observational study that directly compared aromatase inhibitors with tamoxifen showed increased risks of myocardial infarction in the aromatase inhibitor group, with relative risks ranging from 1.50 to 2.29.[18 30 32 38 44] However, the effect was statistically significant only in the observational study and one randomised controlled trial. Most (five out of eight) of the studies that explored the addition of tamoxifen observed a lower risk of myocardial infarction in the tamoxifen group,[19 22 24 26 27 35 36 43] including one trial and two observational studies that found a significantly protective relative risk. Although 12 studies explored the risk of stroke in users of endocrine therapy,[21 24 26 27 33 38-44] the picture was much more mixed and included estimates in both directions. Three of the five observational studies that compared tamoxifen use with non-tamoxifen use suggested a decreased risk of stroke in tamoxifen users.[21 24 26 43] Furthermore, the results for angina were consistent with patterns seen for myocardial infarction, but only four studies explored this outcome.[19 28 41-43] However, one randomised controlled trial reported an increased risk of angina in aromatase inhibitor users compared with placebo users (relative risk 1.35, 95% confidence interval 1.17 to 1.56). Finally, only one inconclusive randomised controlled trial explored the risk of peripheral vascular disease in aromatase inhibitor users compared with tamoxifen users.[30]

Myocardial disease

Figure 3 shows relative risks and 95% confidence intervals for all myocardial disease outcomes. One randomised controlled trial suggested an increased risk of heart failure in aromatase inhibitor users compared with tamoxifen users (aromatase inhibitor versus tamoxifen: relative risk 1.20, 1.04 to 1.38),[39] but this was not replicated in an observational cohort study.[42] A fixed effects meta-analysis based on two observational studies pointed towards a decreased risk of heart failure in tamoxifen users compared with non-users, albeit with a wide confidence interval (relative risk 0.84, 0.65 to 1.07; I^2=0; Cochrane T test P=0.33; appendix 8),[42 43] which was replicated in a randomised controlled trial.[36] One inconclusive study explored the risk of arrhythmia in tamoxifen users compared with non-users.[36]

Venous thromboembolism

Figure 4 shows relative risks and 95% confidence intervals for all venous thromboembolism outcomes. Five out of six randomised controlled trials directly comparing the risk of venous thromboembolism in aromatase inhibitor users versus tamoxifen users estimated large protective relative risks ranging from 0.25 to 0.61,[28 29 32 38 39 44] with one further randomised controlled trial finding no association (I^2=0%, Cochrane Q test P=0.70, fig 4). A fixed effects meta-analysis (appendix 9) suggested a decreased risk of thromboembolic events in aromatase inhibitor users compared with tamoxifen users (relative risk 0.61, 0.58 to 0.63). Five randomised controlled trials in which the key difference between treatment arms was use of tamoxifen reported an increased risk of thromboembolic events in the tamoxifen arm, with relative risks ranging from 1.06 to 4.49.[27 31 35-37] The three randomised controlled trials comparing tamoxifen with placebo reported an increased risk of venous thromboembolism in tamoxifen users, but with 95% confidence intervals that crossed the null association (I^2=45%; Cochrane Q test P=0.16).[31 35 37] A further three observational studies compared the risk of thromboembolic events in tamoxifen users and non-tamoxifen users (I^2=92%, Cochrane Q test P=0.00).[20 23 25] Two found large increased risks in tamoxifen users (relative risks 2.40, 1.60 to 3.40,

doi: 10.1136/bmj.k3845 | BMJ 2018;363:k3845 | the bmj

Table 2 | Risk of bias assessment overview: observational studies

Paper	Study design	Exposure definition	Outcome/case definition	Control selection	Confounding	Missing data	Censoring
Abdel-Qadir 2016	Cohort	High	High	NA	Low	Unknown	Low
Chen 2014	Cohort	High	Low	NA	High	Unknown	Low
Haque 2016	Cohort	High	Low	NA	Low	Low	Low
Hernandez 2008	Cohort	Unknown	Low	NA	Low	Unknown	Low
Hernandez 2009	Cohort	Unknown	Low	NA	Low	High	Low
Ligibel 2012	Cohort	High	Low	NA	High	Unknown	Low
Yang 2014	Cohort	High	Low	NA	High	Unknown	Unknown
Bradbury 2005	Case-control	High	High	Low	High	Low	NA
Geiger 2004	Case-control	Low	Low	Low	High	High	NA
Geiger 2005	Case-control	Low	Low	Low	High	High	NA
Meier 1998	Case-control	Low	Low	Low	High	High	NA

NA=not applicable

and 7.10, 1.50 to 33.00[23 25]), and one reported no difference in risk of thromboembolic events (0.94, 0.78 to 1.13[20]), although this study had a high risk of bias owing to how exposure was defined. One randomised controlled trial reported an increased risk of thromboembolic events in aromatase inhibitor users compared with those given a placebo (relative risk 1.84, 1.11 to 3.04).[41]

Discussion

Among 26 studies providing data on seven specific cardiovascular disease outcomes, we found consistent evidence of an increased risk of venous thromboembolism in tamoxifen users compared with non-users in both randomised controlled trials and observational studies, with a correspondingly decreased risk of venous thromboembolism when aromatase inhibitor users were compared directly with tamoxifen users. However, the direct effect of aromatase inhibitors on venous thromboembolism was less clear, as only a single randomised controlled trial compared aromatase inhibitor with placebo, finding an increased risk in aromatase inhibitor users. The evidence on the effects of endocrine therapies on vascular disease risks was mixed: most studies were consistent with a higher risk of myocardial infarction and angina in aromatase inhibitor users compared with tamoxifen users, and there was a suggestion that this may be partly driven by a protective effect of tamoxifen on these outcomes; inconsistent results were found for the associations with stroke. Of the few studies assessing other outcomes, data were limited and very mixed patterns were observed, making drawing conclusions difficult.

Quality and limitations of evidence

Thirteen of the 15 randomised controlled trials identified disease-free survival as the primary outcome of the study, whereas all observational studies identified either one or several specific cardiovascular disease events as their primary outcome. Women with previous cardiovascular disease were therefore excluded from many observational studies but not from randomised controlled trials. Overall, women included in the randomised controlled trial populations were therefore likely to be at a higher absolute risk of cardiovascular disease during follow-up, which would be problematic only if the prevalence of cardiovascular disease at baseline was different between the treatment arms. In theory, randomisation should result in an equal prevalence of cardiovascular disease at baseline between arms. However, many included studies did not report information on randomisation and concealment of allocation, meaning that selection bias in relation to prevalent cardiovascular disease at the point of randomisation cannot be disregarded. Randomised controlled trials also did not report data on cardiovascular disease risk at baseline, but participants in trials are likely to be healthier than the general population and thus may have had less previous cardiovascular disease. As

Table 3 | Risk of bias assessment overview: randomised controlled trials

Paper	Random sequence generation	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other sources of bias
Bliss 2012	Low	Unknown	Low	Low	Low	Low
Boccardo 2006	Unknown	Unknown	Unknown	Low	Unknown	Low
Coombes 2007	Low	Unknown	Low	Low	Unknown	Low
Fisher 1999	Unknown	Unknown	Unknown	Low	Unknown	Low
Fisher 2001	Low	Unknown	Unknown	Low	Unknown	Low
Forbes 2008	Low	Low	Unknown	Unknown	Unknown	Low
Jakesz 2005	Low	Low	High	Unknown	Unknown	Low
Kaufmann 2007	Low	Low	Unknown	Low	Unknown	Low
McDonald 1995	Unknown	Unknown	Unknown	Unknown	Unknown	Low
Colleoni 2011	Low	Unknown	Low	Low	Unknown	Low
Rutqvist 1993	Unknown	Unknown	Unknown	Low	Unknown	Low
van de Velde 2001	Low	Low	High	Unknown	Unknown	Low
Abo-Touk 2010	Low	Unknown	Unknown	Low	Unknown	Low
Goss 2005	Low	Unknown	Low	Low	Unknown	Low
Pagani 2014	Low	Unknown	High	Low	Unknown	Low

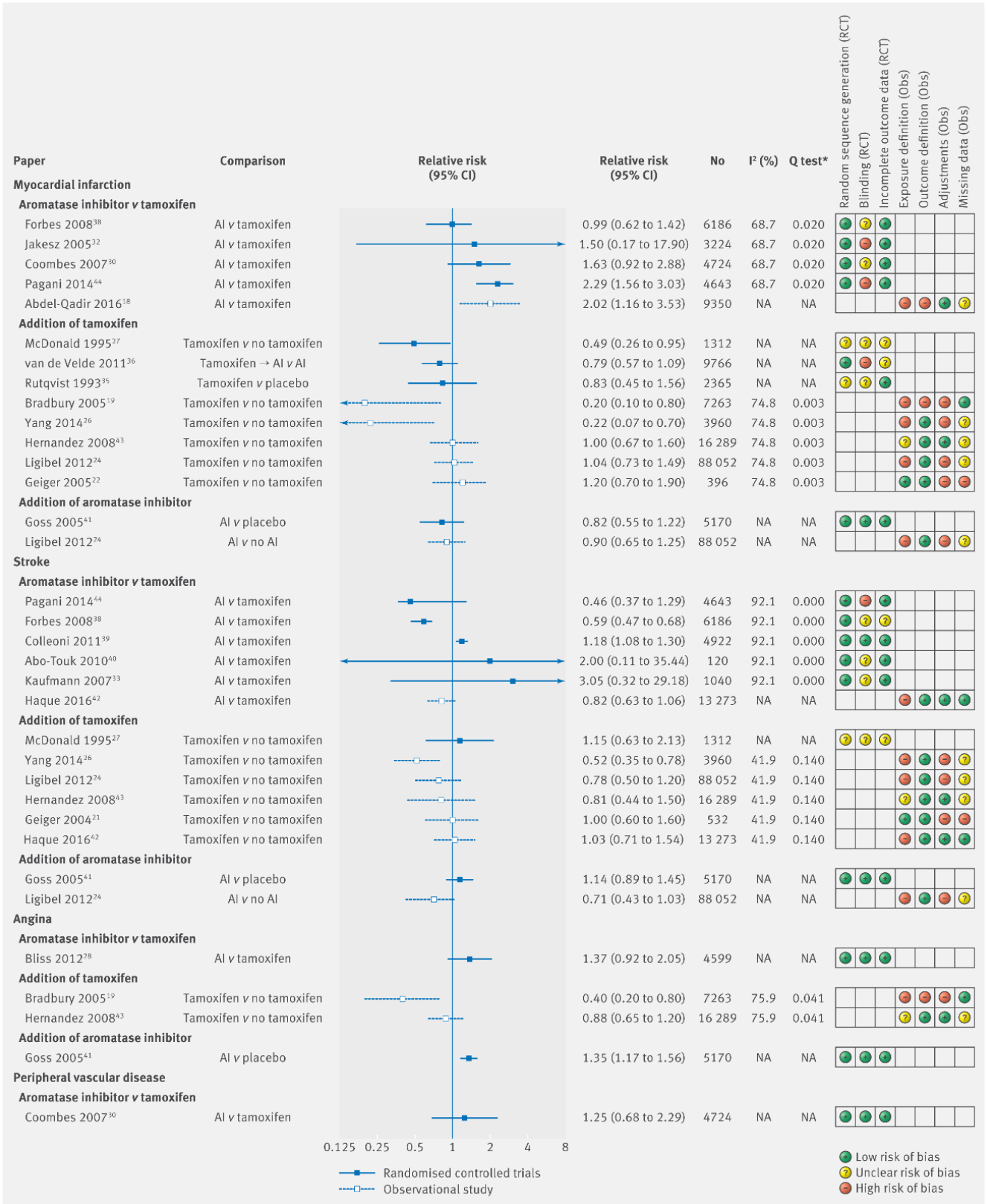


Fig 2 | Estimated relative risk (95% CI) for studies examining use of endocrine therapy and risk of specific vascular diseases, with corresponding I^2 tests, Q tests, and assessment of bias according to prespecified criteria. *P value. AI=aromatase inhibitor; NA=not applicable

doi: 10.1136/bmj.k3845 | *BMJ* 2018;363:k3845 | the bmj

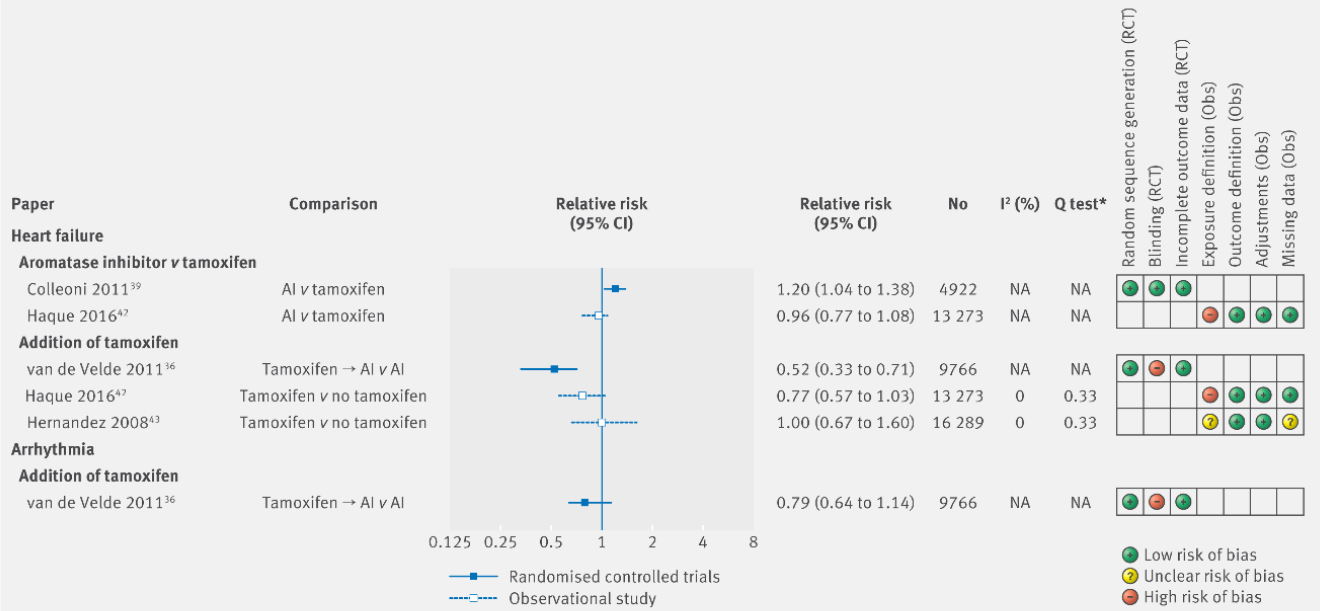


Fig 3 | Estimated relative risk (95% CI) for studies examining use of endocrine therapy and risk of specific myocardial diseases, with corresponding I^2 tests, Q tests and assessment of bias according to prespecified criteria

people with previous cardiovascular disease are likely to be underrepresented in randomised controlled trials, and are excluded from observational studies, the evidence on the association between endocrine therapies and risk of cardiovascular disease in this population remains limited. Furthermore, as the randomised controlled trials were mainly designed to assess disease-free survival, they were not always adequately powered to detect relative differences in the risk of clinical cardiovascular disease outcomes between treatment arms.

Definitions of cardiovascular disease outcomes were generally poorly recorded in the included studies. Between study variation could therefore exist in the measurement or coding of cardiovascular disease outcomes. Most oncology trials use the CTCAE criteria for adverse events, which have definitions that do not align with definitions in cardiology guidelines, although even the latter have variability. Furthermore, as observational studies rely on definitions of outcomes suggested by researchers and clinicians, differences in coding of outcomes could be a further source of heterogeneity in the observational studies. Without access to the outcome definitions and code lists used in these studies, fully understanding the extent to which the differences are problematic is challenging.

Explanation of key findings

A biological rationale exists for the use of aromatase inhibitors increasing the risk of cardiovascular disease outcomes, as they reduce oestrogen concentrations and therefore the oestrogen-mediated protective effects on cardiovascular disease, such as regulation of serum lipid metabolism, increasing vasodilation, and inhibition of the development of atherosclerosis.[45] Aromatase inhibitors could also increase the risk of hyperlipidaemia.[36] Evidence from randomised controlled trials suggests that tamoxifen has cardioprotective effects by decreasing lipid concentrations.[46 47] This systematic review postulates that some evidence exists for aromatase inhibitor users having an increased risk of myocardial infarction, relative to women treated with tamoxifen. However, whether this is driven by a decreased risk of myocardial infarction in tamoxifen users or an increased risk in aromatase inhibitor users is unclear, as results on the individual effects of tamoxifen and aromatase inhibitor are inconclusive.

We found evidence of heterogeneity between all four strata of observational studies exploring the same exposure and outcome (risk of myocardial infarction, stroke, angina, and venous thromboembolism with the addition of tamoxifen), which was potentially driven by the differences in study populations, statistical techniques used, and covariates adjusted for (appendix 5). For example, one cohort and one case-control study reported an increased risk of venous thromboembolism in tamoxifen users compared with non-tamoxifen users in European study populations (Denmark and the UK, respectively), whereas a study in Taiwan reported no evidence of a difference in risk. More broadly, as this systematic review attempted to cover a wide range of clinical cardiovascular disease outcomes, some included observational studies focused on one cardiovascular disease outcome, whereas others covered a broad range of cardiovascular disease outcomes. Different statistical techniques and covariate adjustments were therefore needed. The effect of this heterogeneity between studies was witnessed in the varying relative risks reported within these strata of

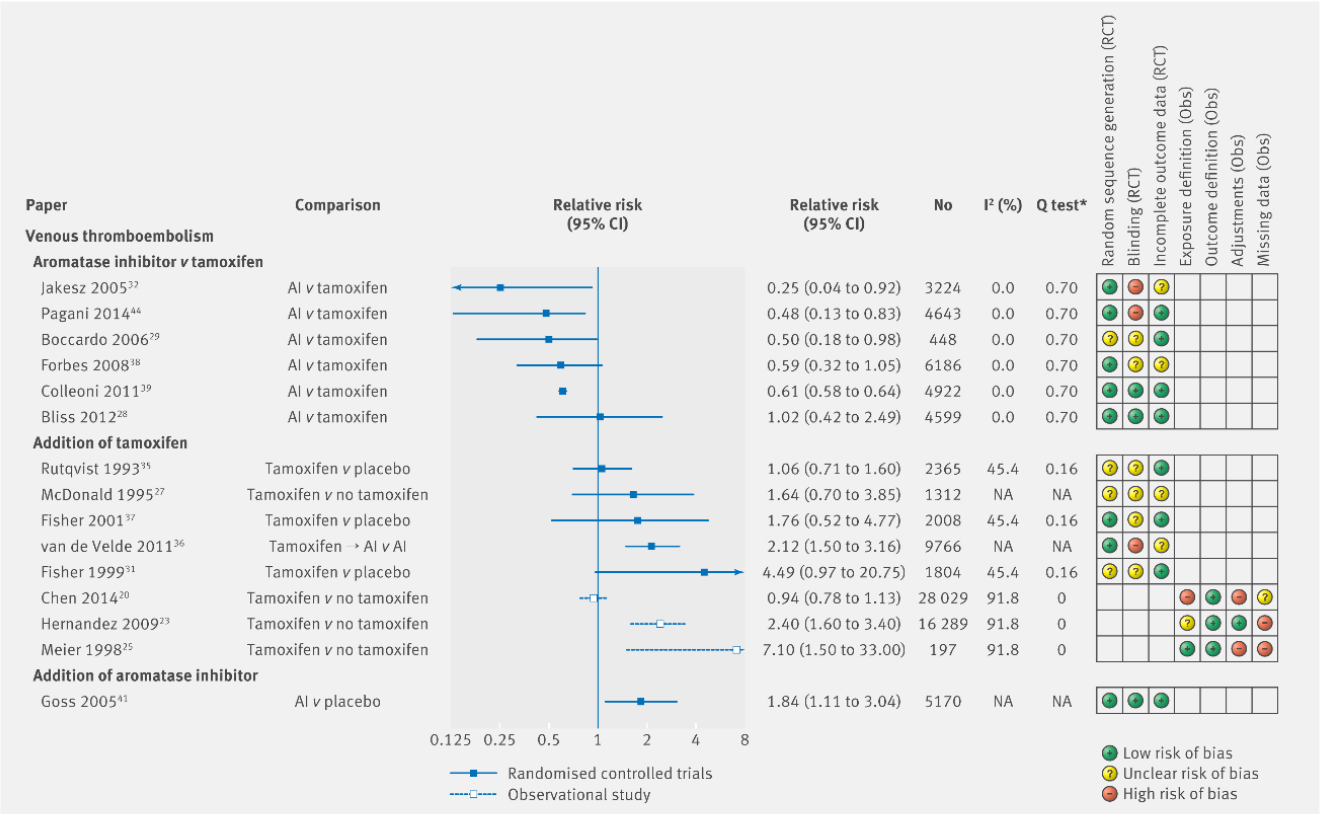


Fig 4 | Estimated relative risk (95% CI) for studies examining use of endocrine therapy and risk of venous thromboembolism, with corresponding I^2 tests, Q tests, and assessment of bias according to prespecified criteria

observational studies, which could be either a product of genuine discrepancies in risks between contrasting populations or the effect of residual confounding and different statistical techniques.

Comparison with other studies

The addition of observational studies in this review allowed comparison of results between real world populations and randomised controlled trials that generally use homogeneous study populations. Overall, we mostly found agreement in the direction of effect between randomised controlled trials and observational studies, but several observational studies reported more extreme effect estimates in comparison with randomised controlled trials (the risk of myocardial infarction with the addition of tamoxifen and the risk of venous thromboembolism with the addition of tamoxifen). However, these observational studies all had a high risk of bias in at least two assessment of bias categories. Including observational studies also enabled further evidence to be gathered where little or no evidence from randomised controlled trials existed. For example, we identified six observational analyses of vascular endpoints finding good or strong evidence for a higher risk for aromatase inhibitor compared with tamoxifen or a lower risk for tamoxifen compared with no tamoxifen. Most randomised controlled trial analyses were underpowered to detect differences in vascular endpoints, with only three finding similar clear evidence despite several others being suggestive of associations in the same direction.

We grouped comparisons on the basis of the drug women were given at the beginning of follow-up. For example, in several randomised controlled trials, women were given two to three years of tamoxifen before being randomised to either aromatase inhibitor or the continuation of tamoxifen for a further two to three years, with follow-up beginning at the point of randomisation. We classed these as a direct comparison of aromatase inhibitor versus tamoxifen, whereas previous reviews classed these comparisons as sequenced therapy versus tamoxifen alone. As all women had had the same treatment regimen before randomisation, classing these as aromatase inhibitor versus tamoxifen was a reasonable comparison to make.

The most recent meta-analysis by Khosrow concluded that randomised controlled trials directly comparing aromatase inhibitors with tamoxifen suggest that aromatase inhibitors are associated with an increased risk of cardiovascular events, but the cardioprotective effects of tamoxifen may account for this increased risk.[5] However, Khosrow et al used composite cardiovascular disease endpoints (excluding venous thromboembolism), which are defined slightly differently within each trial. We

doi: 10.1136/bmj.k3845 | *BMJ* 2018;363:k3845 | thebmj

stratified cardiovascular disease events into more specific outcomes and found a similar pattern for several vascular cardiovascular disease outcomes. The results for other cardiovascular disease outcomes including heart failure suggest a similar trend, but few studies have specifically explored these outcomes, so definite conclusions are unattainable. Like our study, that of Khosrow et al was inconclusive about the effects of endocrine therapy on cerebrovascular events. Another recent review by Rydén reported, with a high quality of evidence, that the risk of venous thromboembolism was higher in tamoxifen users than aromatase inhibitor users in randomised controlled trials.[2] Our study agrees with these results but also shows that this may be accounted for by the increased risk in tamoxifen users.

Strengths and limitations of this review

This study focused on individual clinical cardiovascular disease outcomes, excluding studies that reported composite outcomes. Understanding the effect of endocrine therapies on cardiovascular disease as a whole has several advantages, such as the potential to change the modifiable risk factors weight, smoking, statin use, and alcohol intake, which are present across all clinical cardiovascular diseases. However, understanding the effect of endocrine therapies on more specific clinical cardiovascular disease outcomes has the potential to enable clinicians to be targeted in their approach to preventing these outcomes in breast cancer survivors. The only composite outcome that we explored was venous thromboembolism, as some studies in this group included only deep vein thrombosis outcomes whereas others also included pulmonary embolism within a venous thromboembolism outcome. However, this grouping is relevant owing to the clinical similarities of deep vein thrombosis and pulmonary embolism.

Some relevant studies may have been missed, as searches of literature database take into account only indexed key terms or words used in the title and abstract. Studies in which the secondary outcome was a cardiovascular disease would therefore not have been identified in the original search. For example, several randomised controlled trials that focused on anticancer efficacy do not mention cardiovascular diseases in the title, abstract, or indexed keywords. However, we searched multiple large databases, manually searched the included studies' reference lists and relevant meta-analyses, and searched all endocrine therapy trials since the most recent meta-analysis, which was an indirect way of identifying the aforementioned randomised controlled trials and made this review as comprehensive as possible within the restricted framework imposed by the literature databases.

Implications of findings

This review establishes the need for clinical vigilance and possible preventive measures when prescribing endocrine therapies to women at risk of venous thromboembolism. Knowledge has also been progressed on the effects of endocrine therapies on the risk of vascular cardiovascular diseases, for which little evidence previously existed. However, we also showed that little or no evidence is available on the effect of endocrine therapies on several other specific cardiovascular disease outcomes, although substantial trial evidence outlines the effect on cardiovascular diseases generally. This is unlikely to be studied in future randomised controlled trials, so it is vital that large observational studies are carried out with details of baseline cardiovascular disease risk and drug treatment and clear definitions of cardiovascular disease events to fully understand the effects that endocrine therapies have on potentially fatal cardiovascular disease outcomes such as myocardial infarction, stroke, and heart failure.

Conclusions

Overall, the totality of the randomised controlled trial and observational evidence suggests a decreased risk of venous thromboembolism in aromatase inhibitor users compared with tamoxifen users, which is probably accounted for by an increased risk in tamoxifen users. The evidence also suggests that tamoxifen may have a protective association with vascular cardiovascular diseases, which may drive the higher risk of these outcomes in aromatase inhibitor users when directly compared with tamoxifen users. The results for some cardiovascular disease outcomes is still a mixed picture, many of the existing studies are susceptible to various sources of bias, and cardiovascular disease outcomes collected in oncology trials are generally limited. Nevertheless, the addition of observational studies alongside randomised controlled trials has substantially increased the amount of evidence supporting these conclusions. However, further high quality evidence is still needed for several cardiovascular disease outcomes. Although choice of aromatase inhibitor or tamoxifen will primarily be based on the effectiveness against recurrence of breast cancer, the individual patient's risk of venous or arterial vascular disease is an important secondary consideration, and the totality of evidence we present will thus help to inform prescribing.

Contributors: AM did the initial literature search, as well as the paper selection and data extraction. This process was replicated by REF and HS. AM wrote the first draft. All authors contributed to further drafts and approved the final manuscript. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. AM is the guarantor.

Funding: This study was funded through a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (grant No 107731/Z/15/Z) held by KB. The Wellcome Trust and the Royal Society had no role in the design, analysis, or writing up of this study.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organisation for the submitted work other than that described above; SS has received personal fees from Roche, Clinigen, Eli Lilly, and Novartis; AL has received personal fees from Servier, Novartis, Pfizer, Roche, Ferring Pharmaceuticals, Clinigen Group, Boehringer Ingelheim, Amgen, Eli Lily, and BMS; LS has received grants from Wellcome, MRC, NIHR, BHF, and Diabetes UK and grants and personal fees from GSK and is a trustee of the British Heart Foundation; no

other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Not needed.

Data sharing: All data are freely available within the appendices. No additional data available.

Transparency: The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

1 Davies C, Godwin J, Gray R, et al, Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *Lancet* 2011;378:771-84. doi:10.1016/S0140-6736(11)60993-8

2 Rydén L, Heibert Arnlind M, Vitols S, Höistad M, Ahlgren J. Aromatase inhibitors alone or sequentially combined with tamoxifen in postmenopausal early breast cancer compared with tamoxifen or placebo - Meta-analyses on efficacy and adverse events based on randomized clinical trials. *Breast* 2016;26:106-14. doi:10.1016/j.breast.2016.01.006

3 Mendelsohn ME. Protective effects of estrogen on the cardiovascular system. *Am J Cardiol* 2002;89(12A):12E-7E, discussion 17E-8E. doi:10.1016/S0002-9149(02)02405-0

4 Patnaik JL, Byers T, DiGuiseppi C, Dabelea D, Denberg TD. Cardiovascular disease competes with breast cancer as the leading cause of death for older females diagnosed with breast cancer: a retrospective cohort study. *Breast Cancer Res* 2011;13:R64. doi:10.1186/bcr2901

5 Khosrow-Khavar F, Filion KB, Al-Qurashi S, et al. Cardiotoxicity of aromatase inhibitors and tamoxifen in postmenopausal women with breast cancer: a systematic review and meta-analysis of randomized controlled trials. *Ann Oncol* 2017;28:487-96.

6 Aydiner A. Meta-analysis of breast cancer outcome and toxicity in adjuvant trials of aromatase inhibitors in postmenopausal women. *Breast* 2013;22:121-9. doi:10.1016/j.breast.2013.01.014

7 Amir E, Seruga B, Niraula S, Carlsson L, Ocaña A. Toxicity of adjuvant endocrine therapy in postmenopausal breast cancer patients: a systematic review and meta-analysis. *J Natl Cancer Inst* 2011;103:1299-309. doi:10.1093/jnci/djr242

8 Cuppone F, Bria E, Verma S, et al. Do adjuvant aromatase inhibitors increase the cardiovascular risk in postmenopausal women with early breast cancer? Meta-analysis of randomized trials. *Cancer* 2008;112:260-7. doi:10.1002/cncr.23171

9 Braithwaite RS, Chlebowski RT, Lau J, George S, Hess R, Col NF. Meta-analysis of vascular and neoplastic events associated with tamoxifen. *J Gen Intern Med* 2003;18:937-47. doi:10.1046/j.1525-1497.2003.20724.x

10 Foglietta J, Inno A, de Iuliis F, et al. Cardiotoxicity of Aromatase Inhibitors in Breast Cancer Patients. *Clin Breast Cancer* 2017;17:11-7. doi:10.1016/j.clbc.2016.07.003

11 Zagar TM, Cardinale DM, Marks LB. Breast cancer therapy-associated cardiovascular disease. *Nat Rev Clin Oncol* 2016;13:172-84. doi:10.1038/nrclinonc.2015.171

12 Giordano G, Spagnuolo A, Olivieri N, et al. Cancer drug related cardiotoxicity during breast cancer treatment. *Expert Opin Drug Saf* 2016;15:1063-74. doi:10.1080/14740338.2016.1182493

13 Bird BR, Swain SM. Cardiac toxicity in breast cancer survivors: review of potential cardiac problems. *Clin Cancer Res* 2008;14:14-24. doi:10.1158/1078-0432.CCR-07-1033

14 Higgins J, Green S. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0.* Cochrane Collaboration, 2011.

15 Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557-60. doi:10.1136/bmj.327.7414.557

16 Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011;342:d549. doi:10.1136/bmj.d549

17 Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011;343:d4002. doi:10.1136/bmj.d4002

18 Abdel-Qadir H, Amir E, Fischer HD, et al. The risk of myocardial infarction with aromatase inhibitors relative to tamoxifen in post-menopausal women with early stage breast cancer. *Eur J Cancer* 2016;68:11-21. doi:10.1016/j.ejca.2016.08.022

19 Bradbury BD, Lash TL, Kaye JA, Jick SS. Tamoxifen-treated breast carcinoma patients and the risk of acute myocardial infarction and newly-diagnosed angina. *Cancer* 2005;103:1114-21. doi:10.1002/cncr.20900

20 Chen TW, Chen HM, Lin CH, et al. No increased venous thromboembolism risk in Asian breast cancer patients receiving adjuvant tamoxifen. *Breast Cancer Res Treat* 2014;148:135-42. doi:10.1007/s10549-014-3140-2

21 Geiger AM, Fischberg GM, Chen W, Bernstein L. Stroke risk and tamoxifen therapy for breast cancer. *J Natl Cancer Inst* 2004;96:1528-36. doi:10.1093/jnci/djh285

22 Geiger AM, Chen W, Bernstein L. Myocardial infarction risk and tamoxifen therapy for breast cancer. *Br J Cancer* 2005;92:1614-20. doi:10.1038/sj.bjc.6602562

23 Hernandez RK, Sørensen HT, Pedersen L, Jacobsen J, Lash TL. Tamoxifen treatment and risk of deep venous thrombosis and pulmonary embolism: a Danish population-based cohort study. *Cancer* 2009;115:4442-9. doi:10.1002/cncr.24508

24 Ligibel JA, James O'Malley A, Fisher M, Daniel GW, Winer EP, Keating NL. Risk of myocardial infarction, stroke, and fracture in a cohort of community-based breast cancer patients. *Breast Cancer Res Treat* 2012;131:589-97. doi:10.1007/s10549-011-1754-1

25 Meier CR, Jick H. Tamoxifen and risk of idiopathic venous thromboembolism. *Br J Clin Pharmacol* 1998;45:608-12. doi:10.1046/j.1365-2125.1998.00733.x

26 Yang TL, Wu TC, Huang CC, et al. Association of tamoxifen use and reduced cardiovascular events among asian females with breast cancer. *Circ J* 2014;78:135-40. doi:10.1253/circj.CJ-13-0266

27 McDonald CC, Alexander FE, Whyte BW, Forrest AP, Stewart HJ, The Scottish Cancer Trials Breast Group. Cardiac and vascular morbidity in women receiving adjuvant tamoxifen for breast cancer in a randomised trial. *BMJ* 1995;311:977-80. doi:10.1136/bmj.311.7011.977

28 Bliss JM, Kilburn LS, Coleman RE, et al. Disease-related outcomes with long-term follow-up: an updated analysis of the intergroup exemestane study. *J Clin Oncol* 2012;30:709-17. doi:10.1200/JCO.2010.33.7899

29 Boccardo F, Rubagotti A, Guglielmini P, et al. Switching to anastrozole versus continued tamoxifen treatment of early breast cancer. Updated results of the Italian tamoxifen anastrozole (ITA) trial. *Ann Oncol* 2006;17(Suppl 7):vii10-4. doi:10.1093/annonc/mdl941

30 Coombes RC, Kilburn LS, Snowdon CF, et al, Intergroup Exemestane Study. Survival and safety of exemestane versus tamoxifen after 2-3 years' tamoxifen treatment (Intergroup Exemestane Study): a randomised controlled trial. *Lancet* 2007;369:559-70. doi:10.1016/S0140-6736(07)60200-1

31 Fisher B, Dignam J, Wolmark N, et al. Tamoxifen in treatment of intraductal breast cancer: National Surgical Adjuvant Breast and Bowel Project B-24 randomised controlled trial. *Lancet* 1999;353:1993-2000. doi:10.1016/S0140-6736(99)05036-9

32 Jakesz R, Jonat W, Gnant M, et al, ABCSG and the GABG. Switching of postmenopausal women with endocrine-responsive early breast cancer to anastrozole after 2 years' adjuvant tamoxifen: combined results of ABCSG trial 8 and ARNO 95 trial. *Lancet* 2005;366:455-62. doi:10.1016/S0140-6736(05)67059-6

33 Kaufmann M, Jonat W, Hilfrich J, et al. Improved overall survival in postmenopausal women with early breast cancer after anastrozole initiated after treatment with tamoxifen compared with continued tamoxifen: the ARNO 95 Study. *J Clin Oncol* 2007;25:2664-70. doi:10.1200/JCO.2006.08.8054

34 Mouridsen H, Giobbie-Hurder , Goldhirsch A, et al, ABIG 1-98 Collaborative Group. Letrozole therapy alone or in sequence with tamoxifen in women with breast cancer. *N Engl J Med* 2009;361:766-76. doi:10.1056/NEJMoa0810818

35 Rutqvist LE, Mattsson A, The Stockholm Breast Cancer Study Group. Cardiac and thromboembolic morbidity among postmenopausal women with early-stage breast cancer in a randomized trial of adjuvant tamoxifen. *J Natl Cancer Inst* 1993;85:1398-406. doi:10.1093/jnci/85.17.1398

36 van de Velde CJ, Rea D, Seynaeve C, et al. Adjuvant tamoxifen and exemestane in early breast cancer (TEAM): a randomised phase 3 trial. *Lancet* 2011;377:321-31. doi:10.1016/S0140-6736(10)62312-4

37 Fisher B, Anderson S, Tan-Chiu E, et al. Tamoxifen and chemotherapy for axillary node-negative, estrogen receptor-negative breast cancer: findings from National Surgical Adjuvant Breast and Bowel Project B-23. *J Clin Oncol* 2001;19:931-42. doi:10.1200/JCO.2001.19.4.931

38 Forbes JF, Cuzick J, Buzdar A, Howell A, Tobias JS, Baum M, Arimidex, Tamoxifen, Alone or in Combination (ATAC) Trialists' Group. Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 100-month analysis of the ATAC trial. *Lancet Oncol* 2008;9:45-53. doi:10.1016/S1470-2045(07)70385-6

39 Colleoni M, Giobbie-Hurder A, Regan MM, et al. Analyses adjusting for selective crossover show improved overall survival with adjuvant letrozole compared with tamoxifen in the BIG 1-98 study. *J Clin Oncol* 2011;29:1117-24. doi:10.1200/JCO.2010.31.6455

40 Abo-Touk NA, Sakr HA, Abd El-Lattef A. Switching to letrozole versus continued tamoxifen therapy in treatment of postmenopausal women with early breast cancer. *J Egypt Natl Canc Inst* 2010;22:79-85.

41 Goss PE, Ingle JN, Martino S, et al. Randomized trial of letrozole following tamoxifen as extended adjuvant therapy in receptor-positive breast cancer: updated findings from NCIC CTG MA.17. *J Natl Cancer Inst* 2005;97:1262-71. doi:10.1093/jnci/dji250

42 Haque R, Shi J, Schottinger JE, et al. Cardiovascular Disease After Aromatase Inhibitor Use. *JAMA Oncol* 2016;2:1590-7. doi:10.1001/jamaoncol.2016.0429

43 Hernandez RK, Sørensen HT, Jacobsen J, Pedersen L, Lash TL. Tamoxifen treatment in Danish breast cancer patients and 5-year risk of arterial atherosclerotic events: a null association. *Cancer Epidemiol Biomarkers Prev* 2008;17:2509-11. doi:10.1158/1055-9965.EPI-08-0570

44 Pagani O, Regan MM, Walley BA, et al, TEXT and SOFT Investigators, International Breast Cancer Study Group. Adjuvant exemestane with ovarian suppression in premenopausal breast cancer. *N Engl J Med* 2014;371:107-18. doi:10.1056/NEJMoa1404037

45 Mendelsohn ME, Karas RH. The protective effects of estrogen on the cardiovascular system. *N Engl J Med* 1999;340:1801-11. doi:10.1056/NEJM199906103402306

46 Dewar JA, Horobin JM, Preece PE, Tavendale R, Tunstall-Pedoe H, Wood RA. Long term effects of tamoxifen on blood lipid values in breast cancer. *BMJ* 1992;305:225-6. doi:10.1136/bmj.305.6847.225

47 Love RR, Wiebe DA, Feyzi JM, Newcomb PA, Chappell RJ. Effects of tamoxifen on cardiovascular risk factors in postmenopausal women after 5 years of treatment. *J Natl Cancer Inst* 1994;86:1534-9. doi:10.1093/jnci/86.20.1534

Supplementary appendices

2.2 MOTIVATION FOR THESIS

Given the results of the systematic review, it is clear that there are several gaps in knowledge for the association between endocrine therapies and the risk of CVD. There is a lot of evidence available for the effect of both endocrine therapies (both comparatively and individually) on the risk of some clinically specific CVD events (MI, stroke, and VTE), but there are still several outcomes for which there is little or no evidence (HF, arrhythmia, angina, revascularisation, SCA, cardiomyopathy, VHD). Given the widespread use of AIs in postmenopausal women diagnosed with ER+/PR+ breast cancer, it is important to fully classify such risks. The remainder of this thesis will therefore aim to assess the risk of a full range of clinically specific CVD outcomes associated with endocrine therapy use in postmenopausal women diagnosed with ER+/PR+ breast cancer.

3 DATA SOURCES

3.1 INTRODUCTION

This chapter summarises the data sources used in the UK and US studies, as well as an outline of my contribution to the data acquisition, cleaning, and curating.

3.2 UK STUDY

3.2.1 Data sources

Analyses carried out in UK data for this project utilised two EHR databases: the UK Clinical Practice Research Datalink (CPRD) and linked Hospital Episode Statistics (HES).

3.2.1.1 CPRD

The CPRD Gold (hereafter referred to as the CPRD) is a primary care database containing anonymised data from over 700 GP practices in the UK that use the Vision IT system, covering around 6.9% of the UK population.[98] Those included have been shown to be broadly representative of the UK population in terms of age, sex and ethnicity.[99] The CPRD includes patient-level information from GP surgeries such as consultations, clinical diagnoses, prescriptions in primary care, tests, immunisations, and referral to hospitals. There is also information on lifestyle measures that are collected by GPs such as smoking, alcohol use, and BMI. In addition to these routinely collected data, there may be records of information from some secondary sources that have been passed on to GPs, such as major diagnoses made in hospitals. All diagnoses in the CPRD are recorded using the hierarchical Read code classification system. Read codes have a corresponding description, allowing users of these data to identify diagnoses and their related codes. However, GPs can also input patient information in the form of free text and scans of hospital letters without coding, which could lead to missing information as these data are not available for research. Prescription records are created automatically when a GP issues a prescription, so capture of primary care prescriptions should be almost 100%. Prescription drugs are recorded with a unique CPRD generated product code and description of the drug for each type of prescription. Once data are recorded by GP practices, they undergo a series of validity checks, both at the patient and practice level. Patient level acceptability is based on their registration status, recording of events in the patient record, and valid age and gender. Each practice is also given an 'up to standard date',

which indicates the date from which an individual practice includes reliable data.[98] In depth documentation provided by CPRD on the definition of patient and practice level acceptability is available in Appendix 3.1. CPRD data have been shown to be highly valid in regards to most clinical diagnoses. A systematic review of validation studies reported that the median proportion of neoplasm and circulatory system cases in CPRD confirmed by another source was 95% and 85% respectively.[100] Specifically breast cancer diagnoses are highly valid, with 96% of cases in CPRD also being recorded in either the UK national cancer registry or another data source.[101] Lifestyle measures such as smoking and BMI have also been shown to be highly valid in the CPRD,[102, 103] but can be prone to substantial missingness.

Data Structure

Each patient in the CPRD has a unique patient identifier (patient ID), which can be used to link the files in which all data are stored to build a chronological set of records for each patient. The patient ID is present in all files that provide patient level data. An overview of all data files is included in Table 3.1.

Table 3.1: Description of CPRD data files

File	Description
Patient	Patient level demographic data: year of birth, gender, date of death, date of transfer out of the practice.
Clinical	Patient level clinical diagnoses: date of event and diagnostic code.
Therapy	Patient level information on prescriptions: date of prescription, related drug identifier code, strength, and formulation.
Test	Patient level information on tests: date of test, test type, and result.
Referral	Patient level referrals: date of referral, diagnosis give, method of referral, referral specialty.
Additional	Further information relating to clinical events. This includes information that cannot be stored within the Clinical file such as blood pressure readings, number of cigarettes smokes, BMI, level of alcohol drinking etc.
Consultation	Patient level data on consultations within the GP surgery: date of consultation, type of consultation, and duration of consultation.
Practice	Practice level information on the GP surgery: geographical region, up to standard date, last data collection date .

3.2.1.2 Linked HES

HES is an administrative secondary care database that contains information on admissions from all National Health Service (NHS) hospitals in England. A subset of English GP practices that contribute data to the CPRD take part in a linkage scheme where their patients' records are linked to HES data (75% of practices in England, which equates to 56% of practices in the whole of the UK).[104] These data are linked through deterministic matching using a combination of the patients' NHS number, gender, date of birth, and postcode. The linkage is undertaken by NHS Digital, acting as a trusted third party, on behalf of CPRD. No personal identifiers are held by CPRD. Information on inpatient and day care admissions, outpatient appointments, and accident and emergencies, with comprehensive information on diagnoses, are available in the linkage scheme. However, no information is currently available on prescriptions given to patients in hospital. Once diagnoses are input into hospital systems by clinicians, clinical coders convert these unstructured clinical notes into International Classification of Diseases 10th revision (ICD-10) codes. Procedures are coded using the Office of Population Censuses and Surveys Classification of Interventions and Procedures version 4 (OPCS-4) coding system.

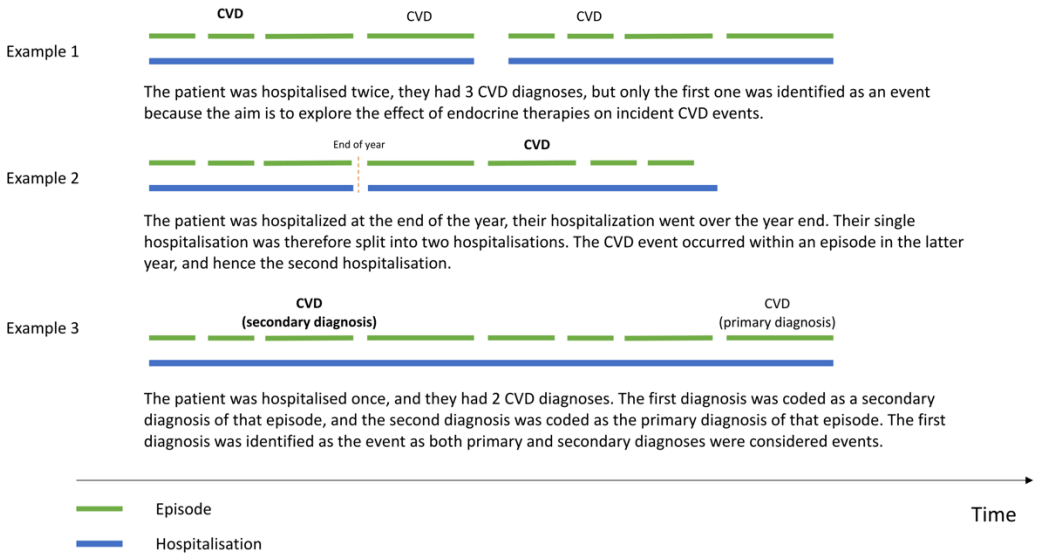
Data structure

As part of the linkage process outlined above, each patient's HES data are tied to the appropriate CPRD identifier (patient ID). The data are arranged into files relating to hospitalisations, episodes, and files for events that are linked to specific episodes. Hospitalisations refer to the total period of inpatient hospital stay from admission to discharge. When a hospitalisation spans the end of the HES year, it is artificially modelled as two hospitalisations, from admission to end of HES year (in the first year's HES data) and from start of the HES year to final discharge (in the second HES year). An episode is a time-period within a hospitalisation, which corresponds to the period where the patient is in the continuous care of one consultant using the beds of one health care provider. This is not always the same as a single stay in hospital, because a patient may be transferred from one consultant to another during their stay. In these cases, there will be two or more episode records for the hospitalisation. Consultant episodes will also terminate when a patient is transferred between health care provider organisations, even though their inpatient care may be continuous. Each patient may have one or more HES hospitalisations. Each hospitalisation can consist of one or more episodes. For each episode, up to 20 diagnoses and 24 procedures may be recorded. Additionally, each episode can have up to nine periods of augmented care. Each episode has one primary diagnosis, and can have several secondary diagnoses, with the primary diagnosis relating to the main

reason for that episode occurring.[105]

In this thesis, a CVD diagnosis in HES was classed as an event if it was within any episode of a hospitalisation. The diagnosis could also be either the primary or secondary diagnosis of that episode (see Chapter 4). Furthermore, the incident event of interest was captured, but no further events. However, the aim of the thesis is to explore the effect of endocrine therapies on incident CVD events, so further events would not matter in the context of the research question. Figure 3.1 gives three examples of how CVD events were identified in HES.

Figure 3.1: Examples of how CVD events were identified in HES



3.2.2 Other potential data sources

There are several other data sources within the UK that could have been used to answer the research question of the thesis. It would be possible to use RCT data, and there have been results from RCT studies reported for the effect of endocrine therapies on several CVD outcomes, as outlined in the systematic review. However, none of these RCTs reported the information on the full range of CVD outcomes, and they were mainly set up to assess the efficacy of tamoxifen or AIs, so are prone to several sources of bias when attempting to explore the CVD side effects of treatment. It would not be possible to set up an RCT to answer the research question due to ethical considerations as AIs have already shown to be more efficacious in comparison with tamoxifen in post-menopausal women diagnosed with ER+/PR+ breast cancer.

Although I have used the CPRD and HES data, which are routinely collected health records, it would also have been possible to use The Health Improvement Network database (THIN).[106] This database is comparable to the CPRD, but the process of linking data to HES is easier for CPRD as it is done in house at CPRD. A third-party company needs to link THIN to HES, which can be expensive and time consuming. LSHTM also has a license with CPRD, but not with THIN, which means CPRD data were cheaper and more easily accessible.

There are currently no appropriate registries in the UK (other than the UK cancer registry, which is discussed later) that would be suitable to identify the source population and answer this research question. However, if there was such a registry it would likely include information on a wider selection of potential confounders that are not available in the CPRD, such as breast cancer stage, grade, and other treatments that were administered. It also should include 100% of all breast cancers diagnosed in the UK.

It would be possible to answer the research question by using data from a prospective cohort study. Although there are several prospective cohort studies currently taking place in the general area of breast cancer research (such as the Generations Study,[107] which examines breast cancer risk and recruited both women with and without a breast cancer diagnosis), none are suitable for answering this specific research question. A prospective study set up to answer this question would require follow up of women diagnosed with breast cancer, with information on their treatment and subsequent CVD outcomes. It is likely that study populations for any prospective study set up to answer this question would be too small to have adequate statistical power to detect an effect of endocrine therapies on each CVD of interest.

There are no suitable administrative databases in the UK that will give sufficient information to answer this research question. Such databases could be from healthcare insurance providers that are mainly used for administrative rather than care purposes. However, the size of these databases would be small given the proportionally small number of people covered by private healthcare insurers in the UK, and the cost of accessing these data for research purposes are likely expensive.

3.2.3 Contribution

I led all data acquisition, cleaning, and curating throughout the duration of the UK study. This included conceptualising and writing a protocol for the study to submit to the Independent Scientific

Advisory Committee board. Approval is required from this board prior to access to anonymised patient level CPRD data for research purposes. LSHTM holds a CPRD license, so once I gained protocol approval, I was able to extract the defined dataset from the files held at LSHTM and begin the cleaning and curating process. I was solely responsible for this process, which included defining and creating the study population, exposure, outcome, and covariates within the data. This was also the stage at which several of the study decisions were taken. Although I initiated all decisions, I also had conversations with both of my supervisors to get their opinions on several of the decisions made. This process took around 8 months to a year from the beginning of the PhD (2016) before I was able to start the main analyses.

3.3 US STUDY

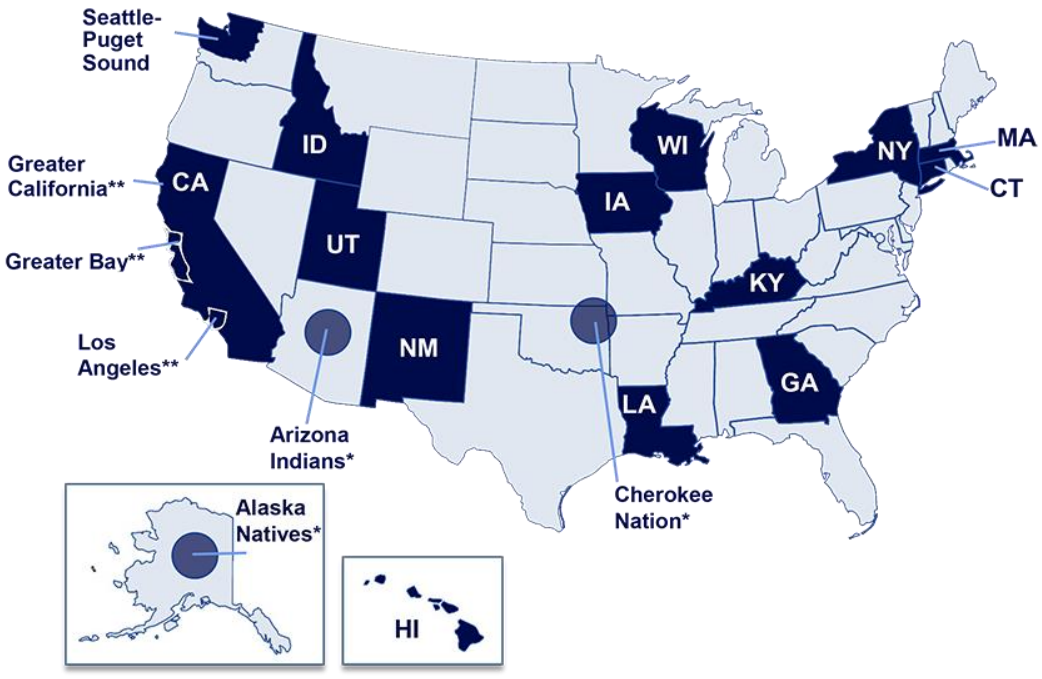
3.3.1 Data sources

Analyses carried out in the US utilised the linkage of two large population-based databases: the Surveillance, Epidemiology and End Results Program (SEER) and Medicare parts A, B, and D.

3.3.1.1 SEER

The SEER program of the US National Cancer Institute is a surveillance system that began collecting information in 1973 on demographics, clinical and tumour characteristics, initial surgical and

Figure 3.2: SEER reporting states



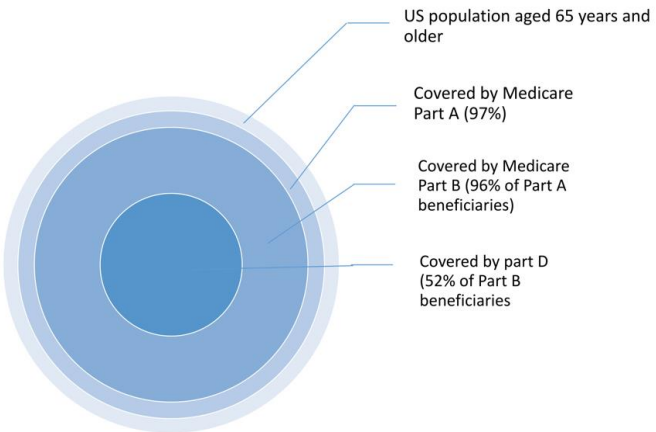
*Subcontract under New Mexico
**Three regions represent the state of California: Greater Bay, Los Angeles, and Greater California

radiation treatment, vital status, and cause of death for all individuals who are diagnosed with cancer and reside within one of the 12 SEER states (Figure 3.1). Geographic areas were selected for inclusion in the SEER Program based on their ability to operate and maintain a high quality population-based cancer reporting system and for their epidemiologically significant population subgroups. The program covers approximately 34.6% of the US population. Although it is only collected from 12 states and is not a nationwide sample, SEER is representative of the wider US population in terms of poverty and education, but includes more urban areas and a greater proportion of foreign-born residents.[108, 109] SEER is considered the gold standard for data quality among cancer registries in the US, with a 95% complete case identification and a 95% annual rate of follow-up to determine survival.[110]

3.3.1.2 Medicare parts A, B, and D

The Medicare systems provides governmental funded health insurance for approximately 97% of US citizens aged 65 years and over.[111] Medicare health insurance is also available to individuals under 65 years if they have been diagnosed with end-stage renal disease or medical disability. As of 2016, 57 million people were covered by Medicare, of whom 48 million were age 65 years or over (84%). Everyone covered by Medicare is entitled to Part A coverage, which includes hospital inpatient care. Around 96% of people then pay for Part B coverage, which covers physician and outpatient services.[112] Claims data from Medicare Parts A and B are available from 1998 onwards. 52% of Medicare beneficiaries also have Part D enrolment, which began in 2007 and offers outpatient coverage for medications. There is an extra cost for Medicare parts B and D, meaning this may only cover a more affluent population. Figure 3.3 gives an outline of how Medicare Parts A, B, and D are

Figure 3.3: Medicare coverage overview



related. Medicare data are highly accurate due to being claims based and the centres for Medicare services regularly evaluating accuracy of billing information by taking random samples of medical records, and reviewing them against the claim for accuracy. Hospitals are subject to penalties if they have up-coded claims.[113] Several CVD outcomes from Medicare have been externally validated: heart failure claims were validated against medical discharge records from the Atherosclerosis Risk in Communities Study with high levels of agreement (Kappa coefficient (κ) 0.92);[114] and myocardial infarction and coronary revascularisation were validated against data from the Women's Health Initiative with equally high agreement (κ=0.74 and κ=0.91, respectively).[115]

3.3.1.3 Linked SEER-Medicare

The SEER-Medicare database is a linkage of the two above databases using a deterministic algorithm based on name, social security number, sex, and date of birth, and includes cancer cases until 2013 and Medicare claims until 2014. This time restriction could mean that any results are not generalisable to the post-menopausal women diagnosed with ER+/PR+ breast cancer today because there are likely less women being prescribed tamoxifen due to the changing use of endocrine therapies.

3.3.2 Contribution

This study was a collaboration with colleagues from the Department of Epidemiology at the University of North Carolina (UNC), Chapel Hill in the US. I led on the study design and protocol development, supported by advisors at both institutions. The protocol was submitted to and approved by the SEER advisory board prior to any work began. Sharon Peacock Hinton, an employee at UNC, then produced the specified dataset from the raw data, with guidance from myself. In depth information on the data structure is therefore not provided. As this was a collaboration with UNC, variables were defined using their well-established definitions, hence there was less scope for validation of variables (i.e. as in chapter 4 for CVD outcomes in the UK study). Once I received the cleaned dataset, I was then responsible for identifying the correct exposures and exposure categories. The project required me to spend three months in at UNC in Chapel Hill, North Carolina, where I was a fully integrated member of the Department of Epidemiology. The collaboration with colleagues at UNC began at the beginning of 2017. Protocol and study development took place throughout 2017, then I visited North Carolina from March-June 2018 to carry out all data analyses.

Prior to forming the collaboration with UNC, I had originally planned to collaborate with the

Department of Clinical Epidemiology at Aarhus University in Denmark. The project was very similar to the one that I eventually carried out in the US. I, again, wrote a protocol for the project which was approved in Denmark. However, following this approval there was a change of the laws in Denmark, which made it illegal for non-Danish researchers to access prescription data. This meant it would not have been possible for me to carry out any analyses. The project was therefore put on hold, and we have applied for funding to hire a Danish researcher to carry out the analyses.

3.4 SUMMARY

- This thesis utilises linked data from the CPRD and HES for the UK study.
- 75% of GP practices that contribute data to the CPRD in England have linked HES data, which equates to 56% of practices in the whole of the UK.
- CPRD data have been shown to be highly valid in regards to most clinical diagnoses.
- Linked SEER-Medicare claims data were used for the US study.
- SEER-Medicare include all cancer cases from 1973 until 2013, and Medicare claims from 1998 until 2014.

4 CAPTURE OF CVD OUTCOMES IN THE UK STUDY

4.1 INTRODUCTION

A range of clinical CVDs make up the main outcomes to be explored within this thesis. This chapter contains a series of exploratory analyses using both CPRD and HES with the aim of developing CVD outcome definitions in UK data, assessing the validity of UK primary and secondary care data to detect CVD outcomes, and deciding if the main analyses in the UK should use either the CPRD data alone or the CPRD and HES linked data. This will be achieved through the following objectives.

4.2 OBJECTIVES

1. Describe the process of creating CVD outcome code lists in CPRD.
2. Identify and describe definite and possible CVD events in the CPRD data.
3. Examine the concordance of events between CPRD and HES.
4. Identify events recorded in HES alone, and explore CPRD data around this time to see if any diagnoses in primary care were originally missed.
5. Assess the validity of CVD events in HES.
6. Calculate the statistical power of analyses using either CPRD data alone, or linked CPRD and HES data.

4.3 METHODS

4.3.1 Study population

The following analyses use the study population definition that is fully outlined in the UK study in Chapter 5. For objective 2, all eligible patients in the CPRD were included (until 2017). The study population was then restricted for objectives 3-5 to only include those eligible for the CPRD and HES linkage scheme, and for follow up time up until the latest CPRD and HES linkage date (2016).

4.3.2 CVD outcomes

A list of clinical CVD outcomes was created to explore in all future analyses (both UK and US). These outcomes were discussed with a cardiologist and GP and included both composite and specific outcomes. The final agreed outcomes were as follows:

- Coronary artery disease
    - Angina
    - Myocardial infarction
    - Revascularisation procedures
    - Sudden cardiac arrest
- Peripheral vascular disease
- Stroke
- Arrhythmia
- Heart failure (consisting of heart failure and cardiomyopathy)
- Pericarditis
- Valvular heart disease
- Venous thromboembolism
    - Deep vein thrombosis
    - Pulmonary embolism

Although it would have been possible to include hypertension as an outcome, the study team decided that hypertension contributes to CVD, rather than being a clear clinical CVD outcome itself. If endocrine therapies altered blood pressure, it is possible that hypertension could partly mediate any effect of endocrine therapies on CVD.

4.3.3 Identification of codes used to define CVD outcomes

4.3.3.1 CPRD

CPRD uses Read codes to identify clinical events that have been diagnosed in primary care. Read codes are a coded thesaurus of clinical terms that have been used in the National Health Service (NHS) since 1985, which provide a standard vocabulary by which clinicians can record patient findings and procedures in health and social care IT systems across primary and secondary care. Creation of Read code lists representing a certain disease allows identification of patients in CPRD with a diagnosis of that disease by merging the code lists with the patients' raw data files.

A systematic approach was used to define code lists in order to identify clinical diagnoses of CVD outcomes. The dictionaries of codes were searched using STATA do files,[116] so that all decisions on inclusion and exclusion criteria were recorded and easily replicated.

A code list for each clinical diagnosis was created using the following algorithm:

- A list of inclusion search terms, which were synonyms of the medical event, was agreed through discussion with the clinical collaborators involved in the study.
- The CPRD Read code dictionary was then searched to identify any codes with one of the search terms in the read term field, which is used to describe the Read code.
- Using a list of pre-specified terms and scanning the codes initially identified, a list of exclusion terms was created and applied.
- Codes and their descriptors were manually reviewed to decide if they were appropriate for the final code list .

Initial inclusion terms to create CVD code lists utilised the search terms in the CALIBER study, where available.[117] These search terms were reviewed and expanded if necessary.

4.3.3.2 HES

HES uses ICD-10 codes to identify clinical events diagnosed in secondary care. ICD-10 codes are a comprehensive classification of causes of morbidity and mortality that is published by the World Health Organisation (WHO). The 10th revision of ICD codes was published in April 1995, and followed the 9th revision (ICD-9) that was published in 1975 and came into use in hospital health systems in 1979. Furthermore, OPCS-4 codes are used to classify interventions and procedures, and were originally published in 1987 by the Office of Population Censuses and Surveys, and was followed by a 4th revision in 1992**.**

CVD outcome events were identified in HES. Clinicians guided the creation of ICD-10 code lists relating to all CVD outcomes. OPCS-4 codes lists were also created to identify revascularisation procedures carried out in secondary care. These codes lists were then used to search for relevant diagnoses in the patients' raw HES data.

4.3.4 Definite, possible, and history of events

For all CVD events in CPRD, the final Read code list was reviewed and all codes were categorised as either a definite event, a possible event, or history of an event. Codes indicating a definite event were those showing a clear clinical diagnosis. Possible event codes were those codes in which a diagnosis was not clear, but there was an indication that the patient suffered from the related CVD. For example, a possible stroke Read code had the description "stroke self-management plan agreed", which is a process of care code and shows that the patient has most probably suffered from a stroke, and is now been given guidance on how to deal with the consequences. This could also come in the form of a code indicating the patient had been referred to a clinic relating to the diagnosis. For some CVD outcomes, there were also other closely related diagnoses that were classed as possible codes. For example, a pulmonary oedema diagnosis usually occurs when the diseased or overworked left ventricle cannot pump out enough of the blood it receives from the lungs, which is essentially congestive HF. So, Read codes relating to pulmonary oedema were considered as possible HF codes. All Read codes that indicated a "suspected" event were also classed as possible codes. Finally, a code was classified as a history of event code if the Read code indicated an event that had happened in the past, or it showed monitoring or review of a past diagnosis.

4.3.5 Analysis

4.3.5.1 Objective 1 - Describe the process of creating CVD outcome code lists in CPRD

For each CVD outcome, the number of codes identified at each stage of the CPRD code list creation algorithm (as above) was outlined. This included this number of codes identified: after inclusion terms had been applied; after exclusion terms had been applied; and after the resulting code list had been manually reviewed. A full list of inclusion and exclusion terms for each CVD outcome are available in Appendix 4.1, and the final Read code lists created are available in appendices 4.2-4.13.

4.3.5.2 Objective 2 - Identify and describe definite and possible CVD events in the CPRD data

After excluding those with a past event (a diagnosis with any code in the final code list), incident CVD events in CPRD during follow up were identified from patients' Clinical and Referral files using the Read code lists created in objective 1 (excluding diagnoses recorded with "history of" codes). All CVD outcomes were assessed separately. The number of events classed as definite and possible events was calculated. If an incident event was classed as possible, the patients' files were further explored to identify how many of these possible events were followed by a definite event within the following

year. The top five Read codes and descriptions that identified incident events for each CVD outcome were then presented.

4.3.5.3 Objective 3 - Concordance of CVD outcomes in linked CPRD and HES data

The study population was restricted to those in the linkage scheme, and women were excluded if they had been diagnosed with a CVD event prior to the index date in either CPRD or HES. Incident CVD events during follow up were identified from an appropriate Read code in patients' Clinical and Referral files (again excluding diagnoses recorded with "history of"' codes), and from patients' HES diagnosis files using the ICD-10/OPCS-4 codes in appendix 4.14. The number of events during follow up were initially identified using CPRD and HES records separately, and then when using both CPRD and HES records together. This allowed the overlap between the two databases to be calculated. The median time between events that occurred in both CPRD and HES (in either direction) was also calculated.

4.3.5.4 Objective 4 - Exploration of primary care codes recorded around the time of HES only events

To identify Read codes missed in the original code lists, all diagnoses occurring in CPRD within a six-month period around an event (three months before and after) that only occurred in HES were identified for each outcome. All events and related Read codes in this period were then accumulated and manually scanned to find any Read codes that were missing from the original Read code list. For outcomes in which new Read codes were found, above analyses examining the concordance between CPRD and HES were repeated. Furthermore, for the four CVD outcomes in which the smallest proportion of total events could be identified in CPRD alone (angina, pericarditis, SCA, and VHD), five women were picked at random, and their individual CPRD diagnosis, prescription, and test records were explored in detail for a year around the time of the HES event. This helped to understand if there was any other way to identify a record in primary care records that was otherwise only recorded in a woman's HES records.

4.3.5.5 Objective 5 - Validity of HES events

For each episode in HES, there can up to 20 diagnoses, with one primary diagnosis and several secondary diagnoses. When originally identifying incident CVD events in HES, all recorded diagnoses were used, regardless of if the event was the primary or secondary diagnosis of the episode. To understand how many events were the main reason for the time in hospital, and to help decide if

only primary events should be used, the proportion of events recorded as the primary diagnosis of that episode were calculated. The ICD-10 letter of the primary diagnoses for that episode was tabulated for all events for which the diagnosis was not the primary diagnosis (i.e. events that indicated a secondary diagnosis). This process was replicated for all outcomes other than revascularisation, as this is classed as a procedure not a diagnosis.

4.3.5.6 Objective 6 - Calculate the statistical power of analyses using either CPRD data alone or linked CPRD and HES data

Detectable hazard ratios for survival analysis were calculated based on the sample size, event numbers, and standard deviation of endocrine therapy exposure (calculated from those exposed to tamoxifen or AIs in the study population) for analyses using CPRD data alone to define the study population and capture CVD events, and analyses that defined the study population in the linked CPRD and HES data, but captured events that were in either CPRD or HES. All detectable hazard ratios were calculated with a statistical power of 80% and type 1 error rate of 5%. The detectable hazard ratio calculates the minimum possible hazard ratio that is detectable based on the sample size and number of events and gives a good indication of size of effects that could be found with the available data. Given there are two potential data sources, comparing the minimum detectable hazard ratio between the two sources enables one to see the source in which smaller effect sizes can be detected.

4.4 FINDINGS

4.4.1 Objective 1 - Describe the process of creating CVD outcome Read code lists in CPRD

After initial inclusion and exclusion criteria were applied, and the codes had been manually scanned to remove any that were inappropriate, the final number of Read codes ranged from 128 in the revascularisation outcome to 20 in the PE outcome (Table 4.1). All Read code created are available in appendices 4.2-4.13**.**

Table 4.1: Number of codes within CVD code lists after inclusion, exclusion, and manual searching

Outcome	After inclusion	After exclusion	Number manually excluded	Final
Angina	71	43	2	41
MI	217	82	9	73
Revascularisation	142	128	0	128
SCA	62	24	6	18
PVD	1874	87	0	87
Stroke	250	250	126	124
Arrhythmia	85	72	1	71
HF	186	150	17	133
Pericarditis	38	37	0	37
VHD	53	48	0	48
DVT	72	55	12	43
PE	22	21	1	20

4.4.2 Objective 2 - Identify and describe definite and possible CVD events in the CPRD data

Once women with prior CVD events in CPRD were excluded, the final number of women included in the study population ranged from 15,262 in the arrhythmia analysis to 16,568 in the pericarditis analysis. Table 4.2 outlines the total number of incident events, the proportion of those that were definite, and the proportion of possible events followed by a definite event within a year. For the majority of outcomes, all incident events were definite, with stroke reporting the lowest number of incident events that were definite (73%). However, 15% of the possible stroke events were followed by a definite event within a year. Furthermore, over 59% of all incident events were captured by 5 Read codes for all outcomes (Table 4.3), with over 90% of all events being captured by 5 Read codes for 6/12 of the outcomes.

Table 4.2: Events in CPRD between 2002 and 2017

Outcome	**Women in study population**	**Total number of incident events in follow up (% of total in study pop)**	**Definite incident events (% of all events)**	**Possible events that were confirmed by a definite event within a year (% of all possible events)*†**
Angina	15655	97 (0.62)	97 (100)	N/A
MI	16206	162 (1.00)	162 (100)	N/A
Revascularisation	16484	36 (0.22)	36 (100)	N/A
SCA	16564	18 (0.11)	18 (100)	N/A
PVD	16260	123 (0.76)	119 (97)	0 (0.00)
Stroke	15934	365 (2.29)	267 (73)	15 (15.31)
Arrhythmia	15262	635 (4.16)	635 (100)	N/A
HF	15997	318 (1.99)	277 (87)	7 (17.07)
Pericarditis	16568	6 (0.04)	6 (100)	N/A
VHD	16413	113 (0.69)	113 (100)	N/A
DVT	16115	284 (1.76)	232 (82)	6 (11.54)
Pulmonary Embolism	16324	189 (1.16)	188 (99)	0 (0.00)

* N/A if all incident events were definite events

† 0 if there are some incidents events are possible events, but none of them have a definite event within a year

Table 4.3: Top five Read codes used to identify incident CVD events between 2002 and 2017

Outcome	Read code	Description	Definite /possible events	Number of incident events	Proportion of all incident events	Cumulative proportion of all incident events
Angina	G33..00	angina pectoris	Definite	75	77.3	77.3
	G311100	unstable angina	Definite	7	7.2	84.5
	G311.13	unstable angina	Definite	5	5.2	89.7
	G332.00	coronary artery spasm	Definite	2	2.1	91.8
	G311200	angina at rest	Definite	2	2.1	93.8
MI	G307100	acute non-st segment elevation myocardial infarction	Definite	67	41.4	41.4
	G30..00	acute myocardial infarction	Definite	42	25.9	67.3
	G30..15	mi - acute myocardial infarction	Definite	31	19.1	86.4
	G30X000	acute st segment elevation myocardial infarction	Definite	12	7.4	93.8
	G30z.00	acute myocardial infarction nos	Definite	5	3.1	96.9
Revascularisation	7911300	replacement of aortic valve nec	Definite	13	21.7	21.7
	7911.12	replacement of aortic valve	Definite	10	16.7	38.3
	792..11	coronary artery bypass graft operations	Definite	10	16.7	55
	792E.00	percutaneous coronary intervention	Definite	8	13.3	68.3
	7911.00	plastic repair of aortic valve	Definite	6	10	78.3
SCA	G575.00	cardiac arrest	Definite	11	61.1	61.1
	7937500	implantation of internal cardiac defibrillator	Definite	4	22.2	83.3
	8532.11	cardiopulmonary resuscitation	Definite	1	5.6	88.9
	2241.00	o/e - collapse -cardiac arrest	Definite	1	5.6	94.4
	G575.12	asystole	Definite	1	5.6	100
PVD	G73z000	intermittent claudication	Definite	34	27.6	27.6
	G73z.00	peripheral vascular disease nos	Definite	13	10.6	38.2
	G73..00	other peripheral vascular disease	Definite	12	9.8	48
	G73zz00	peripheral vascular disease nos	Definite	7	5.7	53.7
	C107.00	diabetes mellitus with peripheral circulatory disorder	Definite	7	5.7	59.3
Stroke	G66..00	stroke and cerebrovascular accident unspecified	Definite	73	20	20
	G66..11	cva unspecified	Definite	63	17.3	37.3
	8HBJ.00	stroke / transient ischaemic attack referral	Possible	62	17	54.2
	8HTQ.00	referral to stroke clinic	Possible	32	8.8	63
	G64z.00	cerebral infarction nos	Definite	25	6.8	69.9
Arrhythmia	181..00	palpitations	Definite	386	42.5	42.5
	G573000	atrial fibrillation	Definite	236	26	68.5
	G573.00	atrial fibrillation and flutter	Definite	86	9.5	78
	G573200	paroxysmal atrial fibrillation	Definite	27	3	80.9
	R050.00	[d]tachycardia, unspecified	Definite	24	2.6	83.6
HF	G58..00	heart failure	Definite	77	23.1	23.1
	G580.11	congestive cardiac failure	Definite	44	13.2	36.3
	G581.00	left ventricular failure	Definite	40	12	48.3
	G580.00	congestive heart failure	Definite	34	10.2	58.6
	G5yy900	left ventricular systolic dysfunction	Definite	23	6.9	65.5
Pericarditis	G534.00	pericardial effusion - acute	Definite	8	36.4	36.4
	G536.00	pericardial effusion	Definite	6	27.3	63.6
	G50..00	acute pericarditis	Definite	3	13.6	77.3
	G50z.00	other and unspecified acute pericarditis	Definite	3	13.6	90.9
	G50..11	pericardial effusion - acute	Definite	1	4.5	95.5
VHD	G540.16	mitral regurgitation	Definite	58	30.9	30.9
	G541500	aortic stenosis	Definite	53	28.2	59
	G541100	aortic stenosis, non-rheumatic	Definite	25	13.3	72.3
	G541012	aortic regurgitation, non-rheumatic	Definite	10	5.3	77.7
	G540.14	mitral valve regurgitation	Definite	10	5.3	83
DVT	G801.11	deep vein thrombosis	Definite	202	71.1	71.1
	1JH..00	suspected deep vein thrombosis	Possible	42	14.8	85.9
	G801.13	dvt - deep vein thrombosis	Possible	10	3.5	89.4
	8HTm.00	referral to deep vein thrombosis clinic	Definite	10	3.5	93
	G801.00	deep vein phlebitis and thrombophlebitis of the leg	Definite	7	2.5	95.4
Pulmonary Embolism	G401.00	pulmonary embolism	Definite	182	96.3	96.3
	G401.12	pulmonary embolus	Definite	6	3.2	99.5
	1JC..00	suspected pulmonary embolism	Possible	1	.5	100

4.4.3 Objective 3 - Concordance of CVD outcomes in linked CPRD and HES data

The pulmonary embolism outcome demonstrated the largest overlap between events in CPRD and HES (49.66%, Figure 4.1. Table 4.4), and pericarditis the smallest (4.55%). Generally, a larger proportion of incident events could be identified using only HES in comparison with using only CPRD, with over 70% of all events in 8/12 of the CVD outcomes identified when using only HES. Whereas over 70% of all events were identified in 5/12 of the CVD outcomes when using only CPRD. Using only HES also captured a higher proportion of the total incident events in comparison with using only CPRD in 8/12 of the CVD outcomes. When there were events in both CPRD and HES, the median time between those events was zero days for both SCA and stroke (Table 4.4). However, the median time between events was 311.5 days for VHD events, which could either be due to delayed input into the patients' records, or the record of a different, unrelated event that followed the incident event.

Figure 4.1: Overlap of incident CVD events between CPRD and HES

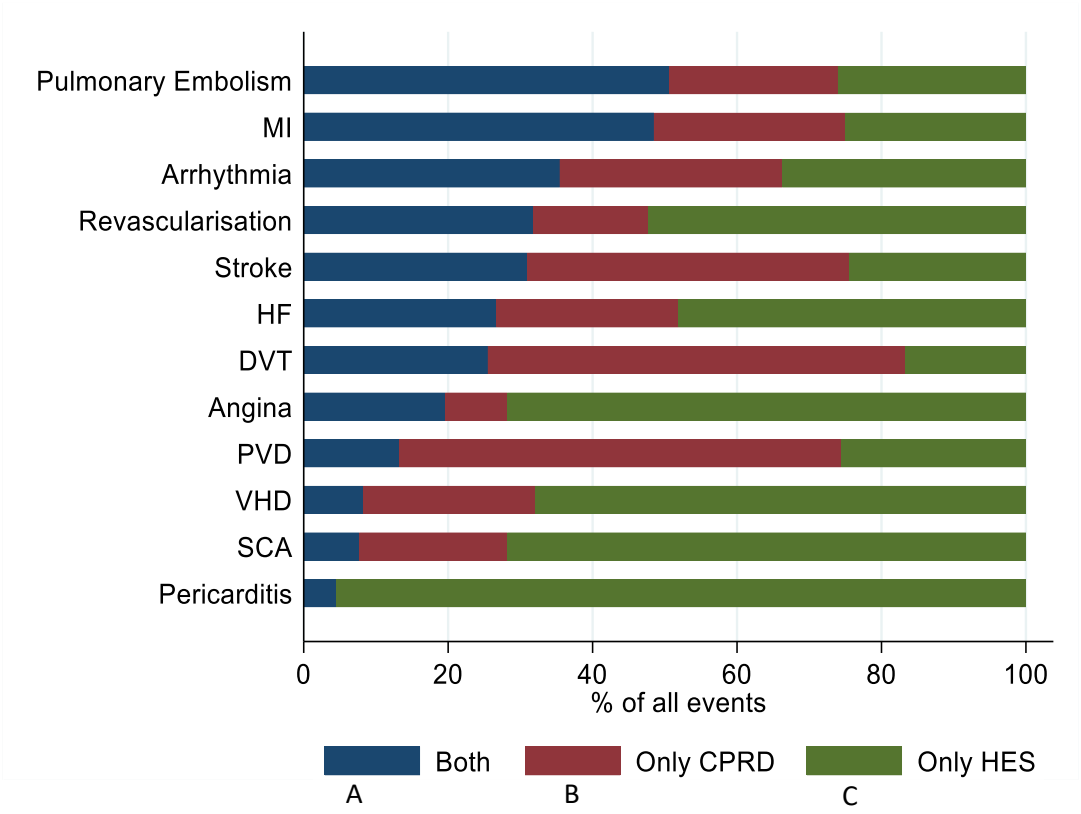


*Table 4.4: Concordance of CVD outcomes in linked CPRD and HES**

Outcome	Women in study population	Incident events when using both CPRD and HES (A+B+C); (% of study population)	Incident events identified from CPRD (A+B); (% of total events from both sources)	Incident events identified from HES (A+C): (% of total events from both sources)	Overlap between CPRD and HES (C); (% of total events from both sources), C)	Median time between events in CPRD and HES
Angina	9234	167 (1.81)	46 (27.54)	149 (89.22)	32 (19.16)	82
MI	9779	132 (1.35)	99 (75.00)	97 (73.48)	64 (48.48)	1
Revascularisation	9916	44 (0.44)	21 (47.73)	37 (84.09)	14 (31.82)	5
SCA	9995	39 (0.39)	11 (28.21)	31 (79.49)	3 (7.69)	0
PVD	9772	98 (1.00)	73 (74.49)	38 (38.78)	13 (13.27)	103
Stroke	9582	296 (3.09)	222 (75.00)	163 (55.07)	91 (30.74)	0
Arrhythmia	9060	542 (5.98)	344 (63.47)	359 (66.24)	184 (33.95)	102.5
HF	9532	344 (3.61)	173 (50.29)	249 (72.38)	89 (25.87)	33
Pericarditis	9986	22 (0.22)	1 (4.55)	22 (100.00)	1 (4.55)	12
VHD	9755	223 (2.29)	70 (31.39)	166 (74.44)	18 (8.07)	311.5
DVT	9681	204 (2.11)	170 (83.33)	86 (42.16)	52 (25.49)	2
Pulmonary Embolism	9832	149 (1.52)	108 (72.48)	112 (75.17)	74 (49.66)	5

*Coverage for CPRD (B) 2002-2017, coverage for HES (C) 2002-2016, and coverage for both data sets (A) 2002-2016

4.4.4 Objective 4 - Exploration of CPRD file around the time of HES only events

After assessing all Read codes in the CPRD clinical and referral files within a 6-month period around an event that was only recorded in HES, new Read codes were identified for arrhythmia, valvular heart disease, and revascularisation. The additional Read codes and descriptions were as follows:

- ***Arrhythmia***
    - 181..00 – Palpitations
- ***Pericarditis***
    - G536.00 – Pericardial effusion
    - G50..11 – Pericardial effusion – acute
    - G534.00 - Pericardial effusion – acute
    - G533.00 – Pericardial effusion – non-inflammatory
    - G536000 – Chronic pericardial effusion
- ***Valvular heart disease***
    - 7911300 - Replacement of aortic valve NEC
    - G541500 - Aortic stenosis
    - G541100 - Aortic stenosis, non-rheumatic
- ***Revascularisation***
    - 7911300 - Replacement of aortic valve NEC
    - 7911.00 - Plastic repair of aortic valve

- 7911.12 - Replacement of aortic valve

The concordance of CPRD and HES events following the addition of these codes to the original Read code lists are outlined in Appendix 4.15. The addition of one new Read code in the arrhythmia code list made the largest difference, identifying 137 additional events in CPRD. However, the addition of new codes for the other three outcomes had minimal impact (between three and nine further events identified in CPRD).

For the five patients that were picked at random for each of the four CVD outcomes with low coverage of events in CPRD (angina, pericarditis, SCA, and VHD, Table 4.5), there was generally no firm conclusions as to why events were recorded in HES, but not CPRD. All patients had evidence of CVD related problems and/or prescriptions for CVD-related drugs, but the specific outcome of interest, as identified in their HES records, could not be ascertained in their CPRD records.

Table 4.5: Overview of the exploration of patients CPRD records around the time of a HES event between 2002 and 2016

Outcome	**Patient**	**Overview of what was found in CPRD**
Angina	1	Been in hospital, on other CVD drugs
	2	Had a fall and seen in casualty, on other CVD drugs
	3	Lots of tests in primary care, pain killers and GI drugs
	4	Other CVD drugs prescribed
	5	Prescribed digoxin for AF, had a stroke not long after
Hypothesis for no Read code	Patients have many other problems. Hard to identify if not coded specifically as angina	
Pericarditis	1	Admitted to hospital and had endoscopy, there was a gastric outlet obstruction
	2	A lot of indigestion tablets, but no other information. Cancer may have gone metastatic
	3	Meningococcal diagnosis
	4	Painkiller prescriptions, and not any diagnoses
	5	Diabetic, and other CVD drugs
Hypothesis for no Read code	Very hard to identify these events in CPRD, with very little evidence of the HES diagnosis	
SCA	1	A lot of tests not long after HES diagnosis, including CVD related tests
	2	Other CVD diagnoses such as heart failure and atrial fibrillation. Also had chest x-rays.
	3	A lot of painkiller prescriptions, and kidney function tests, seems like they have a lot of other problems
	4	A lot of CVD related prescriptions, but no diagnoses
	5	Had a mammography the same day, so possible re-diagnosis of breast cancer
Hypothesis for no Read code	Women seem to have a lot of other problems, so possible that the SCA hasn't been transferred to GPs records	
VHD	1	On several CVD drugs and CVD monitoring
	2	On several CVD drugs
	3	On several CVD drugs, and has diabetes
	4	Has epilepsy and on pain killers for what seems like musculoskeletal problems
	5	On painkillers, and had an x-ray not long after date of VHD
Hypothesis for no Read code	Women have a lot of CVD related problems, but no evidence of VHD diagnoses	

4.4.5 Objective 5 - Validity of HES events

The proportion of all HES events that were classed as the primary diagnosis of the episode ranged from 8.9% for VHD to 69.5% for stroke (Table 4.6). For those events that were classed as the secondary diagnosis, the most common related primary diagnosis within the same episode was from the same ICD-10 chapter as the intended outcome (chapter I) for all outcomes, except pulmonary embolism (Appendix 4.16).

Table 4.6: Total number of HES events, and proportion of events that were the primary diagnoses of the episode between 2002 and 2016

Outcome (ICD letter of outcome)	Total number of events in HES	Number of HES events that were the primary diagnosis (proportion of all HES events)
Angina (I)	167	42 (25.10)
MI (I)	132	79 (59.80)
SCA (I)	38	10 (26.30)
PVD (I)	41	10 (24.40)
Stroke (I)	197	137 (69.50)
Arrhythmia (I)	416	68 (16.30)
HF (I)	322	75 (23.30)
Pericarditis (I)	33	3 (9.10)
VHD (I)	192	17 (8.90)
DVT (I)	100	57 (57.00)
Pulmonary Embolism (I)	147	79 (53.70)

4.4.6 Objective 6 - Calculate the statistical power of analyses using either CPRD data alone or linked CPRD and HES data

The CPRD and HES linked population includes a subset of around 56% of the CPRD population. However, additionally including events recorded in HES allowed accumulation of events not identified in CPRD. Therefore, although there were fewer women in the study population, the statistical power of analyses was generally similar when using the CPRD and HES linked population and identifying CVD events in either CPRD or HES, compared with using CPRD alone to generate the study population and identify CVD events. A lower HR was detectable for 5/12 of the CVD outcomes when using the linked population (Table 4.7). However, the differences for many other outcomes were negligible.

Table 4.7: Detectable HRs for all CVD outcomes when using CPRD alone and linked CPRD and HES data between 2002 and 2016

Outcome	Detectable HR	
	CPRD only (events)	CPRD and HES (events)
Angina	1.77 (97)	1.54 (167)
MI	1.56 (162)	1.63 (132)
Revascularisation	2.07 (60)	2.27 (47)
SCA	3.78 (18)	2.46 (39)
PVD	1.66 (123)	1.76 (98)
Stroke	1.34 (365)	1.39 (296)
Arrhythmia	1.21 (908)	1.24 (679)
HF	1.36 (333)	1.35 (344)
Pericarditis	3.33 (22)	2.94 (27)
VHD	1.51 (188)	1.45 (232)
DVT	1.4 (284)	1.48 (204)
Pulmonary Embolism	1.51 (189)	1.58 (149)

4.5 DISCUSSION

The analyses in this chapter carried out a full exploration of CVD outcomes in the UK data, allowing decisions to be made on how to define these outcomes when exploring the effect of endocrine therapies.

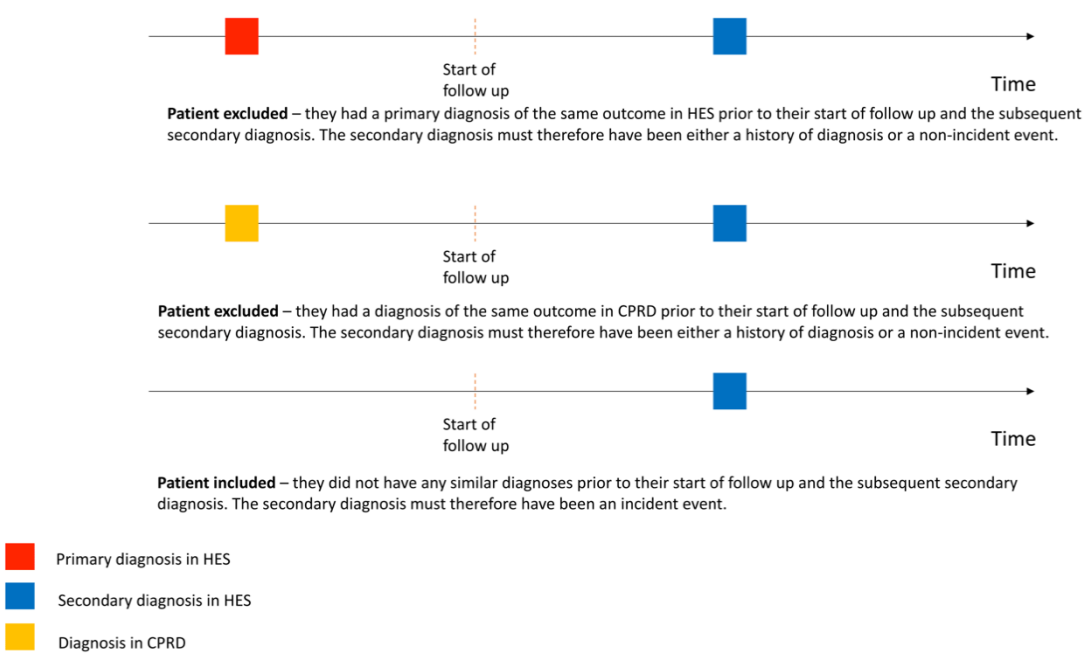
When exploring CVD events in CPRD alone, most of the incident events within follow-up were classed as definite (based on the definition of definite and possible events outlined in the methods). However, even if the incident event was a possible event, a large proportion of these events were confirmed with a definite event within a year. It was therefore decided to include both definite and possible codes in the Read code lists used to identify CVD events in CPRD going forward. Although possible events are not clear clinical diagnoses, the most regularly used possible codes indicate a diagnosis or referral that represents a diagnosis of the CVD of interest, so events will be identified when the GP has not clearly coded the required outcome.

There was generally a low proportion of overlapping events between CPRD and HES when including HES data to identify CVD events. HES captured a larger proportion of events for a small majority of the outcomes, which demonstrates the advantage of including HES data to identify events. However, it also shows that some events may have been missed within CPRD due to incomplete Read code lists. Read code lists were therefore explored further, and additional Read codes for four CVD

outcomes were identified when exploring the time around HES events. However, when further exploring patients' individual records, it was clear that for many patients with an event in HES but no corresponding CPRD record, it would be extremely hard to confidently define the event in CPRD. There are several reasons that an event could be in HES but not CPRD. Firstly, the event in HES could be real, but never fed back to the GP. As there is a distinct disconnect between primary and secondary care in the UK healthcare system, it is possible that a patient received a diagnosis of a disease or condition in hospital, without the GP being made aware. This scenario is even more likely if the patient has several comorbidities, which is probable in older patients diagnosed with breast cancer. Secondly, information of the diagnosis was fed back to the GP, but did they not code the diagnosis correctly in their electronic health record system. Letters from hospital clinicians to patients' GPs inform them of patients' clinical diagnosis in hospital, but if the GP only saves the letter on file, or writes a note in their records, without assigning a Read code, there will be no capture of the diagnosis in CPRD data. Finally, the event may have been miscoded in HES, so the records do not represent a real event. Although possible due to errors by clinicians or coders in hospitals, it is unlikely that this will happen regularly, and any miscoded events are likely to be random.

For many of the outcomes, a low proportion of the events in HES were the primary diagnosis of that episode. However, if the event was a secondary diagnosis, the primary diagnosis of that episode was regularly from the same ICD chapter as the intended outcome. Secondary diagnoses describe those conditions that coexist at the time of admission, or develop subsequently and affect the patient care for the current episode. It is therefore plausible that a secondary diagnosis could be an historical rather than incident event, and it is not possible to distinguish between the two in HES data. However, if the secondary diagnosis was an historical event, then the patient should have records of this event (at the time it occurred) in either their hospital or GP records. As women with historical CVD events are excluded from the study population, it is likely that secondary diagnoses represent conditions that developed subsequently after admission to the hospital, so multiple counting of HES secondary diagnoses is avoided (further diagrammatical explanation in Figure 4.2). Therefore, it was decided to include both primary and secondary diagnoses events identified in HES, as only including primary diagnoses would mean missing a large proportion of genuine incident events.

Figure 4.2: Exclusions of women based on historical CVD events in HES



There were minimal differences in the detectable hazard ratios for most outcomes when using the linked dataset to identify CVD events in CPRD or HES, in comparison with only identifying CVD events from CPRD records in the full CPRD population alone. As including events recorded in HES gives a higher probability of identifying CVD events, with minimal or no loss in statistical power, it was decided that the UK study should use linked CPRD and HES data for all analyses, whilst identifying events that were recorded in either CPRD or HES. This method will likely have higher sensitivity but lower specificity of outcome detection compared with only using events recorded in both CPRD and HES, as there is a higher probability that the women has had the event if it was recorded in both databases. However, only using events that are in both CPRD and HES could cause under ascertainment of the outcome, potentially missing legitimate events only recorded in one database, and would severely affect statistical power. Using events recorded in either CPRD or HES (more sensitive method) could cause over ascertainment of CVD outcomes, as it is possible that an event is not real if only recorded in one database. However, given the problems previously discussed (reasons that events are in HES but not CPRD), and as women may never go to hospital for less severe CVDs (only to the GP), it was considered that over ascertainment would be less of a problem for this study.

4.6 CONCLUSION

This chapter showed that the original algorithm for creating Read code lists found most events identifiable in the CPRD, but it is not possible to identify all events using CPRD alone. The inclusion of possible CVD events in CPRD and events recorded in either CPRD or HES increases the sensitivity of CVD outcome identification in the UK study, with limited impact on statistical power due to having to restrict the study population to women included in the CPRD and HES linkage scheme.

4.7 SUMMARY

- The aim of this chapter was to develop CVD outcome definitions in UK data, assess the validity of UK primary and secondary care data to detect CVD outcomes, and decide if the main analyses in the UK should use either the CPRD data alone or the CPRD and HES linked data.
- The algorithm used to create the original CVD Read code lists was presented, and the number of codes detected to identify CVD outcomes ranged from 128 in the revascularisation outcome to 20 in the PE outcome.
- Most incident CVD events in CPRD were definite events, and a large proportion of possible events were followed by a definite event within a year, so it was decided to include both definite and possible codes in the Read code lists used to identify CVD events in CPRD going forward.
- There was generally a low proportion of overlapping events in both CPRD and HES. However, additional Read codes were identified for four outcomes by exploring diagnoses in CPRD in the time around an event that only occurred in HES, which increased the proportion of events identified using CPRD.
- There were still some outcomes for which few events could be identified in CPRD in comparison with HES, however, there were no clear ways to identify these diagnoses in CPRD. It is likely that these events are missed in CPRD due to records not being fed back to the GP, or the information not being coded correctly by the GP surgery.
- For many CVD events identified in HES, a low proportion were the primary diagnosis of the episode. However, for those that were the secondary diagnosis, the primary diagnosis of the episode was regularly from the same ICD chapter as the intended outcome. The decision was therefore made to include both primary and secondary diagnoses as CVD outcomes in HES.
- Analyses that defined the study population in linked CPRD and HES data, whilst identifying CVD events in either CPRD or HES, had minimal differences in statistical power compared with analyses that only used CPRD to both define the study population and identify CVD events. As including events recorded in HES gives a higher probability of identifying CVD events, with minimal or no loss in statistical power, it was decided that the UK study should use linked CPRD and HES data for all analyses, whilst identifying events that were recorded in either CPRD or HES.

5 ASSESSING THE EFFECT OF ENDOCRINE THERAPY USE ON THE RISK OF CARDIOVASCULAR DISEASE: A COHORT STUDY IN THE UK

This chapter is the final draft of a paper assessing the effect of endocrine therapy use on the risk of cardiovascular disease in a UK population. This draft will soon be submitted to PLOS Medicine. All appendices that will be submitted alongside the paper are available at the end of the thesis.

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	493367	**Title**	Mr
First Name(s)	Anthony		
Surname/Family Name	Matthews		
Thesis Title	Adjuvant endocrine therapy use in female breast cancer survivors and the risk of cardiovascular disease		
Primary Supervisor	Krishnan Bhaskaran		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	PLOS Medicine
Please list the paper's authors in the intended authorship order:	Anthony Matthews, Susannah Stanway, Alexander R Lyon, Liam Smeeth, Krishnan Bhaskaran
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I formulated the study design with input from my primary supervisor. I then carried out data extraction, data manipulation, analyses, and wrote first draft. All authors then contributed to further drafts.

SECTION E

Student Signature	[redacted]
Date	12/01/2019

Supervisor Signature	[redacted]
Date	12/01/2019

5.1 PRE-SUBMISSION MANUSCRIPT DRAFT

Assessing the effect of endocrine therapy use on the risk of cardiovascular disease: A cohort study in the UK

Anthony Matthews*, research fellow*[1], Susannah Stanway, *consultant in medical oncology*[2], Alexander R Lyon, *clinical senior lecturer*[3,4], Liam Smeeth, *professor*[1], Krishnan Bhaskaran, *associate professor*[1]

1 - Department of Non-Communicable Diseases Epidemiology
London School of Hygiene and Tropical Medicine
London, UK

2 - Royal Marsden Hospital
London, UK

3 – National Heart and Lung Institute
Imperial College London
London, UK

4 – Royal Brompton Hospital
London, UK

Corresponding author
Anthony Matthews
Email: anthony.matthews@lshtm.ac.uk
Address: Department of Non-Communicable Diseases Epidemiology, London School of Hygiene and Tropical Medicine, London, UK
Telephone number: +4420 7927 2040

ABSTRACT

Background

Tamoxifen is known to be associated with an increased risk of thromboembolic events, but the long-term effect of aromatase inhibitor (AI) use on risk of cardiovascular diseases are unclear. We aimed to examine the effect of AIs in comparison with tamoxifen on the risk of a comprehensive range of cardiovascular outcomes in postmenopausal breast cancer survivors.

Methods

Using UK primary and secondary care data, we assembled a cohort of postmenopausal female breast cancer survivors, who were prescribed tamoxifen or AIs from 2002-2016. The outcomes were a range of incident cardiovascular events. Associations between endocrine therapy (AI vs tamoxifen) and CVD outcomes were analysed using Cox regression, adjusted for potential confounders.

Results

10005 women were included: 4716 and 5289 were originally prescribed tamoxifen and AIs respectively. There was a pattern of an increased risk of non-venous CVDs in AI compared with tamoxifen users, with evidence of an increased risk of heart failure and arrhythmias in women ever exposed to AI compared with tamoxifen (adjusted hazard ratio (HR): 1.70, 95% confidence interval (CI): 1.26-2.29; adjusted HR: 1.38, 95% CI 1.12-1.70 respectively), which may be driven by cardio-protective effects of tamoxifen use, as we also found past tamoxifen users at a higher risk of both outcomes compared with current users. As expected, AI use was associated with a lower risk of deep vein thrombosis than tamoxifen

Conclusion

We found results that suggest an increased risk of several non-venous CVD outcomes associated with AI in comparison with tamoxifen use, with varying levels of precision, and evidence of an increased risk of heart failure and arrhythmia during AI use in comparison with tamoxifen use, which is likely driven by a protective effect of tamoxifen. Our results also confirmed the established increased risk of VTE associated with tamoxifen use.

BACKGROUND

Endocrine therapies, namely tamoxifen and aromatase inhibitors (AIs), reduce the recurrence rate of oestrogen and progesterone receptor (ER) positive breast cancer, with AIs being more efficacious in postmenopausal women.[1-3] Tamoxifen use is known to be associated with an increased risk of venous thromboembolic events,[4-14] but there have been recent concerns about excess cardiovascular disease (CVD) side effects of AIs in comparison with tamoxifen.

Several randomised control trials (RCTs) have compared tamoxifen with AIs for efficacy;[7-10 15-23] some have compared CVD risks as a secondary outcome, but there has generally been inadequate statistical power to detect associations; more recently a few observational studies have attempted to explore the long-term CVD effect of AIs relative to tamoxifen;[24 25] but most previous work has focused on composite CVD outcomes, with little evidence available on specific CVDs. A recent meta-analysis of RCTs reported an increased risk of non-venous CVDs in comparison with tamoxifen (RR: 1.19, 95% confidence interval (CI): 1.07–1.34), with authors concluding that this is likely due to cardio-protective effects of tamoxifen.[15] Another systematic review also collated RCT and observational evidence on the effect of endocrine therapies on the risk of specific CVD outcomes, and results were consistent with a higher risk of the vascular CVDs myocardial infarction and angina in AI compared with tamoxifen users, which is again likely due to a protective effect of tamoxifen on these outcomes. The review also suggested an increased risk of venous thromboembolic outcomes in tamoxifen users compared with both non-users and AI users.[26]

Given the ongoing uncertainty, limited real world evidence, and clinical importance, we aimed to examine the effect of AIs in comparison with tamoxifen on the risk of a comprehensive range of CVD outcomes in female post-menopausal breast cancer survivors in the UK.

METHODS

Study Design and Data Source

We carried out a cohort study using prospectively collected data from the UK Clinical Practice Research Datalink (CPRD) and linked Hospital Episode Statistics (HES). The CPRD contains anonymised primary care data from general practitioners (GPs) who agreed at the practice level to participate.[27] CPRD covers 7% of the UK population and is broadly representative of the wider population.[28] The database includes diagnoses, prescriptions and tests from primary care, referrals

to specialists, as well as diagnoses and outcomes from secondary care, which are fed back to GPs. Lifestyle and anthropometric measurements are also recorded. Around 56% of practices in the CPRD have linked data available in HES, all of which are in England. HES includes detailed information on hospitalisations.

Study Population

We identified female patients with linked CPRD and HES data aged 54 years and over (median age of the menopause in Europe [29]) with an incident breast cancer diagnosis in CPRD (at least one year of CPRD follow-up prior to first breast cancer for an incident diagnosis, further information on breast cancer definition in Appendix 5.1), who were newly prescribed an AI or tamoxifen in primary care after their diagnosis, from 1st January 2002 (although national recommendations suggested use of AIs for this population from 2006, preliminary analysis showed that third generation aromatase inhibitors anastrozole, exemestane, and letrozole came into widespread use in 2002) to 31st March 2015 (one year prior to the latest date at which CPRD and HES was linked). Follow-up began either one year after the date of breast cancer (to ensure we were studying breast cancer survivors and to separate any acute cardiotoxic effects of other systemic treatments), or at the date of first AI or tamoxifen prescription, whichever occurred latest (hereafter the 'index date'). Patients were excluded if prior to their index date they: died or transferred out of the CPRD; were diagnosed with the CVD event of interest; or had any other cancer diagnoses.

Exposures

Incident tamoxifen and AI exposures were identified using an appropriate prescription code (code lists available at https://doi.org/10.17037/DATA.177, and information on how drug code lists were created is available in Appendix 5.2). The primary exposure was AI use relative to tamoxifen use. To help elucidate how drug exposure is associated with risk, exposure was parameterised in two ways. First, we considered ever exposure to endocrine therapy (ever use of tamoxifen, ever use of AI, ever use of both drugs, Appendix 5.3). If a woman moved between tamoxifen and AI prescriptions, records were time-updated to indicate they had ever been exposed to both drugs from this point forward. Secondly, current exposure to endocrine therapy (categorised as current tamoxifen use, current AI use, no current therapy and previously ever exposed to an AI, no current therapy and previously exposed to tamoxifen only, Appendix 5.4) was time-updated at any changes in therapy. A prescription was defined as continuous if a further prescription followed within 30-days of the

original prescription ending. Appendix 5.5 contains further information on how length of endocrine therapy prescriptions were defined.

Primary Outcomes

The main CVD outcomes of interest were: coronary artery disease (angina, myocardial infarction (MI), revascularisation procedures, sudden cardiac arrest (SCA)); peripheral vascular disease (PVD); stroke; arrhythmia; heart failure (HF, consisting of HF and cardiomyopathy); pericarditis, valvular heart disease (VHD); and venous thromboembolism (VTE) (deep vein thrombosis (DVT), pulmonary embolism (PE)). Composite CVD outcomes and all individual components of the composite outcomes were analysed separately. Events were identified through clinical diagnoses using NHS Read codes in the CPRD and ICD-10 codes in HES (code lists are available at https://doi.org/10.17037/DATA.177, and information on how clinical diagnosis codes were identified are available in Appendix 5.6). Codes that indicated history of an event were included in the code list used to capture past events prior to index date, but were not included when identifying incident events during follow up.

Covariates

Data on the following covariates at index date were extracted from patients' CPRD files for use in the analyses: age (54-59, 60-69, 70+)*;* smoking status (non-smoker, current smoker, ex-smoker)*;* BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker)*;* Index of Multiple Deprivation (IMD) score (level 1-5 based on GP level IMD data)*;* use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); use of anti-platelets; diabetes*;* chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of non-venous CVD (all outcomes except those within the VTE outcome); history of any VTE outcome; year of breast cancer diagnosis; time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year. (Algorithms and code lists used to define confounders in CPRD can be found in Appendix 5.7, and at https://doi.org/10.17037/DATA.177 respectively). Covariate selection of the most important potential confounders was made through discussions with oncologists and cardiologists to understand the variables that may distort the association between the exposure and outcomes due to being independently associated with both the exposure and outcome, and is not an intermediate factor on the causal pathway.

Statistical Analysis

Observation time began at index date and ended at earliest of the following: a CVD event of interest, diagnosis of another cancer, death, transfer out of the CPRD network, or end of follow-up (31st March 2016, the latest date at which CPRD and HES was linked). Prior to exploring the relationship between endocrine therapies and CVD, baseline characteristic distributions of patients who were initially prescribed tamoxifen or AIs were described.

Primary analyses

Number of events and crude incident rates of each outcome of interest were calculated for both parameterisations of the primary exposures. The primary exposure variables were then included in unadjusted (which was adjusted for age due to using age as a timescale, but referred to as unadjusted throughout) and adjusted (accounting for all covariates) Cox regression models with an underlying age timescale, to obtain hazard ratios. Women with missing BMI, smoking status, or alcohol use data (8.7% overall) were excluded (complete case analysis), which is valid in a regression context if missingness is conditionally independent of the outcome.[30]

Secondary analyses

Effect modification

Ever exposure analyses were tested for evidence of effect modification by any prior CVD, age at index (54-59, 60-69, 70+), and time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs, which also implicitly tests the proportional hazards assumption), in all outcomes. Effect modification was tested by including an interaction term between the primary exposure and the potential effect modifier in the fully adjusted models, which was then tested using a Wald test on the interaction term (this tests if each value of the interaction is equal to zero (on the log scale) and gives the corresponding p value, and was carried out using the Stata *testparm* command). Results were not presented if there were no events within any categorisation of a stratified analysis.

Sensitivity and post-hoc and analyses

As guidelines recommended use of AIs in this study population from 2006, and tamoxifen prior to 2006, baseline characteristic distributions of patients who were initially prescribed tamoxifen or AIs

were described before and after the changes in guidelines.

In case of misclassification of exposure status due to delays in patients obtaining their prescriptions, the grace period used to define a continuous prescription was extended from 30 days to 3 months, 6 months, and 1 year in the current prescription analyses in outcomes that showed evidence of effect in primary analyses.

Primary analyses suggested an increased risk of certain outcomes in time previously exposed to endocrine therapies in comparison with time currently exposed tamoxifen. To explore whether this might be driven by reverse causality (i.e. precursors of the outcome causing the patient to stop therapy), we divided the past endocrine therapy categories into recent past use (<6m from stopping) and distant past use (>6m from stopping); we hypothesised that any reverse causality would only affect the recent past use exposure category.

Ever exposure analyses were additionally adjusted for hormone replacement therapy (HRT) use before index date. Women with ER+/PR+ breast cancer are recommended to stop any hormone replacement therapy (HRT) once diagnosed with ER+/PR+ breast cancer, but any long-term effects could potentially act as a confounder (HRT code list are available at https://doi.org/10.17037/DATA.177).

All analyses were performed in STATA 15.[31], and further discussion regarding the choice of study design, covariate adjustment, and model selection is included in Appendix 5.8.

RESULTS

A total of 10005 women aged 54 years and over were prescribed either tamoxifen or an AI following a breast cancer diagnosis during the study period. A STROBE flow diagram is provided in Appendix 5.9. 4716 (47%) and 5289 (53%) women were initially prescribed tamoxifen or an AI respectively. The characteristics of women based on their initial exposure are shown in Table 1. Women originally prescribed AIs were slightly older, with more comorbidities, prior CVDs, and CVD related prescriptions compared with women originally prescribed tamoxifen.

Primary analyses

Ever exposure analyses

Mean follow-up per person for individual clinical CVD outcomes ranged from 3.98 years in arrhythmia analyses to 4.07 years in stroke analyses. After adjustment for all potential confounders, there was evidence of an increased risk of HF and arrhythmia associated with ever AI in comparison with tamoxifen use (adjusted HR: 1.70, 95% CI: 1.26-2.29; adjusted HR: 1.38, 95% CI 1.12-1.70 respectively; Figure 1, Appendix 5.10). A similar pattern was also seen for other non-venous CVDs (other than stroke), albeit with 95% CIs that crossed unity. There was evidence of a decreased risk of DVT associated with ever AI compared with tamoxifen use following adjustment for all potential confounders at baseline (adjusted HR: 0.63, 95% CI: 0.43-0.93).

Current exposure analyses

There was evidence of an increased risk of arrhythmia in all categories in comparison with current tamoxifen use (current AI vs current tamoxifen adjusted HR: 1.45, 95% CI: 1.17-1.81, Figure 2, Appendix 5.11). A similar pattern was seen for coronary artery disease and most of the outcomes that made up this composite outcome (angina, revascularisation, SCA), but precision was limited for these outcomes as 95% CIs crossed unity. There was also an increased risk of HF associated with current AI in comparison with tamoxifen use (adjusted HR: 1.48, 95% CI: 1.07-2.04). In contrast with other outcomes, the highest incidence in the DVT outcome was within the current tamoxifen category; with evidence of a lower risk of DVT associated with current AI use and patients with in those with past exposure to tamoxifen, compared with current tamoxifen use (current AI vs current tamoxifen adjusted HR: 0.50, 95% CI: 0.35-0.72; past tamoxifen only vs current tamoxifen adjusted HR: 0.50, 95% CI: 0.26-0.98), but estimates were less precise for the association between past use of AI and current tamoxifen.

Secondary analyses

Effect modification

All effect modification analyses are presented in Appendices 5.12-5.23. Within the arrhythmia outcome, there was weak evidence of effect modification between ever exposure to endocrine therapy and both time since index date and past CVD (p=0.07 and p=0.06 respectively, Appendix 5.18), with a higher risk of arrhythmia associated with ever AI compared with tamoxifen use in women shortly after index date, and in those with a CVD event prior to index date. There was also

evidence of effect modification between ever exposure to endocrine therapy and time since index date in the composite coronary artery disease outcome (p=0.03, Appendix 5.12), with a higher risk of coronary artery disease associated with ever AI compared with tamoxifen use in women shortly after index date, which attenuated towards the null in later years. There was no further evidence of effect modification by age, past CVD, or time since index date in any other outcome. However, there were generally few events within the stratified analyses for the majority of outcomes, meaning results were not presented for many analyses due to no events within one arm, and little statistical power to estimate effect modification for those that were, including no results for any stratified analyses for the SCA and pericarditis outcomes.

Sensitivity and Post hoc analyses

Characteristics before and after guideline changes

The characteristics of women based on their initial exposure before and after the changes in guidelines in 2006, are shown in Appendix 5.24. Prior to the guideline changes, women that were prescribed AIs were generally older than those prescribed tamoxifen (median age of AI users prior to 2006: 74 years, IQR 65-83; median age of tamoxifen users prior to 2006: 68 years, IQR: 61-77), whereas after the guideline changes women were generally of a similar age (median age of AI users after 2006: 69 years, IQR 63-67; median age of tamoxifen users prior to 2006: 68 years, IQR: 62-75). Women prescribed AIs prior to the guideline changes were also more likely to have prior CVD, both venous and non-venous, compared with AI users after the changes.

Differing grace periods

Evidence of an increased risk of arrhythmia and HF in time previously exposed to both AI and tamoxifen persisted in all grace period variations year (Appendix 5.25). However, the risk of DVT moved away from the null as the grace period length increased when comparing time previously exposed to both AI and tamoxifen only.

Risk of event in first six months after stopping therapy

There continued to be evidence of an increased risk of arrhythmia associated with past use of AI and tamoxifen therapy compared with current tamoxifen use, even >6 months after stopping therapy (Table 2). There also continued to be an increased risk of HF in all time after stopping AI therapy, with a larger HR in the time <6m following prescription end in comparison with the time >6m

following prescription end. There was weak evidence of a small increase in risk of HF in all time after stopping tamoxifen therapy, albeit with 95% CIs that crossed unity.

Additional adjustment for prior HRT use

All ever exposure analyses were the same or extremely similar to the original results after additionally adjusting for HRT use prior to index date (Appendix 5.26).

DISCUSSION

Main findings

In this large population-based cohort study, there was an overall trend of an increased risk of non-venous CVD outcomes (expect stroke) in ever AI compared with tamoxifen users, with 95% CIs that did not cross unity in the arrhythmia and HF outcomes. The risk of arrhythmia and HF was also sustained in the time after AI prescription had ended. However, the increased risk of these outcomes may be due to cardio-protective effects of tamoxifen use as there was evidence of an increased risk of arrhythmia and HF associated with time after exposure to tamoxifen compared with time currently exposed, which was not due to women stopping their prescription shortly before an event. There was also evidence of a decreased risk of DVT in ever and current AI compared with tamoxifen users, which is likely driven by the known increased risk in tamoxifen users.

Another possible explanation for the for the increased risk of arrhythmias and HF in AI users compared with tamoxifen, other than cardio-protective effects of tamoxifen, includes confounding by other cardio-toxic cancer treatments, such as anthracycline chemotherapy and radiotherapy, due to more advanced disease in AI users. We could not explore this possibility as we did not have access to information on cancer treatment or stage and grade of cancer, and although there is no reason to believe that these factors would differ systematically between the comparison groups, any differences in other treatments may explain part of the association. Furthermore, given that more women who initiate AIs have CVD diagnoses prior to index date in comparison with those that initiate tamoxifen (Table 1), it is conceivable that a possible explanation of the results may be confounding by prior CVD. The persistence of an effect in those without any prior CVDs in the stratified HF analyses (Appendix 5.18) suggests that this not be the case for HF, but the attenuation towards the null in those without prior CVD in the stratified arrhythmia analysis (Appendix 5.17) suggests that such confounding may explain at least part of this association. Moreover, as of 2006, it

has been recommended that post-menopausal women diagnosed with ER+ breast cancer be prescribed AIs over tamoxifen.[32] Although the proportion of all users that initiate tamoxifen is decreasing over time, there are still many women that are given this drug over AIs. This may be because of other conditions such low bone mineral density as women prescribed AIs are at an increased risk of fracture in comparison with those prescribed tamoxifen.[33] There may therefore also be residual confounding by reason for initiation of either tamoxifen or AI, which could not be explored, as GP records contain no information on reason for initiation of a certain drug. Finally, there was some evidence of reverse causality in the HF analysis, as there was an inflated increased risk of HF in the first six months after stopping AIs in comparison with the time greater than six months after stopping. This provides some evidence that women stopped their AI treatment due to the precursors of HF, but this did not change any overall conclusions.

Comparison with other studies

Two recent studies by Abdel-Qadir and Haque are the only observational studies to date that have directly compared the use of tamoxifen and AIs associated with the risk of specific CVD outcomes.[24] [25] Abdel-Qadir reported an increased risk of hospitalisation for MI in AI users compared with tamoxifen users (HR: 2.02, 95% CI: 1.16-3.53).[24] We found a similar trend (adjusted HR for ever AI use vs. ever tamoxifen use: 1.53, 95% CI: 0.95-2.48), albeit with a non-significant effect. Further observational evidence have suggested that this effect may be due to a cardio-protective effect of MI in tamoxifen users as two out of four observational studies reported a decreased risk in tamoxifen users compared with non-users.[34-37] Haque reported no evidence of a difference in risk of HF between those exposed to AI and tamoxifen (HR: 0.96, 95% CI: 0.77-1.08),[25] whereas this study reported an increased risk of heart failure in AI users (adjusted HR for ever AI use vs. ever tamoxifen use: 1.70, 95% CI: 1.26, 2.29), but in the Haque study, all women over the age of 18 years were included and follow-up started at breast cancer diagnosis. Haque also reported evidence of an increased risk of a composite outcome that included arrhythmia, dysrhythmia, and pericarditis in AI users compared with tamoxifen users (HR: 1.29, 95% CI: 1.11-1.50), which is similar to the effect we reported in the arrhythmia analysis. Finally, Haque reported weak evidence of a decreased risk of stroke associated with AI compared with tamoxifen use (HR: 0.82, 95% CI: 0.63-1.06), which we did not find.

A meta-analysis of RCTs by Khosrow-Khavar reported an increased risk of a composite non-venous CVD outcome associated with AI compared with tamoxifen use (RR: 1.18, 95% CI: 1.05-1.33),[15] which

was likely the result of a cardio-protective effect of tamoxifen. This is similar to the trend seen in the non-venous CVD outcomes in this study. Most other meta-analyses focused on composite CVD outcomes, but like this study, four out of five individual RCTs reported evidence of a decreased risk of VTE in AI users compared with tamoxifen users, with relative risks ranging from 0.25-0.61.[7-10 20] One RCT also reported evidence of an increased of heart failure in AI users compared with tamoxifen users (RR: 1.20, 95% CI: 1.04-1.38),[10] which is the same direction of effect seen in this study.

A recent systematic review collated all RCT observational evidence and concluded that AI users were at a higher risk of MI and angina in comparison with tamoxifen users, which is again likely driven by a protective effect of tamoxifen.[26] Although we had little precision for these outcomes, the directions of effect in both the ever and current exposure analyses for MI and angina was in a similar to that suggested in the systematic review. The review also suggested an increased risk of VTE outcomes in tamoxifen users compared with AI users, which is consistent with DVT results in this study. However, the results for PE were not precise enough to provide conclusive evidence of an effect in either direction.

Strengths and limitations

As the CPRD is a large dataset and broadly represents the UK population,[28] the findings are likely to be generalisable to postmenopausal women diagnosed with ER+/PR+ breast cancer in both the UK and other developed populations due to the homogenous indication of endocrine therapy worldwide.

ER+/PR status was not available, but it is likely that all breast cancers were ER+/PR+ as such a diagnosis is a prerequisite of being prescribed endocrine therapies. This also meant that we were unable to identify a population of women diagnosed with ER+/PR+ breast cancer who were unexposed to any endocrine therapy to act as a control group. Having such a group would have allowed us detect the effect of both drugs in comparison with those who were free of any potential CVD effect of endocrine therapies. However, this categorisation may also be problematic due to potential differences between those that do and do not initiate endocrine therapies. It would also not be appropriate to compare women diagnosed with ER+/PR+ breast cancer to those with other subtypes of breast cancer (i.e. triple negative) because of the different treatment regimens given to these women and the resulting potential differences in toxicity.

CPRD captures prescriptions at the point of issue, but we had no data on whether prescriptions were

filled, or the drug taken, which could lead to potential misclassification of exposure status. However, further descriptive analyses, using a 6-month grace period to define a continuous prescription, indicated that 97% of women remained on therapy within 1 year of starting, 95% within 3 years, 85% within 5 years, and 29% within 6 years. Hence, it is likely that most women adhered to therapy for the recommended 5-year treatment because if they are continuing to pick up their prescriptions then it is reasonable to assume that they are using the drugs.

It is possible that some of the difference in CVD risk between AI and tamoxifen users is due to the heterogeneous characteristics of tamoxifen and AI users before and after the guideline changes in 2006. AI users prior to 2006 were older with more prevalent CVD, so are likely at a higher risk of further events.

Within women included in the 'past with AI' group in the current exposure analyses, 20% of the time exposed to any endocrine therapy was exposed to tamoxifen. This time was included within `past with AI' group because the primary aim of the analyses was to investigate the CVD effects of AIs. This could help explain the unexpected lack of difference in DVT risk between the past AI and current tamoxifen groups.

Within the patients diagnosed with arrhythmia, 10% had a previous HF diagnosis. It is therefore not surprising to see the association in a similar direction for both arrhythmia and HF, and some of the association between AI use and arrhythmia may be explained by the increased risk of HF with AI use in comparison with tamoxifen.

CONCLUSION

This large population based cohort using data from UK primary and secondary care is the first to explore the relative effect of endocrine therapies on the whole range of clinically specific CVD outcomes. There was a general pattern of an increased risk of non-venous CVDs in AI compared with tamoxifen users, with varying levels of precision. However, results suggested evidence of an increased risk of HF and arrhythmia in AI compared with tamoxifen users. It is thought that the increased risk in AI users is likely due to cardio-protective effects of tamoxifen. A decreased risk of DVT in AI compared with tamoxifen users is also reported, which is likely due to the well-established increased risk associated with tamoxifen use. Although there are several advantages in understanding the effect of endocrine therapies on CVD as whole, such as the potential to change modifiable risk factors such as weight, smoking and alcohol intake, this study shows that these drugs

are not homogenous in their effects on more clinically specific CVD outcomes. Knowledge of these effects is also increasingly important in post-menopausal women diagnosed with ER+/PR+ breast cancer because of the current recommendations to prescribe AIs to this population, and the growing numbers of older women surviving their initial breast cancer diagnosis.

CONFLICTS OF INTEREST

AM has nothing to disclose. SS reports personal fees from Roche, Clinigen, Eli Lilly, and Novartis, outside the submitted work. AL reports personal fees from Servier, Novartis, Pfizer, Roche, Ferring Pharmaceuticals, Clinigen Group, Boehringer Ingelheim, Amgen, Eli Lily, and BMS, outside the submitted work. LS reports grants from Wellcome, during the conduct of the study; grants from Wellcome, MRC, NIHR, BHF, Diabetes UK, and grants and personal fees from GSK, outside the submitted work; and is a trustee of the British Heart Foundation. KB reports grants from Wellcome Trust and the Royal Society, during the conduct of the study.

CONTRIBUTIONS

Study design was decided on by AM and KB. AM carried out all data extraction, manipulation, and analyses. AM wrote the first draft. All authors contributed to further drafts and approved the final manuscript.

FUNDING

This study was supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (grant No 107731/Z/15/Z) held by KB. The Wellcome Trust and the Royal Society had no role in the design, analysis, or writing up of this study.

REFERENCES

1. Early Breast Cancer Trialists' Collaborative G, Davies C, Godwin J, et al. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *Lancet* 2011;378(9793):771-84. doi: http://dx.doi.org/10.1016/S0140-6736(11)60993-8

2. Cuzick J, Sestak I, Baum M, et al. Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 10-year analysis of the ATAC trial. *The Lancet Oncology* 2010;11(12):1135-41. doi: 10.1016/S1470-2045(10)70257-6

3. Early Breast Cancer Trialists' Collaborative G, Dowsett M, Forbes JF, et al. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. *Lancet* 2015;386(10001):1341-52. doi: 10.1016/S0140-6736(15)61074-1

4. Meier CR, Jick H. Tamoxifen and risk of idiopathic venous thromboembolism. *British journal of clinical pharmacology* 1998;45(6):608-12.

5. Hernandez RK, Sorensen HT, Pedersen L, et al. Tamoxifen treatment and risk of deep venous thrombosis and pulmonary embolism: a Danish population-based cohort study. *Cancer* 2009;115(19):4442-9. doi: http://dx.doi.org/10.1002/cncr.24508

6. van de Velde CJ, Rea D, Seynaeve C, et al. Adjuvant tamoxifen and exemestane in early breast cancer (TEAM): a randomised phase 3 trial. *Lancet* 2011;377(9762):321-31. doi: 10.1016/S0140-6736(10)62312-4

7. Jakesz R, Jonat W, Gnant M, et al. Switching of postmenopausal women with endocrine-responsive early breast cancer to anastrozole after 2 years' adjuvant tamoxifen: combined results of ABCSG trial 8 and ARNO 95 trial. *Lancet* 2005;366(9484):455-62. doi: 10.1016/S0140-6736(05)67059-6

8. Boccardo F, Rubagotti A, Guglielmini P, et al. Switching to anastrozole versus continued tamoxifen treatment of early breast cancer. Updated results of the Italian tamoxifen anastrozole (ITA) trial. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 2006;17 Suppl 7:vii10-4. doi: 10.1093/annonc/mdl941

9. Arimidex TAoiCTG, Forbes JF, Cuzick J, et al. Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 100-month analysis of the ATAC trial. *Lancet Oncology* 2008;9(1):45-53.

10. Colleoni M, Giobbie-Hurder A, Regan MM, et al. Analyses adjusting for selective crossover show improved overall survival with adjuvant letrozole compared with tamoxifen in the BIG 1-98 study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2011;29(9):1117-24. doi: 10.1200/JCO.2010.31.6455 [published Online First: 2011/02/16]

11. McDonald CC, Alexander FE, Whyte BW, et al. Cardiac and vascular morbidity in women receiving adjuvant tamoxifen for breast cancer in a randomised trial. The Scottish Cancer Trials Breast Group. *Bmj* 1995;311(7011):977-80.

12. Fisher B, Dignam J, Wolmark N, et al. Tamoxifen in treatment of intraductal breast cancer: National Surgical Adjuvant Breast and Bowel Project B-24 randomised controlled trial. *Lancet* 1999;353(9169):1993-2000. doi: 10.1016/S0140-6736(99)05036-9

13. Rutqvist LE, Mattsson A. Cardiac and thromboembolic morbidity among postmenopausal women with early-stage breast cancer in a randomized trial of adjuvant tamoxifen. The Stockholm Breast Cancer Study Group. *Journal of the National Cancer Institute* 1993;85(17):1398-406.

14. Fisher B, Anderson S, Tan-Chiu E, et al. Tamoxifen and chemotherapy for axillary node-negative, estrogen receptor-negative breast cancer: findings from National Surgical Adjuvant Breast and Bowel Project B-23. *J Clin Oncol* 2001;19(4):931-42. doi: 10.1200/JCO.2001.19.4.931

15. Khosrow-Khavar F, Filion KB, Al-Qurashi S, et al. Cardiotoxicity of aromatase inhibitors and tamoxifen in postmenopausal women with breast cancer: a systematic review and meta-analysis of randomized controlled trials. *Ann Oncol* 2017;28(3):487-96. doi: 10.1093/annonc/mdw673

16. Ryden L, Heibert Arnlind M, Vitols S, et al. Aromatase inhibitors alone or sequentially combined with tamoxifen in postmenopausal early breast cancer compared with tamoxifen or placebo - Meta-analyses on efficacy and adverse events based on randomized clinical trials. *Breast* 2016;26:106-14. doi: 10.1016/j.breast.2016.01.006

17. Cuppone F, Bria E, Verma S, et al. Do adjuvant aromatase inhibitors increase the cardiovascular risk in postmenopausal women with early breast cancer? Meta-analysis of randomized trials. *Cancer* 2008;112(2):260-7.

18. Aydiner A. Meta-analysis of breast cancer outcome and toxicity in adjuvant trials of aromatase inhibitors in postmenopausal women. *Breast* 2013;22(2):121-9. doi: http://dx.doi.org/10.1016/j.breast.2013.01.014

19. Amir E, Seruga B, Niraula S, et al. Toxicity of adjuvant endocrine therapy in postmenopausal breast cancer patients: a systematic review and meta-analysis. *Journal of the National Cancer Institute* 2011;103(17):1299-309. doi: http://dx.doi.org/10.1093/jnci/djr242

20. Bliss JM, Kilburn LS, Coleman RE, et al. Disease-related outcomes with long-term follow-up: an updated analysis of the intergroup exemestane study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2012;30(7):709-17. doi: 10.1200/JCO.2010.33.7899

21. Coombes RC, Kilburn LS, Snowdon CF, et al. Survival and safety of exemestane versus tamoxifen after 2-3 years' tamoxifen treatment (Intergroup Exemestane Study): a randomised controlled trial. *Lancet* 2007;369(9561):559-70. doi: 10.1016/S0140-6736(07)60200-1

22. Abo-Touk NA, Sakr HA, Abd El-Lattef A. Switching to letrozole versus continued tamoxifen therapy in treatment of postmenopausal women with early breast cancer. *J Egypt Natl Canc Inst* 2010;22(1):79-85. [published Online First: 2011/04/20]

23. Kaufmann M, Jonat W, Hilfrich J, et al. Improved overall survival in postmenopausal women with early breast cancer after anastrozole initiated after treatment with tamoxifen compared with continued tamoxifen: the ARNO 95 Study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2007;25(19):2664-70. doi: 10.1200/JCO.2006.08.8054

24. Abdel-Qadir H, Amir E, Fischer HD, et al. The risk of myocardial infarction with aromatase inhibitors relative to tamoxifen in post-menopausal women with early stage breast cancer. *Eur J Cancer* 2016;68:11-21. doi: 10.1016/j.ejca.2016.08.022

25. Haque R, Shi J, Schottinger JE, et al. Cardiovascular Disease After Aromatase Inhibitor Use. *JAMA oncology* 2016 doi: 10.1001/jamaoncol.2016.0429

26. Matthews A, Stanway S, Farmer RE, et al. Long term adjuvant endocrine therapy and risk of cardiovascular disease in female breast cancer survivors: systematic review. *Bmj* 2018;363:k3845. doi: 10.1136/bmj.k3845

27. Herrett E, Gallagher AM, Bhaskaran K, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *International journal of epidemiology* 2015;44(3):827-36. doi: 10.1093/ije/dyv098

28. Campbell JD, D.J. Eaton, S.C. Gallagher, A.M. Williams, T.J. Is the CPRD GOLD Population Comparable to the U.K. Population? *Pharmacoepidemiology and Drug Safety* 2013;22(Suppl 1):280.

29. Dratva J, Gomez Real F, Schindler C, et al. Is age at menopause increasing across Europe? Results on age at menopause and determinants from two population-based studies. *Menopause* 2009;16(2):385-94. doi: 10.1097/gme.0b013e31818aefef

30. White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. *Stat Med* 2010;29(28):2920-31. doi: 10.1002/sim.3944

31. Stata Statistical Software: Release 15 [program]. College Station, TX: StataCorp LLC, 2017.

32. National Institute for Health and Care Excellence. Hormonal therapies for the adjuvant treatment of early oestrogen-receptor-positive breast cancer 2006 [Available from: https://www.nice.org.uk/guidance/TA112/documents/final-appraisal-determination2 accessed 14/04/2016.

33. Tseng OL, Spinelli JJ, Gotay CC, et al. Aromatase inhibitors are associated with a higher fracture risk than tamoxifen: a systematic review and meta-analysis. *Ther Adv Musculoskelet Dis* 2018;10(4):71-90. doi: 10.1177/1759720X18759291

34. Bradbury BD, Lash TL, Kaye JA, et al. Tamoxifen-treated breast carcinoma patients and the risk of acute myocardial infarction and newly-diagnosed angina. *Cancer* 2005;103(6):1114-21.

35. Yang TL, Wu TC, Huang CC, et al. Association of tamoxifen use and reduced cardiovascular events among asian females with breast cancer. *Circulation Journal* 2014;78(1):135-40.

36. Ligibel JA, James O'Malley A, Fisher M, et al. Risk of myocardial infarction, stroke, and fracture in a cohort of community-based breast cancer patients. *Breast Cancer Research & Treatment* 2012;131(2):589-97. doi: http://dx.doi.org/10.1007/s10549-011-1754-1

37. Geiger AM, Chen W, Bernstein L. Myocardial infarction risk and tamoxifen therapy for breast cancer. *British journal of cancer* 2005;92(9):1614-20.

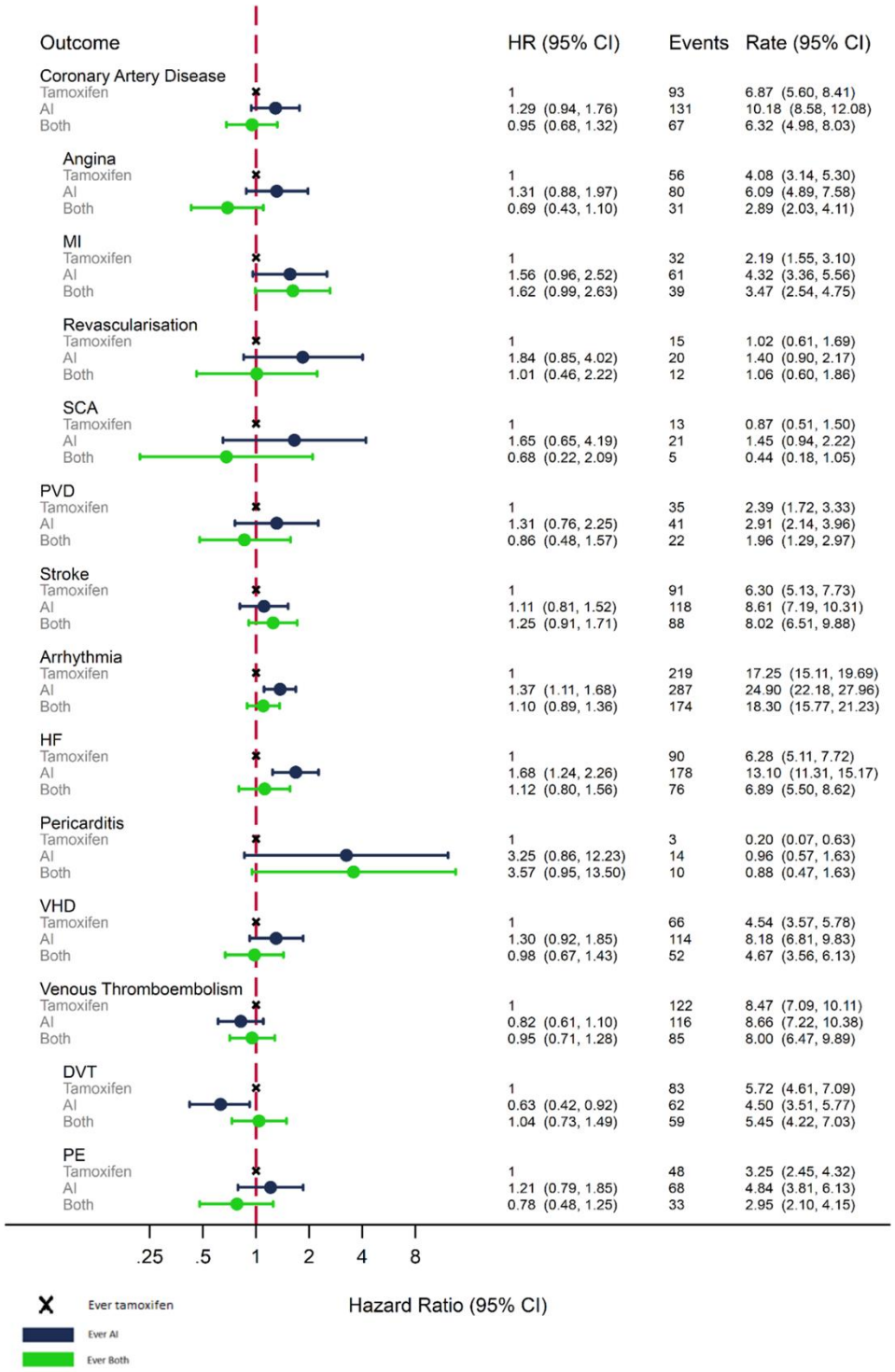
Table 1: Characteristics of study population based on their initial exposure

	Tamoxifen	AI	Total
N	4716 (100)	5289 (100)	10005 (100)
Age (yrs)			
54-59	911 (19.3)	716 (13.5)	1627 (16.3)
60-69	1850 (39.2)	1972 (37.3)	3822 (38.2)
70+	1955 (41.5)	2601 (49.2)	4556 (45.5)
Median (IQR)	68 (62-76)	70 (63-79)	69 (62-78)
Year of breast cancer			
2002	579 (12.3)	72 (1.4)	651 (6.5)
2003	605 (12.8)	123 (2.3)	728 (7.3)
2004	571 (12.1)	186 (3.5)	757 (7.6)
2005	512 (10.9)	289 (5.5)	801 (8)
2006	467 (9.9)	370 (7)	837 (8.4)
2007	386 (8.2)	448 (8.5)	834 (8.3)
2008	380 (8.1)	495 (9.4)	875 (8.7)
2009	290 (6.1)	533 (10.1)	823 (8.2)
2010	223 (4.7)	586 (11.1)	809 (8.1)
2011	216 (4.6)	619 (11.7)	835 (8.3)
2012	210 (4.5)	539 (10.2)	749 (7.5)
2013	174 (3.7)	504 (9.5)	678 (6.8)
2014	93 (2)	451 (8.5)	544 (5.4)
2015	10 (.2)	74 (1.4)	84 (.8)
BMI (kg/m2)			
<18	59 (1.3)	63 (1.2)	122 (1.2)
18-24	1693 (35.9)	1619 (30.6)	3312 (33.1)
25-29	1549 (32.8)	1801 (34.1)	3350 (33.5)
30-34	800 (17)	979 (18.5)	1779 (17.8)
≥35	345 (7.3)	548 (10.4)	893 (8.9)
Missing	270 (5.7)	279 (5.3)	549 (5.5)
Median (IQR)	26 (23-30)	27 (24-31)	27 (24-31)
Smoking status			
Never smoker	2423 (51.4)	2517 (47.6)	4940 (49.4)
Current smoker	503 (10.7)	482 (9.1)	985 (9.8)
Ex-smoker	1761 (37.3)	2268 (42.9)	4029 (40.3)
Missing	29 (.6)	22 (.4)	51 (.5)
Alcohol use			
Non drinker	618 (13.1)	613 (11.6)	1231 (12.3)
Current	3320 (70.4)	3628 (68.6)	6948 (69.4)
Ex-drinker	480 (10.2)	715 (13.5)	1195 (11.9)
Missing	298 (6.3)	333 (6.3)	631 (6.3)
Systolic BP			
Low/ideal	530 (11.2)	599 (11.3)	1129 (11.3)
Pre-high	1862 (39.5)	2327 (44)	4189 (41.9)
High	2314 (49.1)	2355 (44.5)	4669 (46.7)
Missing	10 (.2)	8 (.2)	18 (.2)
Diastolic BP			
Low/ideal	2130 (45.2)	2650 (50.1)	4780 (47.8)
Pre-high	1988 (42.2)	2058 (38.9)	4046 (40.4)
High	588 (12.5)	573 (10.8)	1161 (11.6)
Missing	10 (.2)	8 (.2)	18 (.2)
IMD category			
1	870 (18.4)	962 (18.2)	1832 (18.3)
2	943 (20)	1214 (23)	2157 (21.6)
3	925 (19.6)	1054 (19.9)	1979 (19.8)
4	1052 (22.3)	936 (17.7)	1988 (19.9)
5	926 (19.6)	1122 (21.2)	2048 (20.5)
Missing	0 (0)	1 (0)	1 (0)
Statins before index	1100 (23.3)	1903 (36)	3003 (30)
ACEI before index	1195 (25.3)	1823 (34.5)	3018 (30.2)
CCB before index	1195 (25.3)	1764 (33.4)	2959 (29.6)
ARB before index	493 (10.5)	795 (15)	1288 (12.9)
Anti-platelets before index	1132 (24)	1639 (31)	2771 (27.7)
RA before index	138 (2.9)	137 (2.6)	275 (2.7)
Diabetes before index	462 (9.8)	728 (13.8)	1190 (11.9)
CKD before index	867 (18.4)	1090 (20.6)	1957 (19.6)
Non-venous CVD before index	1031 (21.9)	1636 (30.9)	2667 (26.7)
VTE before index	144 (3.1)	324 (6.1)	468 (4.7)

Table 2: Adjusted HRs for the risk of event with time after stopping prescription split into <6m and >6m after stopping

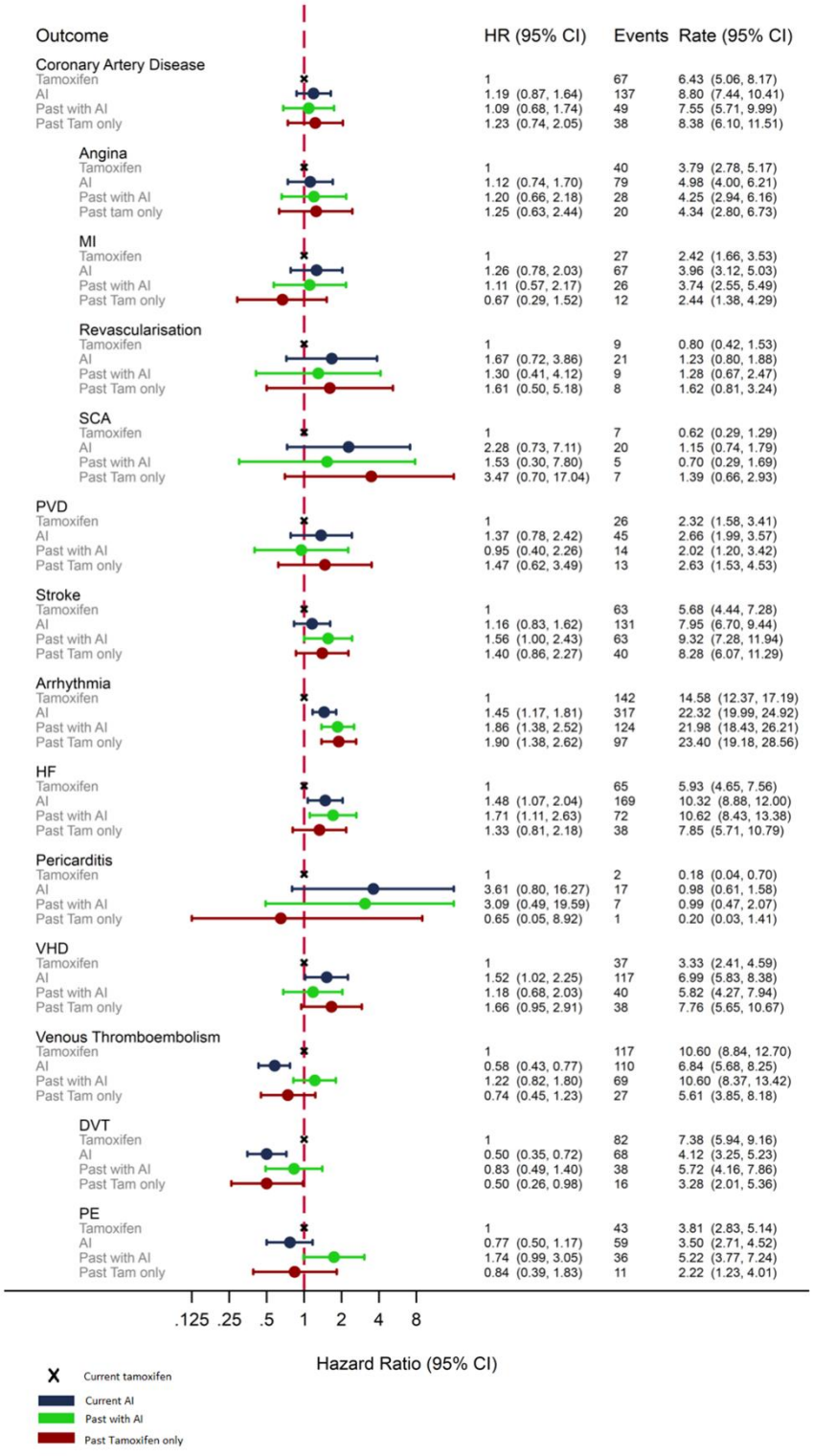
Outcome	Current exposure	Adjusted HR
Arrhythmia	Tamoxifen	1
	AI	1.47 (1.18, 1.83)
	Past with AI <6m	1.70 (1.11, 2.61)
	Past with AI >6m	2.10 (1.50, 2.94)
	Past with tam <6m	1.25 (0.67, 2.33)
	Past with tam >6m	2.19 (1.54, 3.11)
HF	Tamoxifen	1
	AI	1.49 (1.08, 2.06)
	Past with AI <6m	2.09 (1.20, 3.64)
	Past with AI >6m	1.55 (0.95, 2.54)
	Past with tam <6m	1.17 (0.46, 2.98)
	Past with tam >6m	1.29 (0.75, 2.22)

Figure 1: Adjusted HRs, events, and crude rate for the association between ever exposure to endocrine therapy and a range of clinical CVD outcomes



*Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+); smoking status (non-smoker, current smoker, ex-smoker); BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker); IMD score (level 1-5 based on GP level IMD data); use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes; chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of VTE; history of non-venous CVD year of breast cancer diagnosis; time since index (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year.

Figure 2: Adjusted HRs, events, and crude rate for the association between current exposure to endocrine therapy and a range of clinical CVD outcomes



Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+);* smoking status (non-smoker, current smoker, ex-smoker)*;* BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker)*;* IMD score (level 1-5 based on GP level IMD data)*;* use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes*;* chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of VTE; history of non-venous CVD year of breast cancer diagnosis; time since index (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year.

6 ASSESSING THE EFFECT OF ENDOCRINE THERAPY USE ON THE RISK OF CARDIOVASCULAR DISEASE: A COHORT STUDY USING THE US SEER-MEDICARE LINKED DATABASE

This chapter is the final draft of a paper assessing the effect endocrine therapy use on the risk of cardiovascular disease in a US population. This draft will soon be submitted to JAMA Oncology. All appendices that will be submitted alongside the paper are available at the end of the thesis.

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	493367	**Title**	Mr
First Name(s)	Anthony		
Surname/Family Name	Matthews		
Thesis Title	Adjuvant endocrine therapy use in female breast cancer survivors and the risk of cardiovascular disease		
Primary Supervisor	Krishnan Bhaskaran		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	JAMA Oncology
Please list the paper's authors in the intended authorship order:	Anthony Matthews, Sharon Peacock Hinton, Susannah Stanway, Alexander R Lyon, Liam Smeeth, Jennifer L. Lund, Krishnan Bhaskaran
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I formulated the study design with input from my primary supervisor and Jennifer Lund from UNC. I then carried out data manipulation, analyses, and wrote first draft. All authors then contributed to further drafts.

SECTION E

Student Signature	[redacted]
Date	12/01/2019

Supervisor Signature	[redacted]
Date	12/01/2019

6.1 PRE-SUBMISSION MANUSCRIPT DRAFT

Assessing the effect of endocrine therapy use on the risk of cardiovascular disease: a cohort study using the US SEER-Medicare linked database

Anthony Matthews[1,2], Sharon Peacock Hinton[2], Susannah Stanway[3], Alexander R Lyon [4,5], Liam Smeeth[1], Jennifer L. Lund[2]*, Krishnan Bhaskaran[1]*

*Equal contribution

1 - Department of Non-Communicable Diseases Epidemiology
London School of Hygiene and Tropical Medicine
London, UK

2 - Department of Epidemiology
Gillings School of Global Public Health
University of North Carolina at Chapel Hill
Chapel Hill, USA

3 - Royal Marsden Hospital
London, UK

4 – National Heart and Lung Institute
Imperial College London
London, UK

5 – Royal Brompton Hospital
London, UK

Corresponding author
Anthony Matthews
Email: anthony.matthews@lshtm.ac.uk
Address: Department of Non-Communicable Diseases Epidemiology, London School of Hygiene and Tropical Medicine, London, UK
Telephone number: +4420 7927 2040

ABSTRACT

Background

The long-term adverse cardiovascular consequences of endocrine therapies, including tamoxifen and aromatase inhibitors, remain unclear, and clinical trials have not been statistically powered to detect adverse cardiovascular effects. We aimed to examine the effect of tamoxifen and aromatase inhibitors (AI) on the risk of a comprehensive range of cardiovascular disease (CVD) outcomes in female breast cancer survivors aged 66 years and over in the United States.

Methods

We carried out a cohort study using prospectively collected data from the US SEER-Medicare linked database. We identified all women with Medicare Parts A, B and D coverage, aged 66 years and over with, with an incident ER or PR positive and stage 1-3 breast cancer diagnosis between 1st January 2008 and 31st December 2013. Women with endocrine therapy prescriptions prior to their breast cancer diagnosis were excluded. Tamoxifen and AI exposures and CVD outcomes were identified using National Drug Codes, diagnosis codes, and procedure codes. Cox proportional hazards regression models were fitted adjusted for a priori-specified potential confounders.

Results

22027 women with hormone-receptor positive breast cancer were included; initial endocrine treatment was an AI for 15074 (68%), tamoxifen for 2286 (10%) and no endocrine therapy for 4667 (22%) women. There was generally a pattern of a decreased risk of non-venous CVD outcomes associated with ever tamoxifen use with adjusted hazard ratios (HR) ranging from 0.44 (95% confidence interval (CI): 0.30-0.63) in the MI analysis to 0.91 (95% CI: 0.75-1.10) in the PVD analysis. There was also evidence of a lower risk of several non-venous CVD outcomes among AI users compared with the unexposed. As expected there were more deep vein thromboses (DVT) in women ever exposed to tamoxifen compared with the unexposed (adjusted HR and 95% CI: 1.41, 0.98-2.04), but no evidence of association between AI use and DVT (adjusted HR and 95% CI: 1.14, 0.86-1.52).

Conclusions

Apart from the established association between tamoxifen and increased risk of DVT, there was no evidence of increased CVD risk with either tamoxifen or AI use compared with no endocrine therapy. Our results suggested a protective effect of tamoxifen use on the risk of the major arterial cardiovascular diseases. A similar, but weaker, protective association was also seen for arterial disease during AI use, but there was a possibility of residual confounding that bring questions to the causality of these associations.

INTRODUCTION

Breast cancer remains the most common cancer among women worldwide and recent advances in treatment mean that cancer-free survival is increasing.[1-3] However, prolonged survival also means that adverse effects and delayed toxicities of treatment have the potential to substantially impact both patient outcomes and medical resources.

A particular concern regarding treatment effects centres upon CVD risk of endocrine therapies, which are widely used in women with oestrogen or progesterone receptor positive (ER+/PR+) tumours. Several trials, in which the primary outcome was breast cancer reoccurrence, have compared the use of the endocrine therapies tamoxifen and AIs,[4-11] and it has been consistently shown that AIs increase survival compared with tamoxifen in post-menopausal women diagnosed with ER+/PR+ breast cancer. Many of these trials also reported the CVD effects of these drugs, but they were regularly underpowered to detect such associations and reported effects on composite CVD outcomes, with little evidence available for specific CVD events. The most recent meta-analysis of RCTs reported an increased risk of non-venous CVDs in comparison with tamoxifen (RR: 1.19, 95% confidence interval (CI): 1.07–1.34), with authors concluding that this is likely due to cardio-protective effects of tamoxifen.[12] Recent observational evidence has begun to focus on more specific CVD outcomes,[13-19] with many studies showing a protective association between tamoxifen use compared with non-users and the risk of non-venous CVDs, but little conclusive evidence on the risks associated with AI use. Another recent systematic review collated all RCT and observational evidence on the effect of endocrine therapies on the risk of specific CVD outcomes, and results were consistent with a higher risk of the vascular CVDs myocardial infarction and angina in AI compared with tamoxifen users, which is again likely due to a protective effect of tamoxifen on these outcomes. The review also suggested an increased risk of venous thromboembolic outcomes in tamoxifen users compared with both non-users and AI users.[20] To date, no study has explored the effect of endocrine therapies on the risk of the whole range of clinically specific CVD outcomes in postmenopausal ER+/PR+ breast cancer survivors.

Given the ongoing uncertainty, limited real world evidence, and clinical importance, we aimed to examine the effects of tamoxifen and AIs on a comprehensive range of CVD outcomes in female breast cancer survivors aged 66 years and over in the US.

METHODS

Study design and data source

We conducted a cohort study using the Surveillance, Epidemiology, and End Results program (SEER)–Medicare database, a linkage of cancer registry and Medicare enrolment and claims data. This linked database includes cancer cases through 2013 and Medicare claims through 2014. Medicare Part A and B claims provide information on diagnoses and procedures in the hospital and outpatient setting and Part D claims provide information on prescription drug dispensing. SEER data along with Medicare parts A, B, and D claims are used in this study from 2008 (which is the year Medicare Part D data came available to researchers). The data cover 12 states, which are covered by the SEER registry. These states equate to approximately 35% of the US population and have been shown to be highly representative of the wider US population in terms of poverty and education.[21]

Study population

We identified all women with Medicare Parts A, B and D enrolment and no managed care coverage for the 12-months before the month of cancer diagnosis. We included all women aged 66 years and over (Medicare coverage starts at age 65 years), with an incident ER+/PR+ and stage 1-3 breast cancer diagnosed between 1st January 2008 and 31st December 2013. Women were excluded if they had an endocrine therapy prescription prior to their breast cancer diagnosis. Follow up began one year after the date of breast cancer (hereafter the index date). Women were excluded if prior to their index date they: died, discontinued from Medicare Parts A, B, or D, were diagnosed with any cancer relating to sites other than the breast (excluding non-melanoma skin cancer), or were diagnosed with the CVD event of interest (within a 3-year look back period).

Exposures

Incident tamoxifen and AI exposures were identified using claims with the Healthcare Common Procedural Coding System (HCPCS) procedure codes and National Drug Codes (NDCs) included in Appendix 6.1. To help elucidate how drug exposure is associated with CVD risk, exposure was parameterised in two ways. First, we considered ever exposure to endocrine therapy (categorised as unexposed, ever use of tamoxifen, ever use of AI, or ever use of both drugs). If a woman moved between tamoxifen and AI exposure, their records were time-updated to indicate they had ever been exposed to both drugs from this point forward. They could also begin follow up as unexposed

and later move to any of the other exposure categories. A woman's exposure when they began follow up was the relevant categorisation at time of index date (Appendix 6.2). Second, we looked at current exposure to endocrine therapy (categorised as unexposed, current tamoxifen use, current AI use, no current therapy and any prior AI use, no current therapy and prior tamoxifen use only). Current exposure status was time-updated at any change in therapy (Appendix 6.3). An exposure was defined as continuous if a further claim for the same endocrine therapy followed within 30 days (i.e., the grace period) of the end of the prescription dispensing date plus the days supplied.

Outcomes

The main CVD outcomes of interest were: coronary artery disease (angina, myocardial infarction (MI), revascularisation procedures, sudden cardiac arrest (SCA)); peripheral vascular disease (PVD); stroke; arrhythmia; heart failure (HF, including cardiomyopathy), pericarditis; valvular heart disease (VHD); and venous thromboembolism (VTE) (deep vein thrombosis (DVT), pulmonary embolism (PE)). Composite CVD outcomes and all individual components of the composite outcomes were analysed separately. Events were identified using International Classification of Disease, 9th edition (ICD-9) diagnosis codes and HCPCS codes outlined in Appendix 6.4.

Covariates

Data on the following covariates were extracted for use in the analysis: year of breast cancer diagnosis (2008-2013); age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current calendar year. The following covariates were defined using ICD-9 diagnosis and procedure codes, NDC codes, and HCPCS codes based on a 3-year look back period prior to the index date: use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD. All codes used to identify prescribed drugs are outlined in Appendix 6.5. Diagnoses of comorbid conditions were based on the Klabunde adaptation of the Charlson comorbidity index,[22] which searches across the patients inpatient and outpatient claims to retain diagnosis codes that are either in the inpatient setting, or on two outpatient visits separated by over 30 days. All

comorbidities were adjusted for individually. Covariate selection was made through discussions with oncologists and cardiologists to understand the variables that are the most important potential confounders.

Statistical analysis

Observation time began at the index date and ended at the earliest of the following: a CVD event of interest, diagnosis of another non-breast cancer event, death, end of enrolment in Medicare Parts A, B, or D, or end of study period (December 13, 2014). Prior to exploring the relationship between endocrine therapies and CVD, distributions of baseline characteristics of patients who were unexposed at one year after diagnosis, or initially exposed to tamoxifen or AIs were described.

Primary analyses

Number of events and crude incident rates of each outcome of interest were calculated for all parameterizations of the primary exposures. The primary exposure variables were then included in unadjusted (which was adjusted for age due to using age as a timescale, but referred to as unadjusted throughout) and adjusted (accounting for all potential confounders) Cox proportional hazards regression models with an underlying age timescale, to obtain cause specific hazard ratios (HR). By changing the reference category, we calculated HRs both for AI and tamoxifen compared with the unexposed group, and for the direct comparison between AI and tamoxifen. Women with missing data for any covariate (5.1% overall) were excluded (complete case analysis), which is valid in a regression context if missingness is conditionally independent of the outcome.[23]

Secondary Analyses

Effect measure modification

For all outcomes, ever exposure analyses were tested for evidence of effect modification by any CVD (except for the event of interest) prior to index date; current age (66-74, 75-84, 85+; which was time-updated); and time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs, which also implicitly tests the proportional hazards assumption). Effect modification was tested by including an interaction term between the primary exposure and the potential effect modifier in the fully adjusted models, which was then tested using a Wald test. Results were not presented if there were no events within any categorisation of a stratified analyses, and 3 to <5yrs and 5+yrs were combined in the time since

index analysis if there were no events in the 5+yrs category.

Sensitivity analyses

A secondary study population was created, which was a subset of the original population, including only women prescribed either tamoxifen or AIs to address residual confounding by reasons for not initiating any endocrine therapy. Follow up began one year after the date of breast cancer, or at first prescription of tamoxifen or AI (whichever occurred latest). This population was then used to build an adjusted Cox proportional hazards model with an underlying age timescale, to obtain HRs for the association between the ever exposure variable outlined above and all CVD outcomes. Results were compared with those of the original ever exposure model including the unexposed group.

In case of misclassification of exposure status due to delays in patients obtaining their prescriptions, the grace period used to define a continuous prescription was extended from 30 days to 3 months, 6 months, and 1 year in the current prescription analyses.

To ensure results were not driven by additional unmeasured residual confounding, which is more likely in older women included in the study as they have a higher likelihood of being pre-disposed to CVD events given they have survived longer, the ever exposure analyses were repeated, excluding all women over the age of 85.

RESULTS

A total of 22027 women aged 66 and over were diagnosed with ER+/PR+ breast cancer during the study period. A STROBE flow diagram is provided in Appendix 6.6. At the index date, 4667 (22%) women were unexposed to any endocrine therapy, whereas 2286 (10%) and 15074 (68%) were initially exposed to tamoxifen and AIs respectively. The characteristics of women initially in each exposure group are shown in Table 1. Women initially exposed to AIs were generally younger, with later stages of breast cancer, and more systemic cancer treatment compared with women that were initially unexposed or exposed to tamoxifen. There was also an increase in the proportion of women initiating AIs over the study period (of all initiators, 83% were AI in 2008 and 89% were AI in 2013), along with a decrease in the proportion of women initiating tamoxifen (of all initiators, 17% were

tamoxifen in 2008 and 11% were tamoxifen in 2013). There was generally a high proportion of women that were previously diagnosed with non-venous CVD (60% overall), with more prior disease in those unexposed at index (64%) in comparison with those exposed to either tamoxifen or AIs (56% and 59% respectively).

Primary analyses

Ever exposure analyses

Mean follow up per person ranged from 2.26 years in the arrhythmia analysis to 2.50 years in the pulmonary embolism analysis. After adjustment for potential confounders, there was evidence of a decreased risk of MI associated with ever exposure to either tamoxifen or AIs compared with those unexposed to any endocrine therapy (ever tamoxifen vs unexposed adjusted HR: 0.44, 95% CI: 0.30-0.63; ever AI vs unexposed adjusted HR: 0.79, 95% CI: 0.64-0.97, Figure 1, Appendix 6.7), with a larger decreased risk in tamoxifen users (AI vs tamoxifen adjusted HR: 1.81, 95% CI: 1.28-2.58, Appendix 6.7). A similar pattern of a decreased risk in those exposed to tamoxifen was seen across all outcomes that made up the coronary artery disease outcome, albeit with 95% CIs that crossed unity for some outcomes. However, the evidence of a decreased risk in those ever exposed to AIs was not seen in all coronary artery disease outcomes (angina and revascularisation). Across the other non-venous CVD outcomes, there was generally evidence of a decreased risk associated with ever exposure to tamoxifen in comparison with the unexposed group, with varying strength and some 95% CIs that crossed unity. However, there was mixed evidence as to the relationship between ever exposure to AI and the other non-venous CVD outcomes, with some outcomes showing evidence of a decreased risk (stroke, arrhythmia, HF, and pericarditis), and others showing no evidence of effect (PVD, VHD). Within the outcomes that made up the composite VTE outcome, there was weak evidence of an increased risk of DVT associated with ever exposure to tamoxifen compared with those unexposed (adjusted HR: 1.42, 95% CI: 0.98-2.04), and no evidence of effect in those ever exposed to AIs compared with those unexposed (adjusted HR: 1.14, 95% CI: 0.86-1.52). Few PE events led to imprecise estimation of associations for this outcome.

Current exposure analyses

Within the coronary artery disease composite outcome, there was evidence of a decreased risk of both MI and SCA associated with being currently exposed to either tamoxifen or AIs compared with

being unexposed, with the lowest risk consistently in those currently exposed to tamoxifen. The decreased risk of MI persisted when women stopped tamoxifen, but there was no evidence of a difference in risk of MI with past exposure to AIs, and also no difference in risk of SCA with past exposure to either tamoxifen or AIs. Within the other non-venous CVD outcomes, there was consistent evidence of a decreased risk associated with current exposure to both tamoxifen and AIs compared with being unexposed (with less evidence of effect for the association with current exposure to AIs and the outcomes PVD and VHD), with the lowest risk being with exposure to tamoxifen for arrhythmia and HF. However, once exposure stopped, the risk of an outcome typically attenuated towards the null (except for the risk of stroke once stopping tamoxifen). There was evidence of an increased risk of DVT during exposure to tamoxifen compared with being unexposed (adjusted HR: 1.74, 95% CI: 1.23-2.46, Figure 2, Appendix 6.8), which attenuated towards the null among those who had stopped therapy (adjusted HR: 1.31, 95% CI: 0.93-1.84). But there was no evidence of a difference in the risk of DVT with current exposure to AIs compared with being unexposed (adjusted HR: 1.11, 95% CI: 0.83-1.49).

Secondary analyses

Effect modification

Effect modification analyses are presented in Appendices 6.9-6.21. There were generally few events within the stratified analyses for the majority of outcomes, meaning little statistical power to estimate effect modification. No results were presented for the PE outcome, as there were groups with no events for each stratified analysis. There was, however, evidence of effect modification by both time since index date and age within the HF analysis (p for interaction=0.04 and 0.07 respectively). Although there was weak evidence of a decreased risk of HF associated with exposure to both tamoxifen and AIs in the time leading up to 5 years after index, the direction of association changed after this point, albeit with 95% CIs that crossed unity. There was also evidence of effect modification by current age within the VHD outcome (p for interaction=0.04), with a decreasing risk of VHD as age increased.

Sensitivity analyses

Direct comparison of AI vs tamoxifen

There was evidence of an increased risk of angina, MI, arrhythmia, pericarditis, and VHD associated with ever exposure to AI compared with tamoxifen in the head-to-head comparison (Appendix 6.22).

Results were similar in the primary analyses when the reference category was ever tamoxifen use and in the sensitivity analysis restricted to those only given endocrine therapy.

Differing grace periods

Evidence of all effects within all current exposure and outcome analyses persisted when the grace period was extended to 3 months, 6 months, and one year (Appendix 6.23).

Exclusion of women over the age of 85 years

The direction and trend of associations for all outcomes remained the same in the ever exposure analysis when women over the age of 85 years were excluded (Appendix 6.24). However, associations generally moved towards the null for all outcomes within the coronary artery disease and VTE composite outcomes (angina, MI, revascularisation, SCA, DVT, and PE), and away from the null for other non-venous CVD outcomes (stroke, arrhythmia, HF, and pericarditis).

DISCUSSION

Main Findings

In this large, cancer registry-based study in women aged 66 years and over diagnosed with ER+/PR+ breast cancer in the US, there was no evidence of increased non-venous CVD risk associated with either tamoxifen or AI use compared with no endocrine therapy. However, the risk of all non-venous CVD outcomes was lower in those ever exposed to tamoxifen compared with those unexposed, and for six out of 11 of the non-venous outcomes (including the composite coronary outcome), there was good statistical evidence of a protective effect with 95% CIs not including the null association. The largest decreased risk was reported in the MI analysis, with those ever exposed to tamoxifen being at a 56% (95% CI: 37-70%) lower risk of an outcome in comparison with those unexposed. There was mixed evidence on the relationship between exposure to AIs and risk of non-venous CVD outcomes, with evidence of a decreased risk of some outcomes with exposure to AI (MI, stroke, HF, and pericarditis), and no evidence of a difference, or 95% CIs that crossed unity, for other outcomes. However, women exposed to tamoxifen were consistently at the lowest risk of all non-venous CVD outcomes (except for HF, where all endocrine therapy exposed groups had a similar reduced risk in comparison with those unexposed). For those outcomes in which there was a protective effect during exposure to either tamoxifen or AI, the risk of non-venous CVDs typically increased towards

the level of those unexposed once exposure ended. There was also evidence of an increased risk of DVT during time currently exposed to tamoxifen, which decreased when women stopped tamoxifen.

There is biological rationale that tamoxifen possesses cardio-protective effects due to decreasing lipid levels, with evidence of reductions in total serum cholesterol (between 10% and 15%) and low-density lipoprotein cholesterol (between 15% and 22%).[24 25] It has also been suggested that AIs could increase the risk of CVD outcomes in comparison with tamoxifen as they reduce oestrogen levels and therefore the oestrogen-mediated protective CVD effects such as regulation of serum lipid metabolism, increasing vasodilation, and inhibition of the development of atherosclerosis,[26] but several RCTs have compared hypercholesterolemia between AI and tamoxifen users, with inconclusive results.[27-29]

Comparison with other studies

This is the first study to our knowledge to assess the effect of endocrine therapies in comparison with an unexposed breast cancer population on a full range of clinically relevant CVD outcomes. Seven studies (five observational and two RCTs) directly compared tamoxifen use to either no tamoxifen or placebo and the risk of MI, [13-17 30 31] with four studies reporting relative risks in a similar protective direction as reported in our study (RRs ranged from 0.20-0.83).[13 14 30 31] Similar to our study, one observational study and one RCT have reported protective effects of AI use versus no use associated with the risk of MI, albeit with less precise estimates in which the 95% CI crossed unity (RR: 0.90, 95% CI: 0.65-1.25, RR: 0.82, 95% CI: 0.55-1.22 respectively).[16 29] Three out of five studies (four observational and one RCT) exploring the effect of tamoxifen use on risk of stroke also reported a similar protective direction of effect (RRs ranged from 0.52-0.81),[14-16 19 30] but two studies reported effects in opposite directions for the effect of AI use on the risk of stroke,[16 29] whereas we reported a protective effect. Our results are also consistent with the established increased risk of VTE outcomes associated with tamoxifen use, [30 32-35] but we did not find a similar increased risk in AI users reported in one previous RCT.[29] All previous observational and RCT evidence was collated in a recent systematic review.[20] Similar to this study, it concluded that AI users are at a higher risk of several vascular diseases including MI and angina in comparison with tamoxifen users, which may be partly driven by a protective effect of tamoxifen, and tamoxifen users are at a higher risk of VTE outcomes in comparison with both AI users and those unexposed. However, our results also suggest that AI use may have cardio-protective effects versus no endocrine therapy use for some specific

CVD outcomes (MI, SCA, stroke, angina, HF, arrhythmia, pericarditis), but to a lesser extent than tamoxifen.

Strengths and limitations

SEER-Medicare includes a large, diverse population of older women diagnosed with breast cancer and treated in real world settings, in which we were able to account for several potentially cardio-toxic treatments such as anthracyclines and trastuzumab. The inclusion of those with Medicare Part D coverage also allowed adjustment for CVD-related prescription medications. These results likely to be generalisable to women aged 66 and over diagnosed with ER+/PR+ breast cancer in both the US and other developed countries due to the homogenous indication of endocrine therapy worldwide.

Several limitations of the study could be non-causal explanations of the observed associations. The inclusion of only women with ER+/PR+ breast cancer meant we were able to compare CVD risk in those with a clinically similar breast cancer diagnosis. The proportion of women that were unexposed to any therapy at index date in this study (21%) was similar to the proportion of non-initiators reported in a similar population of women diagnosed with ER+/PR+ breast cancer by Farias et al (25%).[36] However, the non-initiators may have an elevated CVD risk at baseline in comparison with women that initiated endocrine therapy due to being generally older, and having more comorbidities such as CKD, and a higher proportion of previous CVD (Table 1). Although, when assessing effect modification by prior CVD for most outcomes in which there was evidence of an effect of tamoxifen use, the protective association mostly persisted in those without previous CVD or there was no evidence of effect modification, which argues against residual confounding by prior CVD. Other reasons for non-initiation could include frailty, poor CVD preventative care, and high BMI, which are also risk factors for CVD. The very large protective effect of tamoxifen use on risk of MI (HR: 0.44, 95% CI: 0.30-0.63), which is beyond even the most established CVD drugs such as statins, or the decreased risk of several outcomes associated with AI use, could therefore be partly explained by unmeasured confounding of these factors, which could not be directly measured in these data. Although it is unlikely that this residual confounding explains all of the observed association for the protective effects of tamoxifen on several non-venous CVD outcomes, it is possible that some of the observed associations with AI use are non-causal.

Within unexposed patients, no information was available on the other reason for non-initiation. This

could be due to patient's individual factors (e.g., beliefs about therapy effectiveness, financial barriers, access to care), physician recommendations, or potential medical contraindications. We were also not able to account for contraindications such as hot flushes and night sweats. The model using the restricted study population without non-initiators (Appendix 6.22) had less potential for confounding due to factors related to reasons for non-initiation,[37] and reported similar results to the original model when directly comparing the difference between ever AI and tamoxifen use.

SEER-Medicare data has no information on patients' lifestyle measures such as smoking, BMI, and alcohol use, which are overlapping risk factors of both breast cancer and CVD. Any differences in these factors between initiators and non-initiators of endocrine therapy will result in residual confounding. However, although there is evidence to suggest that smoking and alcohol use influence adherence to endocrine therapy,[38] there is currently no evidence that any of these factors affect choice of endocrine therapy.

In the interest of minimising the number of exposure arms in the current exposure analysis, all women previously exposed to an AI contributed time to the 'past with AI' group, regardless of if they were also previously exposed to tamoxifen. Within all women included in the 'past with AI group', 10% of the time exposed to any endocrine therapy was exposed to tamoxifen. This could help explain why there are slightly greater protective and toxic effects in the 'past with AI group' in comparison with the current AI group in the angina and DVT current exposure analyses respectively.

CONCLUSION

This large cancer registry-based cohort study in the US is the first to explore the relative effect of tamoxifen and AI use compared with an unexposed population on the risk of a range of clinically specific CVD outcomes in ER+/PR+ older breast cancer survivors. It has been widely suggested that AIs users are at increased risk of non-venous CVDs in comparison with tamoxifen users. Results suggest evidence of a protective effect of tamoxifen therapy on the risk of the non-venous CVDs: MI, SCA, stroke, arrhythmia, HF, pericarditis, and VHD. However, we also report evidence of a decreased risk of several non-venous CVDs during AI use, albeit less than with tamoxifen use, but there is a possibility of residual confounding due to a potentially increased risk of non-venous CVDs in non-initiators of endocrine therapy at baseline, that brings questions to the causality of these associations. The established increased risk of DVT with tamoxifen use is also confirmed. While

choice of AI or tamoxifen will primarily be based on the effectiveness against breast cancer recurrence, the individual's risk of both venous and non-venous CVDs are important secondary considerations, and these results will thus help inform clinical decision-making in women with ER+/PR+ breast cancer.

CONFLICTS OF INTEREST

AM, SPH, and JL have nothing to disclose. SS reports personal fees from Roche, Clinigen, Eli Lilly, and Novartis, outside the submitted work. AL reports personal fees from Servier, Novartis, Pfizer, Roche, Ferring Pharmaceuticals, Clinigen Group, Boehringer Ingelheim, Amgen, Eli Lily, and BMS, outside the submitted work. LS reports grants from Wellcome, during the conduct of the study; grants from Wellcome, MRC, NIHR, BHF, Diabetes UK, and grants and personal fees from GSK, outside the submitted work; and Is a trustee of the British Heart Foundation. KB reports grants from Wellcome Trust and the Royal Society, during the conduct of the study.

CONTRIBUTIONS

Study design was decided on by AM, JL, and KB. SPH carried out data extraction. AM carried out all manipulation and analyses. AM wrote the first draft. All authors contributed to further drafts and approved the final manuscript.

FUNDING

This study was supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (grant No 107731/Z/15/Z) held by KB. The Wellcome Trust and the Royal Society had no role in the design, analysis, or writing up of this study.

REFERENCES

1. Ferlay JS-F, E. Lortet-Tieulent, J. Rosso, S. Coebergh, JWW. Comber, H. Forman, D. Bray, F. . Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *European journal of cancer* 2012;6:1374-403.

2. Cancer Research UK. Breast cancer survival statistics http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/survival#heading-Zero: Cancer Research UK; 2014 [accessed 01/10 2018.

3. American Cancer Society. Cancer Facts & Figures 2018. In: Society AC, ed. Atlanta, 2018.

4. Cuzick J, Sestak I, Baum M, et al. Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 10-year analysis of the ATAC trial. *The Lancet Oncology* 2010;11(12):1135-41. doi: 10.1016/S1470-2045(10)70257-6

5. Regan MM, Neven P, Giobbie-Hurder A, et al. Assessment of letrozole and tamoxifen alone and in sequence for postmenopausal women with steroid hormone receptor-positive breast cancer: the BIG 1-98 randomised clinical trial at 8.1 years median follow-up. *The Lancet Oncology* 2011;12(12):1101-8. doi: 10.1016/S1470-2045(11)70270-4

6. Dubsky PC, Jakesz R, Mlineritsch B, et al. Tamoxifen and anastrozole as a sequencing strategy: a randomized controlled trial in postmenopausal patients with endocrine-responsive early breast cancer from the Austrian Breast and Colorectal Cancer Study Group. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2012;30(7):722-8. doi: 10.1200/JCO.2011.36.8993

7. Boccardo F, Rubagotti A, Guglielmini P, et al. Switching to anastrozole versus continued tamoxifen treatment of early breast cancer. Updated results of the Italian tamoxifen anastrozole (ITA) trial. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 2006;17 Suppl 7:vii10-4. doi: 10.1093/annonc/mdl941

8. Kaufmann M, Jonat W, Hilfrich J, et al. Improved overall survival in postmenopausal women with early breast cancer after anastrozole initiated after treatment with tamoxifen compared with continued tamoxifen: the ARNO 95 Study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2007;25(19):2664-70. doi: 10.1200/JCO.2006.08.8054

9. van de Velde CJ, Rea D, Seynaeve C, et al. Adjuvant tamoxifen and exemestane in early breast cancer (TEAM): a randomised phase 3 trial. *Lancet* 2011;377(9762):321-31. doi: 10.1016/S0140-6736(10)62312-4

10. Ingle JN, Tu D, Pater JL, et al. Intent-to-treat analysis of the placebo-controlled trial of letrozole for extended adjuvant therapy in early breast cancer: NCIC CTG MA.17. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 2008;19(5):877-82. doi: 10.1093/annonc/mdm566

11. Bliss JM, Kilburn LS, Coleman RE, et al. Disease-related outcomes with long-term follow-up: an updated analysis of the intergroup exemestane study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2012;30(7):709-17. doi: 10.1200/JCO.2010.33.7899

12. Khosrow-Khavar F, Filion KB, Al-Qurashi S, et al. Cardiotoxicity of aromatase inhibitors and tamoxifen in postmenopausal women with breast cancer: a systematic review and meta-analysis of randomized controlled trials. *Ann Oncol* 2017;28(3):487-96. doi: 10.1093/annonc/mdw673

13. Bradbury BD, Lash TL, Kaye JA, et al. Tamoxifen-treated breast carcinoma patients and the risk of acute myocardial infarction and newly-diagnosed angina. *Cancer* 2005;103(6):1114-21.

14. Yang TL, Wu TC, Huang CC, et al. Association of tamoxifen use and reduced cardiovascular events among asian females with breast cancer. *Circulation Journal* 2014;78(1):135-40.

15. Hernandez RK, Sorensen HT, Jacobsen J, et al. Tamoxifen treatment in Danish breast cancer patients and 5-year risk of arterial atherosclerotic events: a null association. *Cancer Epidemiology, Biomarkers & Prevention* 2008;17(9):2509-11. doi: http://dx.doi.org/10.1158/1055-9965.EPI-08-0570

16. Ligibel JA, James O'Malley A, Fisher M, et al. Risk of myocardial infarction, stroke, and fracture in a cohort of community-based breast cancer patients. *Breast Cancer Research & Treatment* 2012;131(2):589-97. doi: http://dx.doi.org/10.1007/s10549-011-1754-1

17. Geiger AM, Chen W, Bernstein L. Myocardial infarction risk and tamoxifen therapy for breast cancer. *British journal of cancer* 2005;92(9):1614-20.

18. Haque R, Shi J, Schottinger JE, et al. Cardiovascular Disease After Aromatase Inhibitor Use. *JAMA oncology* 2016 doi: 10.1001/jamaoncol.2016.0429

19. Geiger AM, Fischberg GM, Chen W, et al. Stroke risk and tamoxifen therapy for breast cancer. *Journal of the National Cancer Institute* 2004;96(20):1528-36.

20. Matthews A, Stanway S, Farmer RE, et al. Long term adjuvant endocrine therapy and risk of cardiovascular disease in female breast cancer survivors: systematic review. *Bmj* 2018;363:k3845. doi: 10.1136/bmj.k3845

21. Nattinger AB, McAuliffe TL, Schapira MM. Generalizability of the surveillance, epidemiology, and end results registry population: factors relevant to epidemiologic and health care research. *J Clin Epidemiol* 1997;50(8):939-45. [published Online First: 1997/08/01]

22. Klabunde CN, Potosky AL, Legler JM, et al. Development of a comorbidity index using physician claims data. *J Clin Epidemiol* 2000;53(12):1258-67.

23. White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. *Stat Med* 2010;29(28):2920-31. doi: 10.1002/sim.3944

24. Dewar JA, Horobin JM, Preece PE, et al. Long term effects of tamoxifen on blood lipid values in breast cancer. *Bmj* 1992;305(6847):225-6.

25. Love RR, Wiebe DA, Feyzi JM, et al. Effects of tamoxifen on cardiovascular risk factors in postmenopausal women after 5 years of treatment. *Journal of the National Cancer Institute* 1994;86(20):1534-9.

26. Mendelsohn ME, Karas RH. The protective effects of estrogen on the cardiovascular system. *N Engl J Med* 1999;340(23):1801-11. doi: 10.1056/NEJM199906103402306

27. Arimidex TAoiCTG, Buzdar A, Howell A, et al. Comprehensive side-effect profile of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: long-term safety analysis of the ATAC trial. *Lancet Oncol* 2006;7(8):633-43. doi: 10.1016/S1470-2045(06)70767-7

28. Coates AS, Keshaviah A, Thurlimann B, et al. Five years of letrozole compared with tamoxifen as initial adjuvant therapy for postmenopausal women with endocrine-responsive early breast cancer: update of study BIG 1-98. *Journal of Clinical Oncology* 2007;25(5):486-92.

29. Goss PE, Ingle JN, Martino S, et al. Randomized trial of letrozole following tamoxifen as extended adjuvant therapy in receptor-positive breast cancer: updated findings from NCIC CTG MA.17. *Journal of the National Cancer Institute* 2005;97(17):1262-71.

30. McDonald CC, Alexander FE, Whyte BW, et al. Cardiac and vascular morbidity in women receiving adjuvant tamoxifen for breast cancer in a randomised trial. The Scottish Cancer Trials Breast Group. *Bmj* 1995;311(7011):977-80.

31. Rutqvist LE. Long-term toxicity of tamoxifen. *Recent Results Cancer Res* 1993;127:257-66.

32. Fisher B, Anderson S, Tan-Chiu E, et al. Tamoxifen and chemotherapy for axillary node-negative, estrogen receptor-negative breast cancer: findings from National Surgical Adjuvant Breast and Bowel Project B-23. *J Clin Oncol* 2001;19(4):931-42. doi: 10.1200/JCO.2001.19.4.931

33. Fisher B, Dignam J, Wolmark N, et al. Tamoxifen in treatment of intraductal breast cancer: National Surgical Adjuvant Breast and Bowel Project B-24 randomised controlled trial. *Lancet* 1999;353(9169):1993-2000. doi: 10.1016/S0140-6736(99)05036-9

34. Hernandez RK, Sorensen HT, Pedersen L, et al. Tamoxifen treatment and risk of deep venous thrombosis and pulmonary embolism: a Danish population-based cohort study. *Cancer* 2009;115(19):4442-9. doi: http://dx.doi.org/10.1002/cncr.24508

35. Meier CR, Jick H. Tamoxifen and risk of idiopathic venous thromboembolism. *British journal of clinical pharmacology* 1998;45(6):608-12.

36. Farias AJ, Du XL. Ethnic differences in initiation and timing of adjuvant endocrine therapy among older women with hormone receptor-positive breast cancer enrolled in Medicare Part D. *Med Oncol* 2016;33(2):19. doi: 10.1007/s12032-016-0732-1

37. Lund JL, Richardson DB, Sturmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. *Curr Epidemiol Rep* 2015;2(4):221-28. doi: 10.1007/s40471-015-0053-5

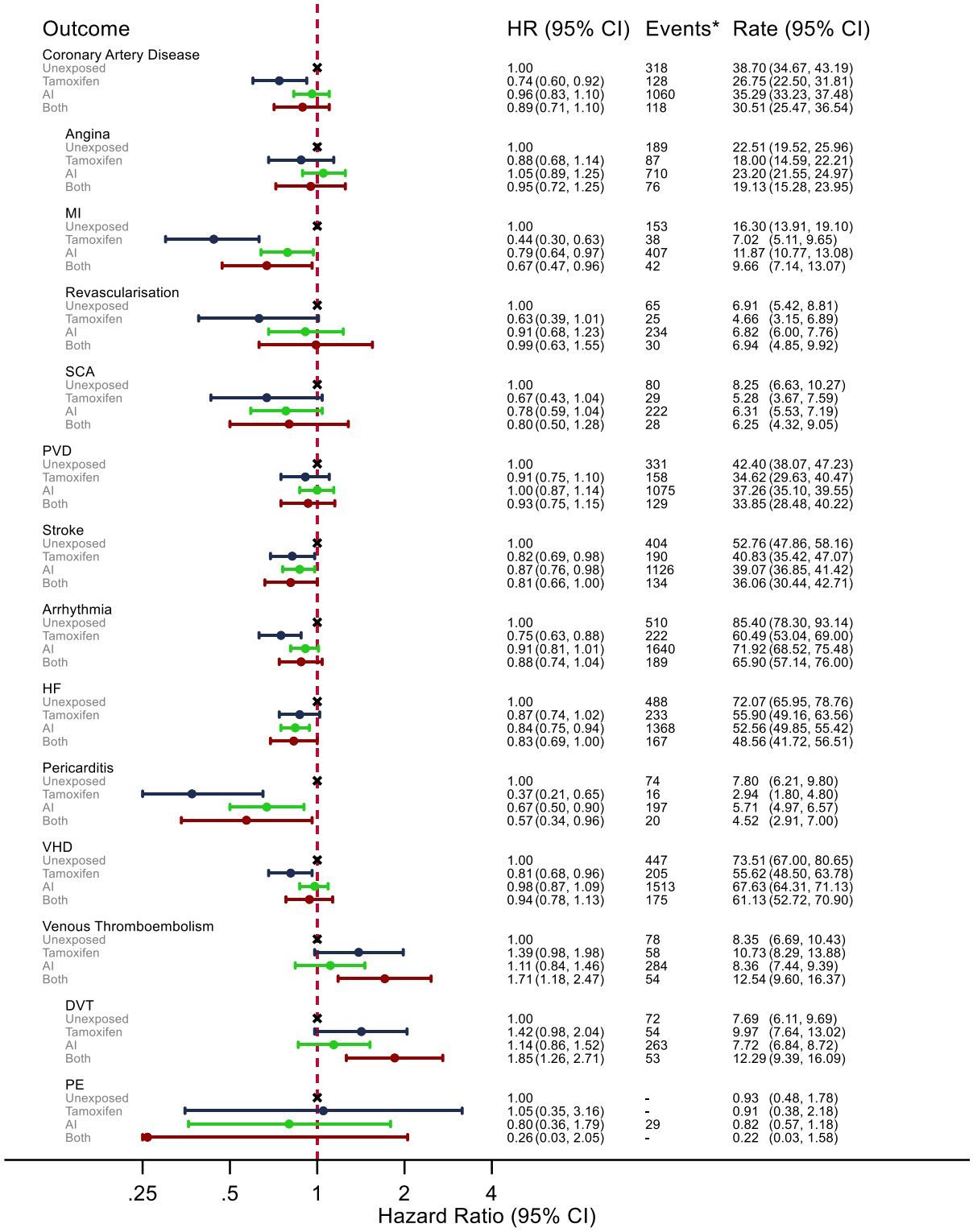
38. Land SR, Cronin WM, Wickerham DL, et al. Cigarette smoking, obesity, physical activity, and alcohol use as predictors of chemoprevention adherence in the National Surgical Adjuvant Breast and Bowel Project P-1 Breast Cancer Prevention Trial. *Cancer prevention research* 2011;4(9):1393-400. doi: 10.1158/1940-6207.CAPR-11-0172

Table 1: Characteristics of study population based on their initial exposure

	Unexposed (%)	Tamoxifen (%)	AI (%)	Total (%)
N	4667 (100)	2286 (100)	15074	22027 (100)
Age at index date (yrs)				
66-74	1538 (33)	897 (39.2)	7505	9940 (45.1)
75-84	1937 (41.5)	994 (43.5)	5894	8825 (40.1)
85+	1192 (25.5)	395 (17.3)	1675	3262 (14.8)
Median (IQR)	79 (73-85)	77 (72-83)	75 (71-	76 (71-82)
Ethnicity				
White	4002 (85.8)	2028 (88.7)	12782	18812 (85.4)
Black	360 (7.7)	102 (4.5)	1099	1561 (7.1)
Other	93 (2)	47 (2.1)	335 (2.2)	475 (2.2)
Asian	123 (2.6)	68 (3)	498 (3.3)	689 (3.1)
Hispanic	71 (1.5)	32 (1.4)	294 (2)	397 (1.8)
Native American	-	-	-	52 (.2)
Missing	-	-	-	41 (.2)
SEER Region				
North East	760 (16.3)	283 (12.4)	3320	4363 (19.8)
South	1023 (21.9)	605 (26.5)	3778	5406 (24.5)
North Central	695 (14.9)	448 (19.6)	1761	2904 (13.2)
West	2157 (46.2)	939 (41.1)	6130	9226 (41.9)
Missing	32 (.7)	11 (.5)	85 (.6)	128 (.6)
Year of breast cancer diagnosis				
2008	847 (18.1)	414 (18.1)	2075	3336 (15.1)
2009	831 (17.8)	449 (19.6)	2140	3420 (15.5)
2010	729 (15.6)	368 (16.1)	2394	3491 (15.8)
2011	744 (15.9)	333 (14.6)	2625	3702 (16.8)
2012	756 (16.2)	350 (15.3)	2709	3815 (17.3)
2013	760 (16.3)	372 (16.3)	3131	4263 (19.4)
Stage of breast cancer				
Stage I	3034 (65)	1486 (65)	8379	12899 (58.6)
Stage II	1275 (27.3)	660 (28.9)	5267	7202 (32.7)
Stage III	358 (7.7)	140 (6.1)	1428	1926 (8.7)
Grade of breast cancer				
1	1522 (32.6)	765 (33.5)	4273	6560 (29.8)
2	2071 (44.4)	1109 (48.5)	7350	10530 (47.8)
3	853 (18.3)	324 (14.2)	2810	3987 (18.1)
Missing	221 (4.7)	88 (3.8)	641 (4.3)	950 (4.3)
Cancer treatments				
Taxane	570 (12.2)	162 (7.1)	2415	3147 (14.3)
Anthracyclines	259 (5.5)	68 (3)	820 (5.4)	1147 (5.2)
Trastuzumab	226 (4.8)	39 (1.7)	687 (4.6)	952 (4.3)
Other treatment	753 (16.1)	244 (10.7)	2992	3989 (18.1)
Comorbidities				
RA	185 (4)	103 (4.5)	547 (3.6)	835 (3.8)
CKD	383 (8.2)	155 (6.8)	1113	1651 (7.5)
Hypertension	3426 (73.4)	1612 (70.5)	11113	16151 (73.3)
Diabetes	1313 (28.1)	598 (26.2)	4545	6456 (29.3)
CVD related treatment				
Statins	1778 (38.1)	948 (41.5)	6988	9714 (44.1)
Hypertensives	169 (3.6)	83 (3.6)	577 (3.8)	829 (3.8)
ACEi	962 (20.6)	477 (20.9)	3251	4690 (21.3)
CCB	850 (18.2)	364 (15.9)	2696	3910 (17.8)
ARB	593 (12.7)	255 (11.2)	2063	2911 (13.2)
Past CVD				
Non venous CVD	2989 (64)	1281 (56)	8896	13166 (59.8)
VTE	162 (3.5)	31 (1.4)	385 (2.6)	578 (2.6)

*Cells with '–' represent those with numbers suppressed due to some cells containing numbers ≤11

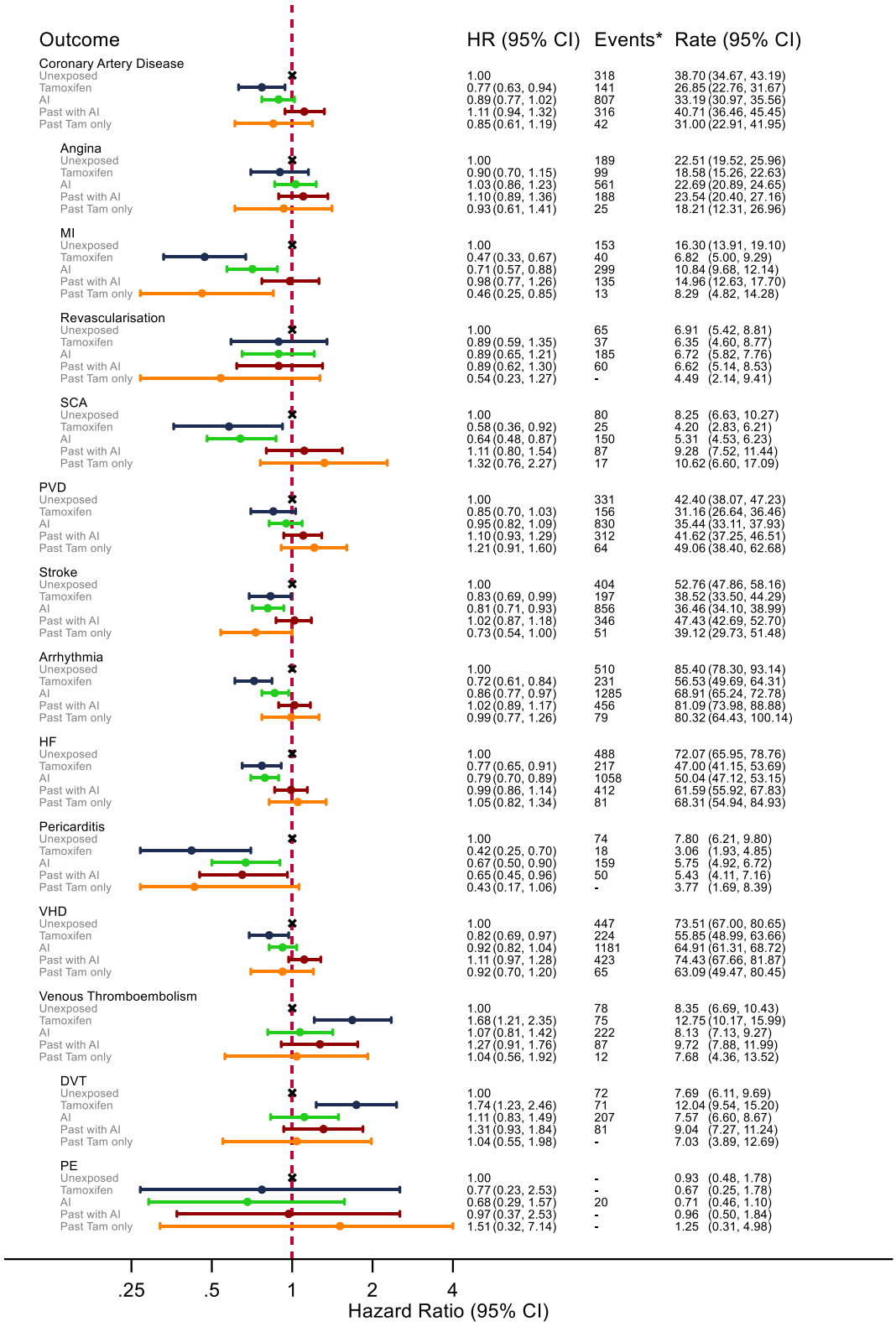
Figure 1: Adjusted HRs, events, and crude rate per 1000 person-years for the association between ever exposure to endocrine therapy and a range of clinical CVD outcomes



*Events suppressed if number of events ≤ 11

† HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

Figure 2: Adjusted HRs, events, and crude rate per 1000 person-years for the association between current exposure to endocrine therapy and a range of clinical CVD outcomes



*Events suppressed if number of events ≤ 11

† HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

7 COMPARISON OF UK AND US STUDIES

7.1 INTRODUCTION

The main aims of both the UK and US studies were to assess the cardiotoxicity of endocrine therapies in breast cancer survivors. However, given the differences between the CPRD/HES and SEER-Medicare data, there were differences in the methods used to achieve this aim in the two studies. This chapter will compare the methodology of the original studies and their respective results, then re-analyse both studies to make them as similar as possible and allow for a direct comparison.

7.2 COMPARISON OF METHODOLOGY

This section will give an overview of the difference in study designs between the UK and US study, as well as explanations of why differences arose.

7.2.1 Study populations

Table 7.1 outlines the concordance of inclusion and exclusion criteria and start of follow up for both studies. In the UK study, all women aged 54 years and over with a breast cancer diagnosis between 2002 and 2013 who received a tamoxifen or AI prescription were included. Whereas, women aged 66 years and over with an ER+/PR+ breast cancer diagnosis between 2008 and 2013 were included in the US study, regardless of if they were prescribed an endocrine therapy. These differences were because SEER-Medicare follow up starts at the age of 65 years (then women needed a year of follow up prior to their breast cancer diagnosis), Medicare part D follow up is complete from 2008, and ER/PR status is not available in CPRD or HES. The only difference in exclusion criteria was that women were excluded from analysis if they had the CVD event of interest at any point prior to index date in the UK, but only within a 3-year look back period in the US. This was because Medicare follow up only starts at the age of 65 years, so the likelihood of identifying prior events would depend on age if there was a longer look back period because older women would have the most prior follow up available. If women were aged 66 or 67 years, their look back period was only 1 and 2 years respectively.

Table 7.1: Concordance of study populations between UK and US studies

	UK study	US study	Concordant	Reason
Inclusion criteria				
Gender	Female	Female	✓	
Age	54 years and over	66 years and over	X	Only data from those aged of 65 year and over available in SEER-Medicare
Cancer diagnosis	Breast cancer with one year of follow up prior to diagnosis	ER+/PR+ breast cancer with one year of follow up prior to diagnosis	X	ER/PR status not available in CPRD or HES
Date	Breast cancer diagnosis between 1/1/2002 and 31/3/2015	Breast cancer diagnosis between 1/1/2008 and 31/12/2013	X	Medicare Part D was complete from 2008 and was needed to identify prescriptions
Endocrine therapy prescription	Only women with a tamoxifen or AI prescription after their breast cancer diagnosis	All women with ER+/PR+ breast cancer, regardless of if they were prescribed tamoxifen, AIs, or neither. Prescriptions had to be after breast cancer diagnosis	X	As ER/PR status was not available in CPRD or HES, it was impossible to identify an unexposed population with a similar breast cancer diagnosis
Exclusion criteria				
Died or discontinued follow up in database	At any point prior to index date	At any point prior to index date	✓	
Diagnosed with another cancer	At any point prior to index date	At any point prior to index date	✓	
Diagnosed with CVD event of interest	At any point prior to index date	Within 3 years prior to index date	X	Women are only eligible for Medicare from the age of 65 years in the US, so including a longer look back period would mean the likelihood of identifying prior events would depend on age because older women would have the most prior follow up available
Start of follow up	Latest of 1 year after breast cancer diagnosis or first prescription	1 year after breast cancer diagnosis	X	Under the inclusion criteria in the UK study, women needed a prescription enter study, so couldn't enter prior to this

7.2.2 Exposure

To define recorded prescriptions, CPRD uses its own product code system and SEER-Medicare uses HCPCS and NDCs. It is not possible to map these systems to each other, so in each study,

independent code lists were created that were as similar as possible.

The ever and current exposure to endocrine therapy variables were created in the same way for both studies, but the US study included an unexposed group. The only difference was how length of prescription was calculated. CPRD supplies information on the number of drugs prescribed and the recommended daily dose, so it was possible to calculate the length of each prescription, whereas Medicare part D directly supplies the number of days for which the drug was supplied. Regardless of how length of prescription was calculated, a 30-day grace period from the end of the prescription defined a continuous prescription; if there was another prescription within this period, prescription continuity was assumed. Both studies also included sensitivity analyses that varied the grace periods.

7.2.3 Outcomes

The UK study used Read codes in CPRD and ICD-10 codes in HES to identify CVD diagnoses, whereas the US study used ICD-9 codes in SEER-Medicare. The ICD-10 codes used in the UK study were provided to UNC as part of the data specification, and were mapped to ICD-9 codes by a UNC student as part of the dataset preparation to ensure concordance between the outcome definitions in both studies. Both studies identified outcomes in both primary and secondary care.

7.2.4 Covariates

Table 7.2 shows the covariates adjusted for in both studies, along with an overview of their definitions. The following section includes further explanations for any differences in covariate definitions.

7.2.4.1 Demographics

IMD score was available in the UK study, but was not available in SEER-Medicare data. Race is available in the CPRD, however there is poor completeness and consistency of this variable, especially prior to 2006,[118] so the decision was taken not to include it. Furthermore, region was not considered an important confounder in the UK due to the smaller geographical differences in the UK.

Table 7.2: Covariates adjusted for in UK and US studies

Covariate	UK study		US study	
	Adjusted	**Definition**	**Adjusted**	**Definition**
Demographic				
Age at index	✓	54-59, 60-69, 70+	✓	66-74, 75-84, 85+
IMD score	✓	level 1-5 based on GP level IMD data	X	N/A
Race	X	N/A	✓	White, Black Asian, Hispanic, Native American, other
Region	X	N/A	✓	North East, South, North Central, West
Lifestyle measures				
Smoking status	✓	At index (non-smoker, current smoker, ex-smoker)	X	N/A
BMI	✓	At index (underweight/healthy weight, overweight, obese)	X	N/A
Alcohol status	✓	At index (non-drinker, current drinker, ex-drinker)	X	N/A
Treatments				
Use of statins	✓	Ever use prior to index	✓	Use within 3-years prior to index
Use of ACE inhibitors	✓	Ever use prior to index	✓	Use within 3-years prior to index
Use of CCBs	✓	Ever use prior to index	✓	Use within 3-years prior to index
Use of ARBs	✓	Ever use prior to index	✓	Use within 3-years prior to index
Use of antiplatelets	✓	Ever use prior to index	X	N/A
Use of anti-hypertensive drugs	X	N/A	✓	Use within 3-years prior to index
Use of taxanes	X	N/A	✓	Use within 3-years prior to index
Use of anthracyclines	X	N/A	✓	Use within 3-years prior to index
Use of trastuzumab	X	N/A	✓	Use within 3-years prior to index
Use of other systemic cancer therapies	X	N/A	✓	Use within 3-years prior to index
Comorbidity diagnoses				
Diabetes	✓	Ever diagnosed prior to index	✓	Diagnosed within 3-years prior to index
CKD	✓	Ever diagnosed prior to index	✓	Diagnosed within 3-years prior to index
Rheumatoid arthritis	✓	Ever diagnosed prior to index	✓	Diagnosed within 3-years prior to index
Systolic blood pressure	✓	At index (low/normal, pre-high, high)	X	N/A
Diastolic blood pressure	✓	At index (low/normal, pre-high, high)	X	N/A
Hypertension	X	N/A	✓	Diagnosed within 3-years prior to index
History of non-venous CVD	✓	Ever diagnosed prior to index	✓	Diagnosed within 3-years prior to index
History of any VTE outcome	✓	Ever diagnosed prior to index	✓	Diagnosed within 3-years prior to index
Information relating to breast cancer				
Year of breast cancer diagnosis	✓	2002-2015	✓	2007-2013
Breast cancer stage	X	N/A	✓	1-3
Breast cancer grade	X	N/A	✓	1-3
Other information				
Time since index date	✓	0-1yr, 1-3yrs, 3-5yrs, 5+yrs	✓	0-1yr, 1-3yrs, 3-5yrs, 5+yrs
Current year	✓	2002-2017	✓	2007-2014

7.2.4.2 Lifestyle measures

No lifestyle measures such as smoking status, BMI, and alcohol status were available in SEER-Medicare data.

7.2.4.3 Treatments

No reliable data on cancer treatments (other than endocrine therapy) were available in the CPRD or HES data. The US study adjusted for use of anti-hypertensive drugs, but the UK study did not, as the presence of hypertension was measured using systolic and diastolic blood pressure. Furthermore, the UK study adjusted for use of anti-platelets, but the US study did not. However, in the UK study, there was minimal difference between the crude HRs and the HRs after adjustment for antiplatelet use. The biggest difference was in the angina analysis, where the absolute difference between the crude and the adjusted HR was 7% (Appendix 7.1). However, the absolute difference between the crude and adjusted HRs was less than 5% for 12/14 of the outcomes. Antiplatelet use was therefore not considered an important confounder.

In the UK study, women were defined as previously being exposed to any of these drugs at index if they had a prescription at any point between their entry into the CPRD network and index date. Whereas in the US study, women were defined as being previously exposed at index if they had a prescription within the three years prior to index date (or between entry into the network and index if this period was less than three years). Like prior CVD events, this was because Medicare follow up only starts at the age of 65 years, so the likelihood of identifying prior prescriptions would depend on age if there was a longer look back period because older women would have the most prior follow up available.

7.2.4.4 Diagnoses

Systolic and diastolic blood pressure readings were not available in SEER-Medicare data, so a record of a hypertension diagnosis, as well as prescription of an anti-hypertensive drug (above) was used as a proxy for hypertension in the US study.

As with treatments above, women in the UK study were defined as previously having a diagnosis of a

comorbidity at index if they had a record at any point between their entry into the CPRD network and index date. Whereas in the US study, women were defined as previously having a diagnosis of a comorbidity at index if they had a record within the three years prior to index date (or between entry into the network and index if this period was less than three years).

7.2.4.5 Information relating to breast cancer

Breast cancer stage and grade was not available in CPRD or HES. It is possible to link the CPRD and HES data to the UK Cancer Registry, but there is a high level of missingness of stage and grade, which has the potential to impact statistical power and bias any results if missingness is dependent on the outcome (in a complete case analysis). It was therefore decided not to link these data.

7.2.4.6 Other information

Both studies adjusted for time since index date and current year.

7.2.5 Statistical Analysis

For both studies, observation time began at index date and ended at earliest of the following: a CVD event of interest, diagnosis of another cancer, death (from CPRD records in the UK study, and Medicare records in the US study), transfer out of the CPRD network/end of enrolment in Medicare Parts A, B, or D, or end of follow-up. Data from both the UK and US studies are from routinely collected data, so complete ascertainment of outcomes during defined follow up can be reasonably assumed, other than misclassification due to improper coding. Although the exposure categorisations were different in the two studies, they both used the same statistical methodology. The primary exposure variables were always included in unadjusted and adjusted (accounting for all covariates) Cox regression models with an underlying age timescale, to obtain hazard ratios.

7.3 COMPARISON OF RESULTS

This section aims to compare the results of the UK and US studies, both as they were reported in chapters 5 and 6, and when re-analysed so the methodology and analyses in each study were modified to make them as similar as possible. To allow for the modified comparison, both study populations were restricted until they were as similar, analyses were adjusted for covariates that were included in both studies, and covariates were identified were using the same methodology.

7.3.1 Methods

7.3.1.1 Comparison of results as originally reported in studies

The primary analysis of the ever exposure analysis in the US study was also carried out with the baseline exposure changed to ever tamoxifen use, so it was possible to make a direct comparison between both ever exposure analyses in the UK and US studies, retaining the original methodology from each study. For simplicity, only the results comparing the risk of CVD outcomes in ever AI compared with tamoxifen users were presented.

7.3.1.2 Modified comparison of results

This section outlines the modified study population, exposures, outcomes, covariates, and analyses used to re-analyse both studies to make them as comparable as possible. The methodology presented below was applied to both the UK and US studies separately, and results were compared.

Study population

Both study populations were modified to include all women with follow up in the data for 12-months before an incident breast cancer diagnosis. All women had to be aged 66 and over, with an incident ER+/PR+ and stage 1-3 breast cancer (in the US study, any breast cancer in the UK study) diagnosed between from 1st January 2002 and 31st March 2015 in the UK and 1st January 2008 and 31st December 2013 in the US, and be newly prescribed an AI or tamoxifen after their diagnosis. In addition, women were excluded if they had an endocrine therapy prescription prior to their breast cancer diagnosis. Follow-up began either one year after the date of breast cancer or at the date of first AI or tamoxifen prescription, whichever occurred latest (hereafter the 'index date'). Women

were excluded if prior to their index date they: died, discontinued from Medicare Parts A, B, or D/transferred out of the CPRD, were diagnosed with any cancer relating to sites other than the breast, or were diagnosed with the CVD event of interest (within a 3-year period prior to index). Overall, in the UK the requirements that were modified from the original study were: age was restricted to 66 years and over, rather than 54 years and over; and women were excluded if they had a CVD event of interest within 3 years prior to the index date rather than at any point prior to the index date. The only modification in the US was that all women were required to have an ER+/PR+ breast cancer diagnosis and an endocrine therapy prescription, rather than just an ER+/PR+ breast cancer diagnosis regardless of if they were prescribed an endocrine therapy.

Exposures

In both UK and US study populations, identification of incident tamoxifen and AI exposures used an appropriate prescription code (based on the code lists outlined in chapters 5 and 6 for CPRD and SEER-Medicare respectively). The primary exposure was AI use relative to tamoxifen use. To help elucidate how drug exposure is associated with risk, exposure was parameterised in two ways. First, ever exposure to endocrine therapy (ever use of tamoxifen, ever use of AI, ever use of both drugs). If a woman moved between tamoxifen and AI prescriptions, records were time-updated to indicate they had been exposed to both drugs from this point forward. Secondly, current exposure to endocrine therapy (categorised as current tamoxifen use, current AI use, no current therapy and previously ever exposed to an AI, no current therapy and previously exposed to tamoxifen only) was time-updated at any changes in therapy. A prescription was continuous if a further prescription followed within 30-days of the original prescription ending. Overall, the exposure groups in the UK were the same as the original study; and exposure groups in the US omitted the unexposed group and was parameterised like the UK version.

Outcomes

The full range of CVD outcomes explored in both the UK and US studies remained the primary outcomes of interest for this comparison analysis. Events were identified using Read/ICD-10 codes in CPRD/HES respectively, and ICD-9 codes in SEER-Medicare.

Covariates

A common set of covariates, available in both datasets, was adjusted for: year of breast cancer; age at index date (66-74, 75-84, 85+); time since index date ((<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of statins; use of ACE inhibitors; use of calcium channel blockers; use of angiotensin receptor blockers; rheumatoid arthritis; chronic kidney disease; diabetes; VTE; and non-venous CVD. Identification of diagnoses and prescriptions were restricted to a 3-year look back window prior to index date.

Analysis

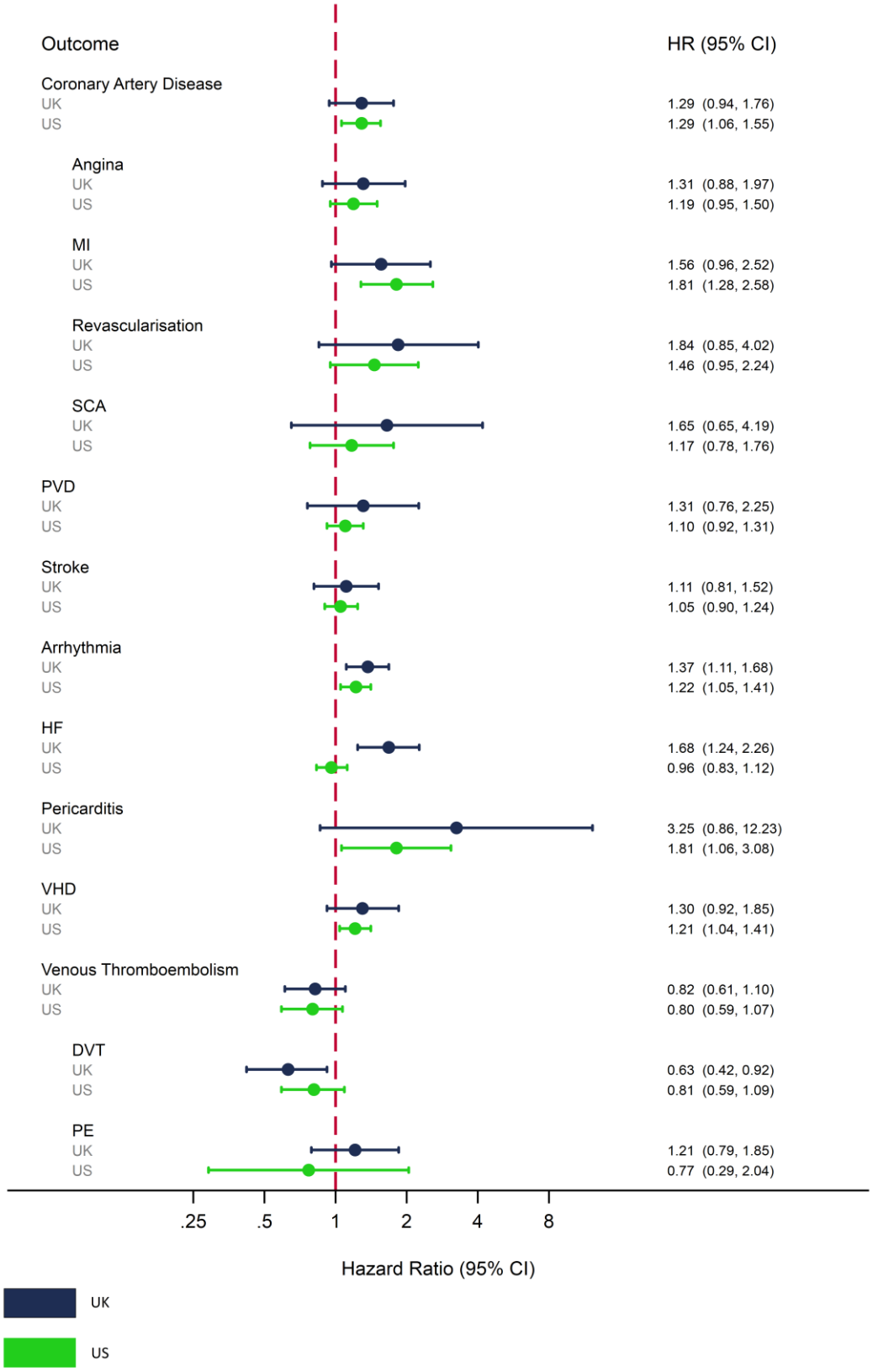
Observation time began at index date and ended at earliest of the following: a CVD event of interest, diagnosis of another cancer, death, transfer out of the CPRD network/end of enrolment in Medicare Parts A, B, or D, or end of follow-up (31st March 2016 in the UK and 31st December 2014 in the US). Prior to exploring the relationship between endocrine therapies and CVD, baseline characteristic distributions of patients who were initially prescribed tamoxifen or AIs were described. The primary exposure variables were then included in unadjusted and adjusted (accounting for all covariates) Cox regression models with an underlying age timescale, to obtain hazard ratios.

7.3.2 Results

7.3.2.1 Comparison of results as originally reported in studies

Generally, the original study results comparing the risk of outcomes associated with ever AI compared with tamoxifen use from the UK and US studies reported effect estimates that were in the same direction, with similar strengths (Figure 7.1). However, the results for HF were not consistent, with the UK study reporting an increased risk of HF associated with ever AI in comparison tamoxifen use (HR: 1.68, 95% CI: 1.24-2.26), and the US study reporting no evidence of an association (HR: 0.96, 95% CI: 0.83-1.12). There were also point estimates in opposing directions for the PE analysis, however 95% CIs crossed unity for both studies due to few outcomes (UK HR: 1.21, 95% CI: 0.79-2.04; US HR: 0.77, 95% CI: 0.29-2.04). Associations were in the same direction for both the UK and US studies for all other outcomes. For all but one of the outcomes (MI), the HR in the US study was slightly closer to the null association. The US study also generally had more precision due to the larger study population and resulting higher number of events

Figure 7.1: Adjusted HRs (adjusted for variables originally reported in each study) for the association between originally reported results for ever AI use compared with ever tamoxifen use and the risk of a range of clinical CVD outcomes in the UK and US studies (with baseline group changed to ever tamoxifen use)



7.3.2.2 Modified comparison of results

There were more women in the US study population in comparison with the UK study population when modified to allow for direct comparison (5665 and 18248 women in the UK and US study populations respectively, Table 7.3). A higher proportion of women were initially prescribed tamoxifen in the UK (44%) compared with the US (13%). Women in the US were generally younger (median age in US: 75 years, IQR: 71-81; median age in UK: 77 years, IQR: 71-83), but tamoxifen users were older in the US, whereas AI users were older in the UK. The proportion of initial tamoxifen users steadily decreased the later the breast cancer diagnosis in the UK, whereas the decline was not as sharp in the US. There was a higher total proportion of ACEi, CCB, and ARB users in the UK in comparison with the US, with a higher proportion of prescriptions in AI users in comparison with tamoxifen users in both countries. The total proportion of women diagnosed with diabetes and non-venous CVDs was much higher in the US in comparison with the UK, but the proportion of women diagnosed with both diabetes and non-venous CVD was higher in AI users in comparison with tamoxifen users in both the UK and the US. Finally, there was a higher total proportion of women diagnosed with CKD in the UK in comparison with the US, with a higher proportion of CKD diagnoses in AI users in comparison with tamoxifen users in both countries.

Ever exposure analyses

Associations between ever AI compared with ever tamoxifen use and the risk of all CVD outcomes were generally in the same direction for both the US and UK studies (Figure 7.2). There was evidence of an increased risk of MI, arrhythmia, pericarditis, and VHD associated with ever AI in comparison with tamoxifen use in the US study, and although there were associations in a similar direction in the UK study, estimates were less precise. The HF analysis was the only analysis in the UK study that showed evidence of an increased risk associated with ever AI compared with tamoxifen use (with a 95% CI that did not cross the null association), but in the US study there was no association. All other effect estimates were generally in the same direction and consistent between the two studies (apart from revascularisation). There were differences in the size of the effect estimates for the stroke, pericarditis, and DVT analyses, but the UK study lacked precision for these outcomes, and the US estimate was contained within the UK 95% CI. In this modified analysis, the precision of the UK study was considerably restricted due to a fewer women eligible for the final study population, and fewer resulting outcome events. There were also no events in the ever tamoxifen arm of the SCA analysis in the modified UK study, so effect measures could not be estimated.

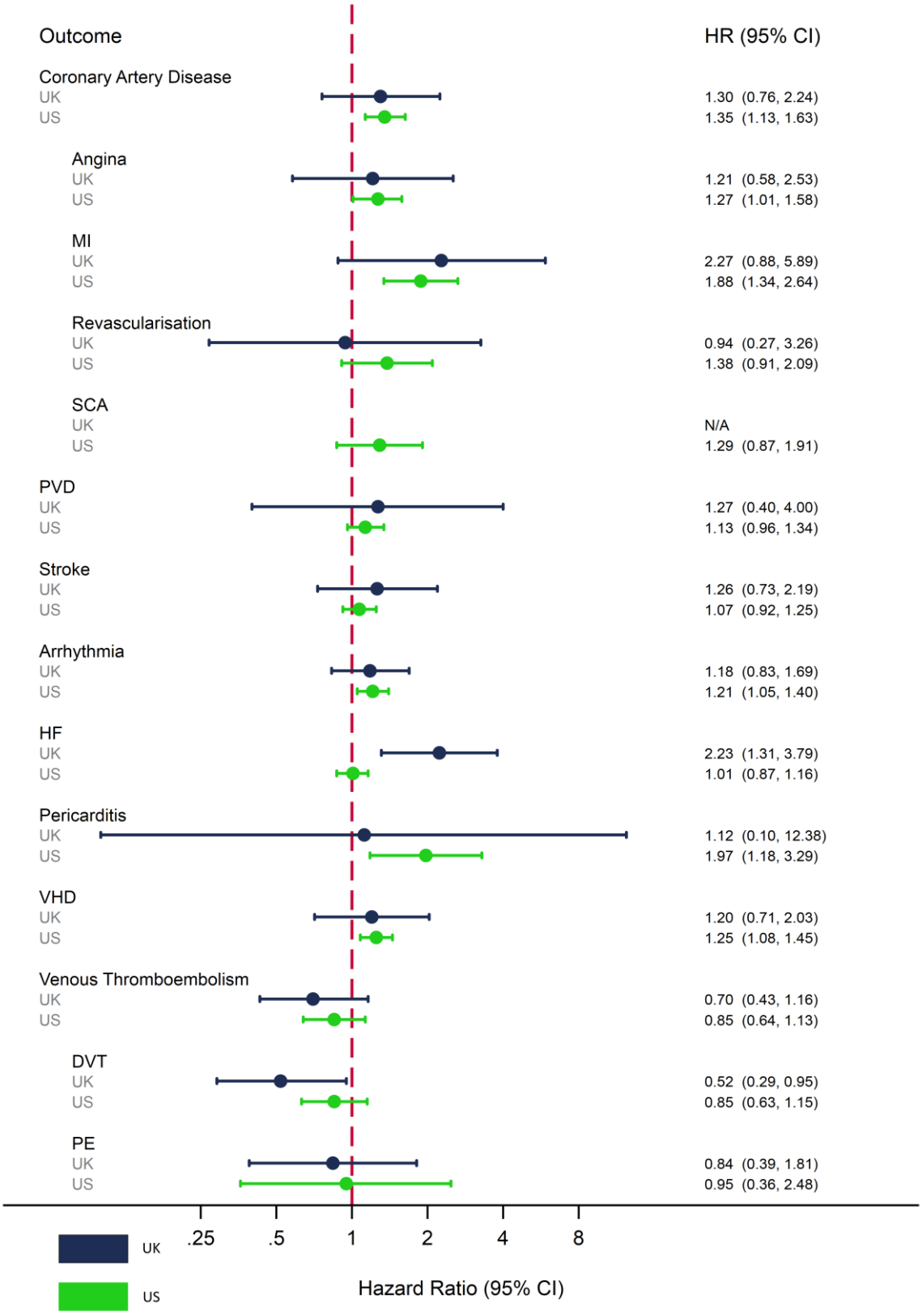
Table 7.3: Characteristics of modified study populations based on their initial exposure for both UK and US studies

	UK			US		
	Tamoxifen (%)	**AI (%)**	**Total (%)**	**Tamoxifen (%)**	**AI (%)**	**Total (%)**
N	2504 (100)	3161 (100)	5665 (100)	2410 (100)	15838 (100)	18248 (100)
Age at index (yrs)						
66-74	1222 (48.8)	1245 (39.4)	2467 (43.5)	944 (39.2)	7879 (49.7)	8823 (48.4)
75-84	896 (35.8)	1282 (40.6)	2178 (38.4)	1047 (43.4)	6181 (39)	7228 (39.6)
85+	286 (15.4)	632 (20.1)	1020 (18)	419 (17.4)	1778 (11.2)	2197 (12)
Median (IQR)	75 (70-82)	78 (71-84)	77 (71-83)	77 (72-83)	75 (71-81)	75 (71-81)
Year of breast cancer diagnosis						
2002	320 (12.8)	48 (1.5)	368 (6.5)			
2003	297 (11.9)	84 (2.7)	381 (6.7)			
2004	295 (11.8)	135 (4.3)	430 (7.6)			
2005	278 (11.1)	181 (5.7)	459 (8.1)			
2006	245 (9.8)	234 (7.4)	479 (8.5)			
2007	202 (8.1)	265 (8.4)	466 (8.2)			
2008	215 (8.6)	297 (9.4)	512 (9)	449 (18.6)	2244 (14.2)	2693 (14.8)
2009	149 (6)	304 (9.6)	453 (9)	477 (19.8)	2304 (14.5)	2781 (15.2)
2010	111 (4.4)	340 (10.8)	451 (8)	395 (16.4)	2513 (15.9)	2908 (15.9)
2011	134 (5.4)	353 (11.2)	487 (8.6)	346 (14.4)	2733 (17.3)	3079 (16.9)
2012	109 (4.4)	296 (9.4)	405 (7.1)	360 (14.9)	2839 (17.9)	3199 (17.5)
2013	88 (3.5)	301 (9.5)	389 (6.9)	383 (15.9)	3205 (20.2)	3588 (19.7)
2014	53 (2.1)	280 (8.9)	333 (5.9)			
2015	8 (0.3)	44 (1.4)	52 (0.9)			
Statins before index	685 (27.4)	1232 (39)	1917 (33.8)	840 (34.9)	6274 (39.6)	7114 (39)
ACEi before index	617 (24.6)	966 (30.6)	1583 (27.9)	425 (17.6)	2804 (17.7)	3229 (17.7)
CCB before index	680 (27.2)	1035 (32.7)	1715 (30.3)	307 (12.7)	2267 (14.3)	2574 (14.1)
ARB before index	330 (13.2)	516 (16.3)	846 (14.9)	219 (9.1)	1694 (10.7)	1913 (10.5)
RA before index	26 (1)	24 (0.8)	50 (0.9)	111 (4.6)	583 (3.7)	694 (3.8)
Diabetes before index	306 (12.2)	499 (15.8)	805 (14.2)	626 (26)	4809 (30.4)	5435 (29.8)
CKD before index	692 (27.6)	954 (30.2)	1646 (29.1)	170 (7.1)	1184 (7.5)	1354 (7.4)
Non-venous CVD before index	422 (16.9)	806 (25.5)	1228 (21.7)	1356 (56.3)	9429 (59.5)	10785 (59.1)
VTE before index	65 (2.6)	104 (3.3)	169 (3)	35 (1.5)	430 (2.7)	465 (2.5)

Current exposure analyses

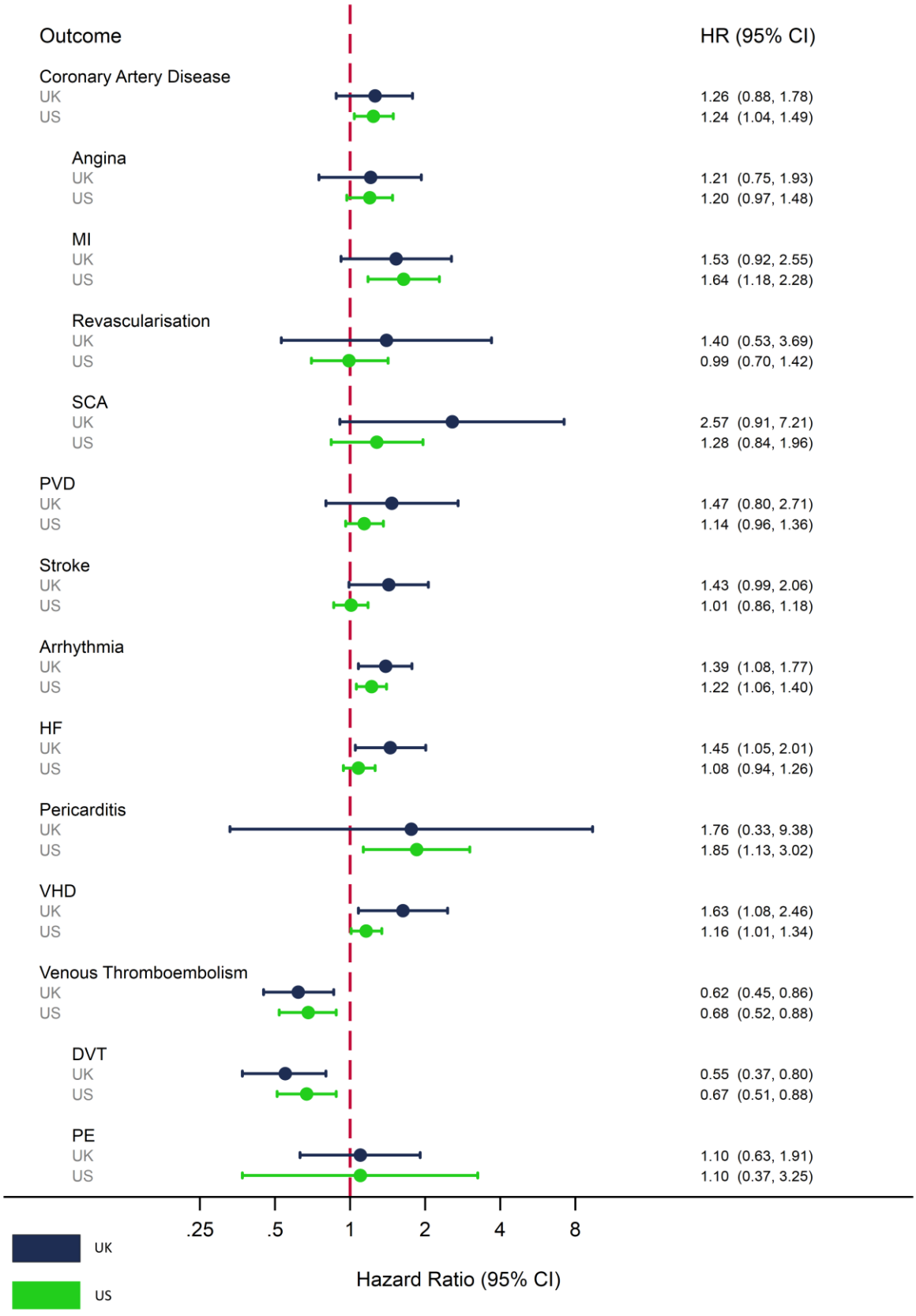
Associations between current AI compared with tamoxifen use and the risk of all CVD outcomes were again generally in the same direction for both the UK and US studies (Figure 7.3). There was evidence of an increased risk of MI, arrhythmia, pericarditis, and VHD associated with current AI compared with tamoxifen use in the US study, with the UK study reporting associations in a similar direction for all these outcomes. However, in the HF analysis, there was again evidence of an increased risk of an event associated with current AI compared with current tamoxifen use in the UK study, but no evidence of a difference in the US study. Although effect estimates were in a similar direction, there were some differences in effect sizes between the UK and US for the stroke, VHD, and SCA analyses, but the UK study again lacked precision for these outcomes. Within the modified analyses exploring the associations between both past with AI and past tamoxifen only, compared with current tamoxifen use, and the risk of all CVD outcomes (Appendix 7.2), there were generally few CVD events in the two past use categorisations, leading to low precision in the analyses from the UK study. However, most effect estimates were consistently in the same direction in the UK and US studies.

Figure 7.2: Adjusted HRs for the association between ever AI compared with ever tamoxifen use and a range of clinical CVD outcomes in modified analyses in both the UK and US



†Adjusted for: year of breast cancer; age at index date (66-74, 75-84, 85+); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of statins; use of ACE inhibitors; use of calcium channel blockers; use of angiotensin receptor blockers; rheumatoid arthritis; chronic kidney disease; diabetes; VTE; and non-venous CVD

Figure 7.3: Adjusted HRs for the association between current AI compared with current tamoxifen use and a range of clinical CVD outcomes in modified analyses in both the UK and US



†Adjusted for: year of breast cancer; age at index date (66-74, 75-84, 85+); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of statins; use of ACE inhibitors; use of calcium channel blockers; use of angiotensin receptor blockers; rheumatoid arthritis; chronic kidney disease; diabetes; VTE; and non-venous CVD

7.3.3 Discussion

7.3.3.1 Main findings

This section enabled a comparison of results from the UK and US studies, both as they were presented in chapters 5 and 6, and after modifying methodology to allow for a more direct comparison. The comparison of the original results generally reported agreement between the UK and US studies, with some small differences in effect sizes. In the analysis that modified both the UK and US study populations, covariates, and analyses, the final restricted population in the UK study was a lot smaller than the US study population, so there was less statistical power to detect associations and some effect estimates had 95% CIs that crossed unity. However, the directions of effect continued to stay the same in both the UK and US studies for most outcomes in both the ever and current exposure analyses.

The main difference in results between the two studies was when directly comparing the use of AI in comparison with tamoxifen and the risk of HF. In both the comparison of the original results, and the modified comparison, there was consistently a greater risk of HF associated with AI compared with tamoxifen use in the UK study, and no difference between the two groups in the US study. Results of secondary analyses from the original UK study suggested that the increased risk associated with AI compared with tamoxifen use was likely due to cardio-protective effects of tamoxifen, and results in the original US study suggested that tamoxifen use was indeed associated with a reduced risk of HF in comparison with an unexposed population. However, the US study also suggested a reduced risk of HF associated with AI use compared with an unexposed population, with no difference in the risk of HF between AI and tamoxifen users. Therefore, although the original US study agreed with the UK study that there is evidence that tamoxifen reduced the risk of HF and AI use does not increase the risk of HF, it also reported contrasting results for the relative difference in risk of HF when comparing AI and tamoxifen use.

7.3.3.2 Possible reasons for differences between studies

Study populations

There were more than double the number of women included in the original US study compared with the UK study, even though the study population was restricted to older women (aged 66 years and over in the US, aged 54 years and over in the UK), and those who were diagnosed with breast cancer within a shorter time period (2008-2013 in the US, 2002-2016 in the UK). There was a similar

discrepancy in size in the modified study populations, although the number of women included in the UK study nearly halved, whereas the US study was around 84% of the original size. The larger drop in size of study populations in the UK compared with the US after modification was due to the revised UK study excluding all women diagnosed with breast cancer between the ages of 54 and 66 years, whereas the US study only excluded women diagnosed with ER+/PR+ breast cancer that were not prescribed endocrine therapy.

In the modified study populations, the US population had a higher proportion of women with CVD diagnoses in the 3 years prior to index date, but the revised UK population had a higher proportion of CVD related prescriptions (ACEi, CCB, ARB), which is somewhat contradictory. However, the higher proportion of CVD related drug prescriptions in the UK could be given for the primary prevention of CVD, which is consequently reducing the burden of CVD. Furthermore, the study starting earlier in the UK may explain the higher proportion of women that were originally prescribed tamoxifen in the UK compared with the US, with women initiating tamoxifen in the earlier years due to limited knowledge of the additional efficacy of AIs over tamoxifen in postmenopausal women. However, the higher proportion of women with previous CVD in the revised US population, and the known association between tamoxifen and VTE could also influence the difference between AI and tamoxifen initiators. Clinicians are less likely to prescribe tamoxifen to those with prior CVD, even though this thesis suggests a protective association between tamoxifen and certain CVD outcomes.

AI users were older than tamoxifen users in the UK study, in both the original and modified analyses, with both a higher median age, and more women in the oldest age category. However, in the US study, tamoxifen users were older than AI users. Results show that for several non-venous CVD outcomes in both the ever and current exposure analyses, the effect estimates for events when comparing AI and tamoxifen users were larger in the UK compared with the US study, albeit with 95% CIs that crossed unity. The higher proportion of older AI compared with tamoxifen users could explain these differences due to the increased risk of CVD associated with increasing age.[119]

Confounding

There were several potential confounders unaccounted for in the modified analyses as the aim was to make the analyses as similar as possible, adjusting only for variables that were available in both studies. Residual confounding due to these variables that were not accounted for could therefore

explain any differences seen between the studies. The most prominent potential confounders omitted from the modified studies, and the original US study, were lifestyle factors such as BMI. Organisation for Economic Cooperation and Development (OECD) data show that the proportion of the population defined as obese is higher in the US (38.2%) than the UK (26.9%).[120] The original UK study reported that the BMI of AI users was higher than that of tamoxifen users. Although adjusting for BMI in the original UK study did not largely influence effect estimates, if there was a greater difference between BMI in AI and tamoxifen users in the US, then not adjusting for BMI could cause considerable residual confounding, and may explain any differences between effect estimated in both the original and modified UK and US studies.

The difference in HF results could be due to AI users receiving more cardio-toxic chemotherapy treatments, such as anthracyclines, compared with tamoxifen users in the UK, as there is a known increased risk of HF associated with increasing cumulative doses of anthracyclines.[75] The US study only included women diagnosed with stages 1-3 breast cancer, which may be a factor for any differences in treatment, as the UK study included all women with a breast cancer diagnosis recorded in the CPRD during the study period. Those diagnosed with DCIS or stage 4 breast cancer, who could be included in the UK study, are likely given different treatment regimens to those with stages 1-3 breast cancer. There were more anthracyclines administered to AI users compared with tamoxifen users in the US study. However, there was generally a low proportion of all women given these agents (5.2% of women overall given an anthracycline). If more women in the UK were given these anthracyclines, which included a higher proportion of AI users compared with tamoxifen users, then it is conceivable that this is driving an increased risk of HF in AI users. Although it is not possible to test this hypothesis within the UK data as information on cancer treatments were not available, the UK National Cancer Registration and Analysis Service reported that approximately 33% of all women diagnosed with breast cancer in the UK were administered any chemotherapy agent in 2013-2014.[121] It is likely that the proportion of post-menopausal women diagnosed with stage 1-3 ER+/PR+ breast cancer who are given an anthracycline is less than 33% as such therapy will not be given to those diagnosed with earlier stage and less aggressive breast cancers, and they could be treated with other, non-anthracycline based, chemotherapy agents. However, it is highly unlikely that the proportion of women given an anthracycline will be as low at the 5.2% reported in the US study. It is therefore reasonable to assume that more treatment was given in the UK compared with the US study population, although inferences into the difference of treatments between AI and tamoxifen users in the UK cannot be made.

Outcome definitions

Another reason for the difference in the HF results in the UK and US studies could be that the definition of HF within the two studies is slightly different. The UK study used Read codes to identify events in primary care, and ICD-10 codes to identify results in secondary care, whereas the US study only used ICD-9 codes. Although this was also case for all other CVD outcomes, a HF diagnosis has a more nuanced diagnosis, with the European Society of Cardiology stating that a HF diagnosis requires: symptoms typical of heart failure (breathlessness at rest or on exercise, tiredness, fatigue, ankle swelling); signs typical of heart failure (tachycardia, tachypnoea, pulmonary rales, pleural effusion, raised jugular venous pressure, hepatomegaly, peripheral oedema); and objective evidence of a structural or functional abnormality of the heart at rest (cardiomegaly, third heart sound, cardiac murmurs, abnormality on echocardiogram, raised B-type natriuretic peptides).[122] To help understand if potentially varying HF definitions were the reason for the differences in results, the UK analysis was replicated using restricted code lists which only included Read codes that fall under the Quality and Outcomes Framework (QOF) definition of HF (Appendix 7.3), and specifically HF ICD-10 codes (as original code list also included cardiomyopathy outcomes). However, even with the restricted outcome definition, ever AI users continued to have an increased risk of HF in comparison with ever tamoxifen users in the UK (adjusted HR: 1.95, 96% CI: 1.42-2.69).

Statistical chance

Statistical chance is also a possible explanation of the difference in HF results, or the results of any other CVD outcomes in both the original and modified analyses. If the true effect estimates for the ever-exposed analyses comparing AI to tamoxifen use and the risk of HF in either the UK or US studies were at the extremities of the 95% CIs, then results could be similar.

Differences in the nature of data

More generally, the nature of the data used could contribute towards differences in both study populations and results. Data from the UK CPRD and HES are EHRs recorded by clinicians in GP surgeries and hospitals, whereas data from SEER-Medicare in the US are claims based. GP surgeries and hospitals use EHRs to record details of all encounters with patients, allowing for a complete overview of their care, whereas generation of claims data come from bills submitted by physicians and hospitals for payment by commercial and government health plans.

Claims based data accurately record filled drug prescriptions as a record generation only happens

after payment for the prescribed drug. However, EHR records contains information on all drugs prescribed by the clinician, regardless of if that prescription was filled, which could lead to an overestimation of what was actually filled and eventually taken by the patient. This could explain the higher proportion of women prescribed statins, ACEis, ARBs, and CCBs in the UK study compared with the US study.

On the other hand, EHRs can extract information that may not be available in claims data. For example, EHRs record lab and test results, which allows a deeper understanding of diagnoses such as hypertension. The original US study used a combination of records of an ICD-9 code for hypertension and prescription of anti-hypertensive to identify hypertension, whereas the UK study used records of the patients' systolic and diastolic blood pressure in the UK. Although both methods will correctly identify if a woman was hypertensive, the method used in the UK will allows for more granular detail, and the US method may be prone to residual confounding as the risk of CVD is known to increase with increasing blood pressure.[123]

Claims data are also more temporally limited than EHR data as a claim reflects the diagnoses and services that occurred on the date the claim was submitted, meaning only incident diagnoses are usually recorded. Claims data do not to convey information about what happened in the past. However, UK based EHRs can capture historical diagnoses, and when a patient registers at a new GP, the standard process is to update their records to include all diagnoses from their previous surgeries. All analyses excluded women with a previous diagnosis of the outcome of interest, and the exclusions in the modified analyses were restricted to CVD diagnoses within three years prior to index date. If a woman registered with a new GP within this three-year period, or the GP had updated their records within this time, it is possible that women with prevalent diagnoses that occurred prior to this three-year period were not included due to GPs back-dating their records. However, exclusions only applied to those with an incident CVD event in the three-years prior to index date in the US study due to the nature of claims data. This could explain some of the differences between results in the two studies if the proportion of women excluded because of a historical, rather than an incident, diagnosis differed between those prescribed tamoxifen and AI in the UK study. It could also explain differences between other diagnoses such as RA, diabetes, and CKD. Theoretically, the difference between prevalent and incident disease should not be a problem when identifying CVD events during follow up, as the UK study dropped all Read codes that represented historical events when identifying events during follow up in CPRD. However, there is a possibility that GPs incorrectly coded historical events as incident events, in which case some events

in follow up may represent prevalent disease.

7.3.4 Conclusion

Overall, although restricting the study populations and methodology of the two studies allowed for a direct comparison between results, it also introduced more uncertainty. Smaller study populations and fewer events meant effect estimates were less precise than in the original studies, especially in the UK. However, there was general agreement between the results in the modified analyses of the UK and US studies, much like the comparison of the original results. HF remained the one outcome in which the studies disagreed, and residual confounding by variables not available in both datasets, such as lifestyle measures and other cancer treatments, or statistical chance, may drive this difference.

7.4 SUMMARY

- A comparison was made of the study populations and covariates adjusted in the UK and US studies. There was longer follow up and a wider age range in the UK study, but women diagnosed with ER+/PR+ breast cancer that were unexposed to any endocrine therapy were included in the US study. Information on breast cancer treatments and severity were not available in the UK data, whereas lifestyle measures were not available in the US data
- There was general agreement for the results of the ever-exposed analyses as originally presented in the UK and US studies. HF was the only outcome for which the two studies were in disagreement, with the UK study reporting evidence of an increased risk of HF associated with AI compared with tamoxifen use, and the US study reporting no association.
- After the modification of both study populations and analyses, making them as similar as possible, effect estimates were generally in the same direction, with some differences in effect sizes. The US study generally reported more precise effect estimates that were closer to the null association. However, there continued to be inconsistent results for the HF analyses.
- The difference in HF results may to be due to different definitions for HF, more cardio-toxic treatments used in AI users compared with tamoxifen users in the UK, or statistical chance.

8 SUMMARY AND CONCLUSIONS

8.1 INTRODUCTION

This chapter will begin with a concise overview of the findings of the main thesis, followed by comparing the findings of all individual CVD outcomes with previous literature identified in the systematic review in chapter 2. A discussion of the biological plausibility and overall strengths and limitations will follow. Finally, the work of the thesis will be contextualised in relation to its implications and the future research that should follow.

8.2 OVERVIEW OF RESEARCH

This section will provide an overview of what was included and the conclusions of each chapter. Section 8.3 will give further detail on the findings of the individual studies.

8.2.1 Chapter 2 - Systematic review

- A systematic review aimed to collate the current literature on the effect of both tamoxifen and AIs on the range of specific clinical CVD outcomes in women with a history of non-metastatic breast cancer.
- 26 eligible studies were identified that investigated the risk of seven specific CVD outcomes (venous thromboembolism, 15 studies; MI, 14 studies; stroke, 12 studies; angina, 4 studies; HF, 4 studies; arrhythmia, 1 study; and PVD, 1 study) associated with either tamoxifen use, aromatase inhibitor use, or a comparison of the two treatments, in all women with a history of non-metastatic breast cancer.
- There were no studies identified investigating revascularisation, SCA, pericarditis, or VHD.
- Results suggested an increased risk of VTE in tamoxifen users compared with both non-users and aromatase inhibitor users. Results were also consistent with a higher risk of the vascular diseases MI and angina in aromatase inhibitor users compared with tamoxifen users, but there was also a suggestion that this may be partly driven by a protective effect of tamoxifen on these outcomes. Data were limited, and evidence was generally inconsistent for all other cardiovascular disease outcomes.

- The main problem in RCTs was incomplete reporting of methods, which made it impossible to judge whether studies were prone to certain biases.
- All observational studies had at least one domain categorised as being at high risk of bias.

8.2.2 Chapter 3 - Data sources and creation of code lists

- All databases that were to be used in the PhD were introduced, including the CPRD and HES from the UK, and SEER-Medicare from the US, and my contribution in each study was documented.

8.2.3 Chapter 4 – Capture of CVD outcomes in the UK study

- A series of exploratory analyses using both CPRD and HES in the UK data were undertaken with aim of assessing the validity of CPRD to detect CVD outcomes, identifying missing Read codes and finalising code lists, and deciding on the use of CPRD and HES linkage data
- Events in the CPRD are identified using Read code lists. The algorithm used to create the original CVD Read code lists was presented, and the number of codes detected to identify CVD outcomes ranged from 128 in the revascularisation outcome to 20 in the PE outcome.
- A description was given of the proportion of definite and possible CVD events in the CPRD data. It was observed that most incident CVD outcomes in the CPRD were classed as definite events, and a large proportion of possible events were followed up by a definite event within a year. The decision was therefore made to include both types of events in all further analyses.
- The concordance of events between CPRD and HES was observed, and there was generally a low proportion of overlapping events in both CPRD and HES. Additional codes were then identified for arrhythmia, pericarditis, VHD, and revascularisation outcomes by exploring diagnoses in CPRD in the time around an event that only occurred in HES.
- There was similar statistical power to detect associations for all CVD outcomes if the linked CPRD and HES study population was used, rather than CPRD alone. To maximize sensitivity of outcome ascertainment, the decision was therefore made to restrict the study population to those that had linked HES data and include CVD outcome events identified from either CPRD or HES, rather than CPRD alone.

8.2.4 Chapter 5 - Assessing the effect of endocrine therapy use on the risk of cardiovascular disease: A cohort study in the UK

- Routinely collected primary care data from the UK CPRD database were used to assemble a cohort of post-menopausal female breast cancer survivors, who were prescribed tamoxifen or an AI from 2002-2016.
- The effect of endocrine therapy use (AI vs tamoxifen) on the risk of a range of clinically specific CVD outcomes were analysed using Cox regression, adjusted for potential confounders including year of breast cancer diagnosis, age, time since index date, current year smoking, BMI, alcohol status, IMD score, statins, ACE inhibitors, CCBs, ARBs, anti-platelets, diabetes*;* chronic kidney disease, rheumatoid arthritis, systolic blood pressure, diastolic blood pressure, history of VTE, and history of non-venous CVD.
- There was a pattern of an increased risk of non-venous CVDs in AI compared with tamoxifen users, with evidence of an increased risk of heart failure and arrhythmias in women ever exposed to AI compared with tamoxifen (adjusted HR: 1.70, 95% CI: 1.26-2.29; adjusted HR: 1.38, 95% CI 1.12-1.70 respectively). It is likely that these associations are driven by a protective effect associated with tamoxifen use as results also suggested past tamoxifen users were at a higher risk of both outcomes compared with current users. Results also confirmed the established reduced risk of DVT associated with AI in comparison with tamoxifen use (adjusted HR: 0.54, 95% CI: 0.36-0.83).
- A key limitation was that ER/PR status was not available in the CPRD data to allow an unexposed population of women diagnosed with ER+/PR+ breast cancer to be identified. Having these data would have meant the individual effects of tamoxifen and AI could have been disentangled, rather than just comparing the two treatments against each other.

8.2.5 Chapter 6 - Assessing the effect of endocrine therapy use on the risk of cardiovascular disease: a cohort study using the US SEER-Medicare linked database

- Claims based data from the US SEER-Medicare database were used to assemble a cohort of female ER+/PR+ stage 1-3 breast cancer survivors aged 66 and over from 2008-2013.
- The effect of endocrine therapy use (AI or tamoxifen) compared with non-use of any endocrine therapy on the risk of a range of clinically specific CVD outcomes were analysed using Cox regression, adjusted for potential confounders including year of breast cancer diagnosis, age, time since index date, current calendar year, race, SEER region, breast cancer

stage, breast cancer grade, taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, CCBs, ARBs, rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, history of VTE, and history of non-venous CVD.

- Apart from the established association between tamoxifen use and increased risk of DVT, there was no evidence of increased CVD risk with either tamoxifen or AI use compared with no endocrine therapy use. Results also suggested a protective association between tamoxifen use and risk of several non-venous CVDs with adjusted HRs ranging from 0.44 (95% CI: 0.30-0.63) in the MI analysis to 0.91 (95% CI: 0.75-1.10) in the PVD analysis.
- A similar, but weaker, protective association was also seen for non-venous CVDs during AI use, but there was a possibility of residual confounding due to an artificial increased risk of CVD in women that never initiated endocrine therapy that brings questions to the causality of these associations.

8.2.6 Chapter 7 - Comparison of UK and US studies

- The pattern of results in the UK study suggested that any difference in risk of non-venous CVD between AI and tamoxifen users may be driven by a protective effect of tamoxifen. The US study then provided evidence of a decreased risk of several non-venous CVDs associated with tamoxifen use. However, the US study also reported a reported a protective association between AI use and several non-venous CVDs, albeit less protective that tamoxifen.
- When directly comparing ever AI to tamoxifen use, the UK and US studies reported similar effect estimates for all CVD outcomes other than HF. The evidence of an increased risk of HF in AI compared with tamoxifen reported in the UK study, was not replicated in the US, which reported evidence of a similar strength of a protective association between both AI and tamoxifen use and the risk of HF.
- Study populations, covariate adjustments, and analyses were then modified so the studies were as similar as possible. Results were again generally consistent, even though there was a lack of precision for the some of the outcomes in the modified UK analysis. Even when analyses were modified, there continued to be a distinctive difference between the results reported for HF.

- It is thought that the difference in HF results between the UK and US studies could be due to either: more cardio-toxic chemotherapy agents being administered to AI users compared with tamoxifen users in the UK; residual confounding by BMI, smoking, or alcohol use in the US; or statistical chance.
- Differences in the nature of the data (EHRs in the UK and claims based data in the US) could also contribute towards any differences in results. For example, EHRs records lab and test results allowing for additional adjustment of potential confounders that are not recorded in claims data. Claims data accurately records filled drug prescriptions, whereas EHRs only record information of what has been prescribed, but not filled. Claims are also more temporally limited in comparison with EHRs.

8.3 OVERVIEW OF RESULTS FOR INDIVIDUAL CVD OUTCOMES AND COMPARISON WITH PREVIOUS LITERATURE

Of the CVD outcomes outlined in chapter 4, which were all explored in the UK and US studies, no previous literature was identified for the systematic review in chapter 2 for the association between endocrine therapy use and the risk of revascularisation, SCA, pericarditis, or VHD. The studies in this thesis are therefore the first to study the effect of endocrine therapy use on the risk of these CVD outcomes. However, results in both the UK and US studies for the effect of endocrine therapy use on the risk of all other CVD outcomes will be discussed in the context of previous literature.

8.3.1 Angina

- Results from the UK and US studies suggested HRs in the direction of a slightly increased risk of angina associated with ever AI compared with tamoxifen use (adjusted UK HR: 1.31, 95% CI: 0.88-1.97; adjusted US HR: 1.19, 95% CI: 0.95-1.50), although the 95% CIs were also compatible with no association. This was similar to results reported by one previous RCT that compared AI to tamoxifen use (RR: 1.37, 95% CI: 0.92-2.05).[124]
- The US study reported an effect estimate that suggested evidence a protective association between ever tamoxifen use and the risk of angina (ever tamoxifen vs unexposed adjusted HR: 0.88, 95% CI: 0.68-1.10), albeit with little precision and again with a 95% CI that is compatible with no association. Although two previous observational studies disagreed on

effect size, they both reported point estimates that also suggested a protective effect of tamoxifen (tamoxifen vs no tamoxifen RR: 0 40, 95% CI: 0.20-0.80; and RR: 0.88, 95% CI: 0.65-1.20).[125, 126] In the US study there was no evidence that any protective effects of tamoxifen persisted once tamoxifen prescription is stopped (adjusted HR for past tamoxifen vs unexposed: 0.93, 95% CI: 0.61-1.41), although there was little precision in the current exposure analysis.

- The US study reported no evidence of association between ever AI use and risk of angina (ever AI vs unexposed adjusted HR: 1.05, 95% CI: 0.89-1.35), which did not agree with the increased risk in AI users suggested in a previous RCT (AI vs placebo RR: 1.35, 95% CI: 1.17-1.56).[127]
- Overall, there is no statistical evidence to suggest that is a difference in risk of angina between AI and tamoxifen users. However, point estimates are consistent with the pattern seen in some other non-venous CVDs, whereby there is a higher risk in AI compared with tamoxifen users, which is driven by a protective effect of tamoxifen, but more statistical power is required to confirm within the data presented in this thesis. The effect of AIs on angina remains unclear, but any difference in risk of angina associated with AI use is likely to be minimal.

8.3.2 MI

- Both the UK and US studies suggested evidence of an increased risk of MI associated with ever AI compared with tamoxifen use (adjusted UK HR: 1.56, 95% CI: 0.96-2.52, adjusted US HR: 1.81, 95% CI: 1.28-2.58), with greater statistical power to detect an association in the US study. A similar direction and strength of association was reported in all but one of five previous studies that compared AI to tamoxifen use (RRs ranged from 1.50-2.29 in both RCT and observational studies).[128-132]
- The US study suggested evidence of a large protective effect of ever tamoxifen use associated with the risk of MI (ever tamoxifen vs unexposed adjusted HR: 0.44, 95% CI: 0.30-0.63). Three RCTs and two prior observational studies reported effect estimates in the same direction when comparing tamoxifen use to either placebo or no tamoxifen, but the strength of associations reported varied greatly (RRs ranged from 0.20-0.83).[125, 133-136] However, three observational studies also reported effect estimates suggesting either no difference in risk of MI associated with tamoxifen use, or a toxic effect of tamoxifen.[126, 137, 138]

- The US study also suggested that any protective association between tamoxifen use and MI may persist once tamoxifen use has ended (past tamoxifen vs unexposed adjusted HR: 0.46, 95% CI: 0.25-0.85), which is a possible argument against a causal effect as results may be due to residual confounding.
- The US study goes further to suggest evidence of a small protective effect of ever AI use associated with the risk of MI (ever AI vs unexposed adjusted HR: 0.79, 95% CI: 0.64-0.97). A similar association was alluded to in a previous RCT and observational study (AI vs placebo RR: 0.82, 95% CI: 0.55-1.22; AI vs no AI RR: 0.90, 95% CI: 0.65-1.25 respectively), but both results were also consistent with no association.[127, 137]
- This thesis has suggested an increased risk of MI in AI compared with tamoxifen users. It also suggests that this difference in risk is likely driven by a protective effect of tamoxifen on MI risk, because although there is mixed evidence of the effect of tamoxifen use on the risk of MI, there is no evidence in either the US study or any prior studies, that AI use increased risk of MI. In fact, the US study suggests a small protective effect of AI use on the risk of MI, but it unclear if this relationship is causal due to potential residual confounding because of unmeasured variables such as frailty, poor CVD preventative care, and high BMI which may be more likely in non-initiators of endocrine therapy, and could mean an artificially increased risk of MI in non-initiators at beginning of follow up. This residual cofounding could also explain part of the large protective association between tamoxifen use and MI reported in the US study.

8.3.3 Revascularisation (no previous evidence)

- Both the UK and US studies suggested effect estimates that pointed towards an increased risk of revascularisation procedures associated with ever AI compared with tamoxifen use (adjusted UK HR: 1.84, 95% CI: 0.85-4.02, adjusted US HR: 1.46, 95% CI: 0.95, 2.24), albeit with 95% CIs that crossed unity.
- Results from the US study also suggested weak evidence of a decreased risk of revascularisation associated with ever tamoxifen use compared with non-users of any endocrine therapies (ever tamoxifen vs unexposed adjusted HR: 0.63, 95% CIL 0.39-1.01). There was, however, an attenuation of this effect estimate when comparing risk in the time currently exposed to tamoxifen to time unexposed to any endocrine therapy (current tamoxifen vs unexposed adjusted HR: 0.89, 95% CI: 0.59-1.35).

- There was no evidence of an association between ever AI use and revascularisation in the US study (ever AI vs unexposed adjusted HR: 0.91, 95% CI: 0.68-1.23).
- Both studies reported effect estimates that suggested an increased risk of revascularisation in AI compared with tamoxifen users, which is likely driven by a decreased risk in tamoxifen users rather than any toxic effect of AI use. However, there were too few events to precisely estimate the reported effect estimates in both studies, so all results were generally inconclusive.

8.3.4 Sudden cardiac arrest (no previous evidence)

- Both the UK and US studies suggested HRs in the direction of an increased risk of SCA associated with ever AI compared with tamoxifen use (adjusted UK HR: 1.65, 95% CI: 0.65-4.49, adjusted US HR: 1.17 95% CI: 0.78-1.76), although the 95% CIs were also compatible with no association.
- The US study suggested some evidence of a decreased risk of SCA associated with both ever tamoxifen and AI use, when compared with non-users (ever tamoxifen vs unexposed adjusted HR: 0.67, 95% CI: 0.43-1.04; ever AI vs unexposed adjusted HR: 0.78, 95% CI: 0.59-1.04). Effect estimates were similar during time currently exposed to either drug (current tamoxifen vs unexposed adjusted HR: 0.58, 95% CI: 0.36-0.92; current AI vs unexposed adjusted HR: 0.64, 95% CI; 0.48-0.87), and attenuated towards the null once exposure ended. The UK study also suggested a direction of association that was consistent with an increased risk of SCA associated with time after stopping tamoxifen in comparison with time currently exposed (past tamoxifen only vs current tamoxifen adjusted HR: 3.47, 95% CI: 0.70-17.04), and although this would also suggest a decreased risk during time currently exposed to tamoxifen, there were too few events to draw conclusions.
- To summarise, the US study suggested weak evidence that use of both drugs individually lowered risk of SCA, but tamoxifen use may have larger decreased risks. However, problems around residual confounding leading to an artificial increased risk of SCA in those that do not initiate therapy (outlined in the MI section above) still persist, and all results are fairly inconclusive given the lack of precision.

8.3.5 PVD

- Only one RCT has previously explored the effect of endocrine therapies on the risk of PVD, which reported an inconclusive association between AI and tamoxifen use, but the direction of the effect estimate alluded to an increased risk of PVD associated with in AI in comparison with tamoxifen use (RR: 1.25, 95% CI: 0.68-2.29).[130] Both the UK and US studies suggested similar effect estimates in the same direction and of a similar strength (adjusted UK HR: 1.31, 95% CI: 0.67-2.25; adjusted US HR: 1.10, 95% CI: 0.92-1.31), but results were again imprecise.
- Although the effect estimate in the US study suggested a protective association between both ever and current tamoxifen use, and the risk of PVD, the strength of association was small, and the estimate had 95% CIs that crossed unity (ever tamoxifen vs unexposed adjusted HR: 0.91, 95% CI: 0.75-1.10; current tamoxifen vs unexposed adjusted HR: 0.85, 95% CI: 0.70-1.03).
- There was no evidence of association between AI use and the risk of PVD reported in the US study in both ever and current users of AIs (ever AI vs unexposed adjusted HR: 1.00, 95% CI: 0.87-1.14; current AI use vs unexposed adjusted HR: 0.95, 95% CI: 0.82-1.09)
- Although all results are somewhat imprecise, the addition of the UK and US study, with similar effect sizes, increases the evidence that there may be a small increased risk of PVD in AI compared with tamoxifen users. Any increase in risk is likely to be driven by a protective effect of tamoxifen on PVD risk, but it is possible that all reported associations could be due to chance because few events and a lack of precision.

8.3.6 Stroke

- Both the UK and US study suggested no evidence for an association between ever AI compared with tamoxifen use, and the risk of stroke (UK adjusted HR: 1.11, 95% CI: 0.81-1.52; US adjusted HR: 1.05, 95% CI: 0.90-1.24). Although effect estimates pointed in a direction that suggested an increased risk associated with AI use, effect estimates were small and imprecise. Previous evidence on the association between AI compared with tamoxifen users and the risk of stroke was mixed, with three studies reporting effect estimates suggesting an lower risk of stroke in AI users,[128, 133, 139] and three studies

reporting effect estimates suggesting a higher risk in AI users.[140-142] Two of the studies suggesting a higher risk in AI users were extremely imprecise.

- The US study suggested evidence of a decreased risk of stroke associated with both ever tamoxifen and AI use (ever-tamoxifen vs unexposed adjusted HR: 0.82, 95% CI: 0.69-0.98; ever-AI vs unexposed adjusted HR: 0.87, 95% CI: 0.76-0.98), which was also observed when comparing current tamoxifen and AI use to those unexposed (current tamoxifen vs unexposed adjusted HR: 0.83, 95% CI: 0.69-0.99; current AI vs unexposed adjusted HR: 0.81, 95% CI; 0.71-0.93). It was also suggested that the effect persisted once tamoxifen exposure ended, but not when AI exposure ended (past tamoxifen vs unexposed adjusted HR: 0.73, 95% CI: 0.54-1.00; past AI use vs unexposed adjusted HR: 1.02, 95% CI: 0.89-1.18).
- Six previous studies, one RCT and five observational, reported evidence for the effect of tamoxifen use on the risk of stroke, with effect estimates generally pointing towards a reduced risk associated with tamoxifen use.[126, 133, 136, 137, 139, 143] However, the strength of association varied (RR's for tamoxifen vs either placebo no tamoxifen ranged from 0.52-1.15), and only one observational study that reported a decreased risk associated with tamoxifen use was statistically powered to detect an association. One RCT and one observational study also reported effect estimates in opposing directions for the effect of AI use on the risk of stroke (RCT, AI vs placebo RR: 1.14, 95% CI:.89-1.45; observational study, AI vs no AI RR: 0.71, 95% CI: 0.43-1.03).
- The addition of evidence from this thesis intersects the previous evidence and suggests no difference in the risk of stroke between AI and tamoxifen users. However, the US study suggests that this is in the context of a small reduction in risk of stroke in both tamoxifen and AI users. As previous evidence on the effect of both tamoxifen and AIs on this risk of stroke is mixed, no definite conclusion can be made regarding these effects.

8.3.7 Arrhythmia

- Both the UK and US studies suggested an increased risk of arrhythmia associated with ever AI compared with tamoxifen use (UK adjusted HR: 1.37, 95% CI: 1.11-1.68; US adjusted HR: 1.22 95% CI: 1.05-1.41), and no other study had previously made this comparison.
- The US study suggested evidence of a decreased risk of arrhythmia associated with both ever tamoxifen and AI use (ever tamoxifen vs unexposed adjusted HR: 0.75, 95% CI: 0.63-

0.88; ever-AI vs unexposed adjusted HR: 0.91, 95% CI: 0.81-1.01). The US study also observed an increased risk of arrhythmia associated with current tamoxifen and AI use in comparison with those unexposed (current tamoxifen vs unexposed adjusted HR: 0.72, 95% CI: 0.61-0.84; current AI vs unexposed adjusted HR: 0.86, 95% CI: 0.77-0.97), which did not persist once exposure ended (past tamoxifen only vs unexposed adjusted HR: 0.99, 95% CI: 0.77-1.26; past with AI vs unexposed adjusted HR: 1.02, 95% CI: 0.89-1.17). The UK study also suggested evidence of an increased risk of arrhythmia associated with time after stopping tamoxifen therapy, in comparison with time currently exposed (past tamoxifen only vs current tamoxifen adjusted HR: 1.90, 95% CI: 1.38-2.62), suggesting a decreased risk during time exposed. One previous RCT also reported an effect estimate, albeit imprecise, suggesting a decreased risk of arrhythmia associated with tamoxifen use (RR: 0.79, 95% CI: 0.64-1.14).[134]

- Evidence presented in this thesis is consistent with an increased risk of arrhythmia in AI users compared with tamoxifen users, and there is increasing evidence to suggest that this is likely driven by a protective effect of tamoxifen. However, this is the first place that has also suggested that a small protective effect of AI use on the risk of arrhythmia. All protective effects reported are also prone to the residual confounding issues described in the MI outcome above, which may wholly explain the association between AI use and risk of arrhythmia, due to the effect estimate being closer to the null association in comparison with the association between tamoxifen use and risk of arrhythmia.

8.3.8 HF

- HF is the only outcome in which the UK and US studies reported results with differing conclusions when comparing the risk of HF in AI compared with tamoxifen users. The UK study suggested evidence of an increased risk of HF associated with ever AI compared with tamoxifen use (adjusted HR: 1.68, 95% CI: 1.24-2.26), and the US study suggested no difference in risk of HF between users of the two types of endocrine therapy (adjusted HR: 0.96, 95% CI: 0.83-1.12). There is a similarly mixed picture in the previous literature; one RCT reported evidence of an increased risk of HF associated with AI compared with tamoxifen use with a smaller effect estimate than the UK study (RR: 1.20, 95% CI:.04-1.38);[140] and an observational study reported no difference in risk between users of the two types of endocrine therapy (RR: 0.96, 95% CI: 0.77-1.08).[139]

- In the US study, the above association was in the context of evidence of a decreased risk of HF associated with both ever tamoxifen and AI use (ever tamoxifen vs unexposed adjusted HR: 0.87, 95% CI: 0.74-1.02; ever AI vs unexposed adjusted HR: 0.84, 95% CI: 0.75-0.94). Similar effect estimates were observed when comparing current tamoxifen and AI use to non-use (current tamoxifen vs unexposed adjusted HR: 0.77, 95% CI: 0.65-0.91; current AI vs unexposed adjusted HR: 0.79, 95% CI: 0.70-0.89), which was not preserved during time after exposure had ended (past tamoxifen only vs unexposed adjusted HR: 1.05, 95% CI: 0.82-1.34; past with AI vs unexposed adjusted HR: 0.99, 95% CI: 0.86-1.14). The UK study correspondingly reported an effect estimate, albeit imprecise, that suggested evidence of an increased risk of HF associated with time after tamoxifen exposure compared with time currently exposed (past tamoxifen only vs current tamoxifen adjusted HR: 1.33, 95% CI: 0.81-2.18). Previous evidence is mixed, with an RCT and observational study reporting evidence pointing towards a protective association between tamoxifen use and HF,[134, 139] but another observational study reporting no evidence of an association.[126]
- Although the UK study suggested evidence of an increased risk of HF in AI users compared with tamoxifen users, the US study suggested that, any difference in risk is not due to a toxic effect associated with AI use, and rather a protective effect of tamoxifen use. However, although the US study adds to the evidence that tamoxifen use decreases risk of HF, there still continues to be uncertainty because of mixed evidence in previous studies. The association between AI use and risk of HF is also still unclear due to the differing results in the UK and US studies, and no previous studies having assessed this association.

8.3.9 Pericarditis (no previous evidence)

- The US study suggested an increased risk of pericarditis associated with ever AI compared with tamoxifen use (adjusted HR: 1.81, 95% CI: 1.06-3.08). The direction of effect was replicated in the UK study, however fewer events meant less precision and an inconclusive result (ever AI vs ever tamoxifen adjusted HR: 3.25, 95% CI: 0.86-12.23).
- The US study suggested evidence of a decreased risk of pericarditis associated with both ever tamoxifen and AI use when compared with non-users, which was considerably larger in tamoxifen users (ever tamoxifen vs unexposed adjusted HR: 0.37, 95% CI: 0.21-0.65; ever AI vs unexposed adjusted HR: 0.67, 95% CI: 0.50-0.90). There was evidence of effect estimates in a similar direction when comparing the risk of pericarditis during time currently exposed

to tamoxifen or AI to non-use (current tamoxifen vs unexposed adjusted HR: 0.77, 95% CI: 0.65-0.91; current AI vs unexposed adjusted HR: 0.79, 95% CI: 0.70-0.89).

- If a real difference in the risk of pericarditis exists between AI and tamoxifen users, the US study suggests that use of both drugs individually have a protective effect on pericarditis, but tamoxifen may have a larger effect. However, the size of the protective effect of tamoxifen appears implausibly large, which raises questions about the causality of the association and reasons for non-causality outlined in the MI outcome could also be present in this outcome. Furthermore, cancer can present with pericarditis,[144] which may explain some of the protective association if the pericarditis recording was delayed and was associated with not being prescribed any endocrine therapy. The UK study was also severely underpowered to detect any associations, meaning the evidence from this study was inconclusive.

8.3.10 Valvular heart disease (no previous evidence)

- The UK and US studies reported effect estimates that pointed towards an increased risk of VHD associated with ever AI compared with tamoxifen use (UK adjusted HR: 1.30, 95% CI: 0.92-1.85; US adjusted HR: 1.21, 95% CI: 1.04, 1.41).
- The US study suggested evidence of a protective association between ever tamoxifen use and risk of VHD (ever tamoxifen vs unexposed adjusted HR: 0.81, 95% CI: 0.68-0.96). There was a similar association when comparing time currently exposed to tamoxifen to non-use (current tamoxifen vs unexposed adjusted HR: 0.82, 95% CI: 0.69-0.97), which attenuated towards the null once exposure finished (past tamoxifen only vs unexposed adjusted HR: 0.92, 95% CI: 0.70-1.20). There was also evidence of effect modification by current age (p for effect modification = 0.04), with no evidence of association between tamoxifen use and risk of VHD in the younger age group (ever tamoxifen vs unexposed adjusted HR in women currently aged 66-74: 1.10, 95% CI: 0.80-1.52), and a decreased risk of VHD associated with tamoxifen use in the oldest age group (ever tamoxifen vs unexposed adjusted HD in women currently aged 85+: 0.56, 95% CI: 0.39-0.80). The UK study also suggested weak evidence of an increased risk of VHD associated with time after tamoxifen exposure compared with time currently exposed (adjusted HR: 1.66, 95% CI: 0.95-2.91), implying a potential increased risk associated with current tamoxifen use.

- There was no evidence of an association between ever AI use and risk of VHD in the US study (ever AI vs unexposed adjusted HR: 0.98, 95% CI: 0.87-1.13).
- Evidence in both the UK and US studies suggests an increased risk of VHD in AI users compared with tamoxifen users, which is likely driven by the protective effect of tamoxifen uses and VHD risk as observed in the US study. However, the US study also suggested that this protective effect was driven by those in the older age groups, with no protective effects seen in younger women. There was also no evidence of any toxic or protective effects associated with AI use. As these studies are the first to report results on these effects, and there is the possibility of residual confounding by factors such as frailty are even greater in older populations, causality of any protective effects of tamoxifen cannot be assumed.

8.3.11 Venous Thromboembolism

- This comparison will focus on the composite VTE outcome, which is consistent with the VTE outcome in the systematic review that included DVT individually, or both DVT and PE together
- Both the UK and US studies suggested effect estimates pointing towards evidence of a decreased risk of VTE associated with ever AI compared with tamoxifen use (UK adjusted HR: 0.82, 95% CI: 0.61-1.10; US adjusted HR: 0.80, 95% CI: 0.59-1.07). Six out of seven previous RCTs reported effect estimates in the same direction, but with effect estimates further away from the null association (RRs comparing AI to tamoxifen use ranged from 0.26-0.61),[128, 129, 131, 140, 145] and another RCT reported no evidence of association.[124]
- The US study suggested evidence of an increased risk of VTE associated with ever tamoxifen use (ever tamoxifen vs unexposed adjusted HR: 1.39, 95% CI: 0.98-1.98), and a larger effect estimate was reported when comparing time currently exposed to non-use (current tamoxifen vs unexposed adjusted HR: 1.68, 95% CI: 1.21-2.35). The apparent toxic effect associated with tamoxifen use also disappeared once tamoxifen exposure ended (past tamoxifen only vs unexposed adjusted HR: 1.04. 95% CI: 0.56-1.92), although few events meant this estimate was imprecise. The UK study also observed a decreased risk of a VTE associated with time after tamoxifen exposure compared with time currently exposed (adjusted HR: 0.74, 95% CI: 0.44-1.23), suggesting an increased risk during time exposed, but results were consistent with chance variation. An increased risk of VTE associated with tamoxifen use was reported in five out of six previous RCTs and two out of three

observational studies, with varying levels of precision (RRs comparing tamoxifen use to placebo or no tamoxifen ranged from 1.64-7.10).[133-135, 146-150]

- There was no evidence of association between ever AI use and risk of VTE in the US study (ever AI vs unexposed adjusted HR: 1.11, 95% CI: 0.84-1.46), which contradicts the one previous RCT that indicated an increased risk associated with AI use (AI vs placebo RR: 1.84. 95% CI: 1.11-3.04).[127]
- The majority of evidence, including that reported in this thesis, suggests that AI users are at a decreased risk of VTE outcomes in comparison with tamoxifen users, which is likely driven by the long established increased risk of VTE in tamoxifen use. This thesis also hypothesises that any toxic effects may stop when exposure ends, but evidence on this effect is not conclusive. Effect sizes presented in this thesis were smaller than in previous studies that reported evidence of effect between both AI compared with tamoxifen use and tamoxifen use compared with non-use, on the risk of VTE. The potential underestimation in risk of VTE in tamoxifen users in comparison with several previous studies could be further evidence of an artificially increased risk of CVD in non-initiators of endocrine therapy due to residual confounding, suggested as a possible reason for non-causal relationships within other CVD outcomes. There is still uncertainty of the effect of AI use and one the risk of VTE outcomes.

8.4 BIOLOGICAL PLAUSIBILITY

As outlined in the introduction, It is known that tamoxifen use lowers total serum cholesterol by 10-20% and low-density lipoprotein levels by 15-22%,[92-95]. This suggests that the results reported in the US study suggesting a possible protective effect associated with tamoxifen use on the risk of several non-venous CVD outcomes may be mediated through tamoxifen reducing the level of CVD risk factors such as cholesterol. The heterogeneous protective effect sizes reported between tamoxifen use and the different non-venous CVD outcomes may therefore partly be a product of any differing effects of cholesterol on the different outcomes. Tamoxifen also has oestrogen agonistic effects that could increase the risk of in thrombogenicity through a reduction in antithrombin and protein C levels, [96, 97] which may explain the reported increased risk of VTE outcomes associated with tamoxifen use in both this thesis, and previous studies.

AIs work by inhibiting the aromatase enzyme and depleting oestrogen levels, which are known to be protective of CVD. Evidence regarding the effect of AIs on cholesterol levels is inconclusive, with two

RCTs reporting a higher incidence of hypercholesterolemia in AI compared with tamoxifen users,[151, 152] but another study finding no differences.[127] However, neither of the studies within the thesis reported evidence that the oestrogen depleting effects of AIs translated into an increased risk of any CVD outcome associated with AI use. There were some results that suggested evidence of decreased risk of some CVD events associated with AI use, but if true, there is no current understanding of a biological mechanism through which such an effect is mediated.

8.5 STRENGTHS AND LIMITATIONS

The study specific strengths and limitations were outlined within individual studies. This section will reiterate and expand on these points as well as discussing the wider implications of using routinely collected data for research and the potential problems of cross-country comparisons.

8.5.1 Strengths

- A key strength of using routinely collected data for research is the large size of the many available data sources and breadth of these data, which allows studies of exposures and outcome combinations that are otherwise too rare to study. It would be extremely expensive and time consuming to carry out a prospective cohort study to detect the effect of endocrine therapies in breast cancer survivors on the risk of each individual CVD outcome outlined in previous chapters.
- Unlike traditional prospective cohort studies that require participants to agree to be followed up, potentially inducing healthy participator bias, data used in this thesis are collected routinely, and are highly representative of the populations from which they originate.
- This thesis used data from two independent databases in the UK and US to address the same research question. This approach helps to guard against incorrect conclusions due to biases present in a single database, and consistent patterns of results increases the confidence in the conclusions.

8.5.2 Limitations

8.5.2.1 Misclassification

Exposure

- Although all prescriptions are captured in primary care in the CPRD and Medicare Part D, no information is recorded on the administration of these drugs. Women who are prescribed endocrine therapies (or any other prescription used as a covariate in each study) but do not take the drug, would be misclassified as exposed. However, given the severity of a breast cancer diagnosis, and the current evidence that endocrine therapies reduce recurrence, women will likely be mindful to take prescribed drugs. Furthermore, 97% of women in the UK and 90% in the US continued to pick up repeat prescriptions 1 year after starting therapy, which is an indication that most prescriptions were being filled and taken.
- Endocrine therapies could be prescribed in hospital at the first instance, followed by prescriptions in primary care. These prescriptions were not included in the UK as HES data does not include information on hospital prescribing, which could have led to delayed date of initiation in CPRD records. However, follow up started one year after breast cancer diagnosis, and any misclassification in regards to time to exposure in CPRD will likely be non-differential as time between breast cancer diagnosis and first prescription in primary care was similar in initiators of both tamoxifen and AIs (median time to initiation was 57 and 63 days in tamoxifen and AI initiators respectively). Missing such prescriptions was not a problem in the US study as prescriptions from both hospital and primary care are captured in either Medicare Parts A, B, or D.
- In the ever exposure analyses in both studies, women were defined as exposed from the point of their first tamoxifen or AI prescription until they were either censored or switched between drugs (at which point their exposure were time updated to be ever exposure to both drugs). Women could therefore receive only one prescription, but stay within the ever exposed group. Although this remains within the definition of being ever being exposed to the drug, it is unlikely that women who were exposed to only one endocrine therapy prescription received any protective or toxic effects of the drug. However, stopping therapy after one prescription is also unlikely given the known reduction in breast cancer recurrence associated with use of both tamoxifen and AIs, and again, descriptive analyses in both studies showed that 97% and 90% of women were exposed for at least 1 year in the UK and US studies respectively.

Outcome

- EHRs may not record acute CVD events that result in death prior to admission to hospital, so some women with an event could have been incorrectly censored at death when they actually suffered a CVD event of interest. This differential misclassification could have biased results if the proportion of acute CVD events leading immediately to death that were not recorded as CVD events varied between exposure groups. Such variation is conceivable given the differences in characteristics that may be associated with a higher risk of acute CVD between exposure groups (higher BMI and more smoking in AI initiators in the UK study, and more taxanes, anthracyclines, and trastuzumab in the AI initiators in the US study), but the assumption is not testable due to no information on events that occurred outside of the hospital resulting in death. This misclassification is likely to be more problematic for outcomes such as sudden cardiac arrest, where overall survival following an out of hospital event is approximately 8.3% in the US.[153] Although some of these events will be missed, many of them will still be recorded in patients' records alongside their record for death. The inclusion of primary care records increases the likelihood of the correct recording of an out of hospital CVD event resulting in death, but the best way to overcome such problems in the UK would be to use the Office of National Statistics mortality data, which includes information on cause of death. These data were not utilised in this thesis, but would be advantageous to any future work.
- This thesis explored a range of clinically specific CVD outcomes, which had varying degrees of severity, and as the study populations included older women, it is likely that those included had several comorbidities. It plausible that less severe CVD outcomes may not be recorded or recognised by clinicians if the woman also has any other, more severe, comorbidities. Women that initiated AIs in both the UK and US study had more comorbidities such as RA, CKD, and diabetes in comparison with tamoxifen initiators. If less CVD events were recorded in the women exposed to AI, then it is possible that there is an underestimation of effect size when comparing AI to tamoxifen use. In the US study, women unexposed to any endocrine therapy generally also had more comorbidities than those that initiated tamoxifen or AIs, which could again have led to underestimation of effect sizes.

Covariates

- Patients are not required to update their lifestyle choice records every time they visit a GP surgery in the UK. For example, a woman's latest record could indicate they currently smoke, but their cancer diagnosis caused them to stop. If they have not been questioned as to their change in habit, their records would not reflect this change. However, given women diagnosed with breast cancer are likely to have regular contact with the GP following their diagnosis, the likelihood of full and up to date lifestyle measures is high, which is also reflected in the low number of missing data within these variables.
- Prescriptions of drugs other than endocrine therapy were identified using the same method as the exposure (a record of a prescription in the patients' files based on a pre-specified list of codes). Therefore, there is still a risk of women being prescribed, but never actually dispensing and taking the drug. However, given that all women were then diagnosed with breast cancer, it is likely that they reviewed their medication with a GP and were encouraged to take any drugs that they had ben indicated.
- Diagnoses of the potential comorbidities diabetes, CKD, and rheumatoid arthritis were identified using only the CPRD in the UK study. It is therefore possible for patients to be misclassified as not being diagnosed with the comorbidity if they had only been diagnosed in secondary care. However, all comorbidities in this thesis were chronic and of sufficient significance that they are likely recorded in the primary care record, regardless of if they were initially diagnosed in secondary care. Such misclassification would be more problematic for acute events.
- The UK and US health systems are set up in very different ways. For example, the Quality and Outcomes Framework financially incentivises all GPs in the NHS to diagnose, give drugs or take readings based on certain recommendations (such as BP recordings in all patients over the age of 45).[154] This does not happen in the US due to the more fragmented nature of the healthcare system. Any difference in recording is, however, likely to be non-differential in relation to exposure to endocrine therapy.

8.5.2.2 Missing data

- Routinely collected health records can be prone to missing data. A complete case analysis, in which crude and adjusted analyses only included women without missing data was carried

out in both the UK and US studies. This approach is valid in a regression context if missingness is conditionally independent of the outcome.[155] In the context of the UK study, this means that the assumption was made of no association between having complete data on BMI (5.5% missing), smoking (0.5% missing), alcohol use (6.3% missing), and blood pressure (0.2% missing for systolic and diastolic) and developing any of the CVD outcomes given fixed covariate values. Although this is an untestable assumption, it is more plausible than the missing at random assumption required for multiple imputation, as recording of lifestyle variables may be associated with the value of the variable itself (those with high or low BMI are more likely to have their BMI recorded). In comparison, the variables with missingness in the US study were region (0.6% missing), ethnicity (0.2% missing), and grade of breast cancer (4.3% missing). Although it is plausible that these variables could be missing at random, meaning multiple imputation would be a valid approach; missingness is likely independent of the CVD outcomes, so a complete case analysis is also valid.

- Diagnoses in EHR studies are defined by either the presence or absence of a clinical code. Therefore, not having a diagnosis is defined by an absence of evidence of disease, rather than an assessment of disease that indicates the person does not have a certain diagnosis. However, it is possible that those with absence of a clinical code were assessed for the disease, but had missing data in relation to the diagnosis. Due to this method of identifying diagnoses, there is no way to identify true missingness.

8.5.2.3 Confounding

- The main potential source of confounding in the US study was likely due to reasons for non-initiation of endocrine therapy. If women did not initiate either tamoxifen or AIs due to reasons that are also associated with the risk of a CVD event (BMI, smoking, alcohol use, financial barriers, access to care), then it is likely that non-initiators are at an increased risk of an event at baseline compared with those that initiate therapy. If this is true, then any protective effect of either endocrine therapy could be over-estimated. Protective effects were generally small when comparing the association between AI use and non-venous CVDs, so residual confounding due to reasons for non-initiation could explain the effects. The measures outlined above that could potentially influence non-initiation were not available in SEER-Medicare data, so it was not possible to explore reasons for non-initiation.

- The increased risk of VTE events associated with tamoxifen use, which was observed in both the UK and US studies, has been reported in other studies and is well known to clinicians. If this known association also meant women at a high risk of CVDs other than VTE were more likely to be prescribed AIs instead of tamoxifen, then the AI users would have higher risk of non-venous CVD at therapy initiation. This could explain some of the increased risk of non-venous CVD outcomes in AI compared with tamoxifen users observed in both studies. However, adjustments were made for many CVD risk factors in both studies; in the UK study it was also possible to further include lifestyle variables that were not available in the US.
- Residual confounding could have also originated from the categorisation of covariates. All prior diagnoses (chronic kidney disease, rheumatoid arthritis, diabetes) were included as binary variables, but the disease severity could differ greatly within those diagnosed. For example, those with mild diabetes have a very different biomarkers and treatment regimens to those with severe diabetes, causing residual confounding within this group of patients. This mechanism of residual confounding is also applicable to all other treatment and lifestyle covariates.
- All adjustments made in analyses are only as good as the data from which they originated. So, any misclassification described above that may lead to measurement error, has the potential to cause residual confounding.
- Hypertension (or blood pressure) was included as a potential confounder. It is however possible that changes in cholesterol levels due to endocrine therapy use could cause hypertension, which is known to contribute to CVD prevalence. If this is the case, then hypertension could be a partial mediator of any association between endocrine therapy and CVD. Adjusting for a potential mediator could therefore bias any reported associations. A possible way to explore this would be to include hypertension as an outcome, but the study team decided not to do so as I wanted to explore clear clinical cardiovascular end points rather than disease that contributes to and causes these outcomes. A limitation of this approach is that by not including hypertension as an outcome, some adverse CVD consequences of therapy that had not yet led to harder CVD diagnoses may be missed.

8.5.2.4 Multiple testing

- As a large range of outcomes are collected in routinely collected EHR data, there can be a tendency to test for a wide range of associations within one study population, which could

potentially lead to an increased risk of false positive results. Study populations for each individual CVD analysis are not identical in both the UK and US studies as they exclude women with the specific CVD outcome of interest prior to index. Caution should therefore be taken when interpreting 95% CIs that are close to the null association, due to problems associated with multiple testing. However, there were consistent patterns of results for the association between endocrine therapies and non-venous CVD outcomes, so for the results of non-venous CVD outcomes where the 95% CI was close to the null association, it unlikely that chance is the full explanation. A way to correct for multiple testing is to apply the Bonferroni correction.[156] However, this correction is usually over-conservative and only changes the significance cut off. This would also place a lot of emphasis on arbitrary p values to assess statistical significance. Furthermore, Bonferroni correction should only be considered if: (1) a single test of the 'universal null hypothesis' that all tests are not significant is required, (2) it is imperative to avoid a type I error, and (3) a large number of tests are carried out without pre-planned hypotheses.[157] Although type I errors should be avoided, requirements 1 and 3 are not satisfied in this context as I was not exploring the possibility that all tests were simultaneously "non-significant" and all association assessed had pre-planned hypotheses.

8.5.2.5 Statistical power

- Although EHR databases are usually very large in comparison with traditional cohort studies, lack of statistical power to detect small associations is still sometimes problematic. In survival analysis, the number of events has a high impact on statistical power. The number of events in some analyses assessing the risk of relatively rare CVD outcomes such as pericarditis and SCA was small in the study populations of female breast cancer survivors.

8.5.2.6 Generalisability

- The study populations of the two studies within this thesis originate from countries that have a high incidence of both breast cancer and CVD in comparison with many other countries in the world. Although these results are likely generalisable to women in other high-income countries, further understanding of the associations are needed in low and middle-income settings.

8.6 BRADFORD-HILL CRITERIA FOR CAUSALITY

The implicit comparison for the question of causality is each endocrine therapy (tamoxifen and AIs) vs no endocrine therapy, which was explored in the US study, because the comparison between AI and tamoxifen is secondary to the causal effects of each drug vs no endocrine therapy.

8.6.1 Does tamoxifen causally affect the risk of CVD outcomes?

The strength of association between tamoxifen and non-venous CVD outcomes varied based on the specific outcome. Some outcomes, such as MI, showed evidence of a very strong protective effect (ever tamoxifen vs no endocrine therapy US HR: 0.44, 95% CI: 0.30-0.63), whereas others, such as PVD, showed weaker evidence of a protective effect (ever tamoxifen vs no endocrine therapy US HR: 0.91, 95% CI: 0.75-1.10). Furthermore, consistency of effects has been discussed in detail in section 8.3, and for the outcomes in which there is a considerable amount of previous literature, results are generally consistent with what has been found in this thesis. Temporality is assumed for the associations between tamoxifen and non-venous CVD outcomes (and all other exposure outcome relationships in this thesis) due to the prospective nature of the cohort studies that have been designed, leading to little risk of reverse causality. Biological gradient was not explicitly tested (do increasing doses of tamoxifen give further non-venous CVD protection? Similar for all other exposure outcome associations) because of limited power, but within patients diagnosed with ER+/PR+ breast cancer, the dose of tamoxifen will likely not be increased due to its known efficacy on reducing the risk of breast cancer recurrence. However, the effect of increasing length of tamoxifen prescription on the risk of non-venous CVD outcomes was explored, and it was generally found that there was no further protection as length of prescription increased. Finally, there is a biological explanation of a protective effect of tamoxifen on non-venous CVD outcomes through lowering of cholesterol (full explanation in section 8.4).

The strength of association between tamoxifen and venous CVD outcomes is reasonably large, with evidence of an increased risk of VTE associated with ever tamoxifen use (ever tamoxifen vs no endocrine therapy US HR: 1.39, 95% CI: 0.98-1.98) which is consistent with most previous RCTs and observational studies (fully outlined in chapter 2 and section 8.3). There was, however, little evidence of an increasing risk of VTE as length of prescription increased, but this could be due to limited power. Finally, there is biological plausibility of this association as tamoxifen has oestrogen agonistic effects that could increase the risk of in thrombogenicity through a reduction in

antithrombin and protein C levels.

8.6.2 Do AIs causally affect the risk of CVD outcomes?

The strength of association between AIs and non-venous CVD outcomes was considerably weaker in comparison with the association between tamoxifen and non-venous CVD outcomes. Some outcomes, such as pericarditis, showed evidence of a reasonably strong protective effect (ever AI vs no endocrine therapy US HR: 0.67, 95% CI: 0.50-0.90), whereas others, such as PVD, showed no evidence of an effect (ever AI vs no endocrine therapy US HR: 1,00, 95% CI: 0.87-1.14). Relatively few previous studies have explored the effect of AIs on the risk of non-venous CVD outcomes, and for those that have, there have been mixed results, with some showing small protective effects and others showing no association. Again, there was little evidence that increasing length of prescription affected the protective effects of AIs. Finally, there is no biological reasoning behind a protective effect of AIs on non-venous CVD outcomes, but instead a plausible mechanism that suggests an increased risk of these outcomes with AI use as they deplete oestrogen levels, which are known to be protective of CVD.

There is little evidence of an association between AI use and venous CVD outcomes in this thesis and only one previous RCT reporting an increased risk of VTE in AI users. There is also no evidence of a dose-response relationship, but it is there is a possible biological mechanism to an increased risk of VTE in AI users through depletion of oestrogen levels.

8.7 IMPLICATIONS FOR CLINICAL PRACTICE AND FUTURE WORK

- This thesis has shown that the hypothesised mechanism for an increased risk of CVD events with AI use does not appear to translate into an increased risk of CVD events associated with AI use, and any difference in CVD risk in AI users compared with tamoxifen users is likely to be predominantly driven by the cardiac effects of tamoxifen (be that a possibly protective of some non-venous CVDs, or toxic of venous CVDs). Therefore the call for clinicians to prioritise assessment and reduction of CVD risk in AI users may not be as important as was previously thought.[158]
- Descriptive analyses reported a marked yearly decrease in the proportion of women initiating tamoxifen, coupled with an increase in women initiating AIs after a breast cancer

diagnosis in both the UK and US studies. This is likely due to the evidence that AIs increase disease free and overall survival in comparison with tamoxifen in post-menopausal women diagnosed with ER+ breast cancer.[159] Although this change will likely mean less VTE events in post-menopausal survivors of ER+/PR+ breast cancer, it is unclear how the loss of an apparently protective effect of some non-venous CVD associated with tamoxifen use will affect the population as a whole. Clinicians should therefore be wary of this possible loss in protection of any women at high risk of CVD who are switching form tamoxifen to AIs.

- As tamoxifen use is becoming less prevalent in post-menopausal women diagnosed with ER+/PR+ breast cancer, and older women diagnosed with breast cancer are dying from causes other than the cancer itself, it would be informative to know if the cardio-protective effects of tamoxifen outweigh the excess reduction in risk of breast cancer recurrence associated with AI use compared with tamoxifen use. A meta-analysis of RCTs showed that there was no difference in overall mortality between those given tamoxifen compared AIs for 5 years, and further understanding of the CVD and breast cancer specific mortality associated with both tamoxifen and AI use is needed.[22] If the long term CVD benefits of tamoxifen outweigh the excess protection effect of breast cancer reoccurrence with AI use, it is conceivable that post-menopausal women diagnosed with ER+/PR+ breast cancer who are at high risk of CVD may receive more benefit from taking tamoxifen rather than AIs.
- The effect of AI use and CVD risk should be explored further. This includes further understanding of the viability of a plausible biological mechanism for a reduced risk of certain non-venous CVD events associated with AI use, and if these reported associations are causal or a result of biases. This could be progressed through the following questions:
    - Are the CVD effects of AI use the same within all types of third generation AIs (anastrozole, letrozole, and exemestane)?
    - What are the risk factors of non-initiation of any endocrine therapy in women diagnosed with ER+/PR+ breast cancer and how similar are these to the risk factors of CVD?
    - Does AI use effect the risk of CVD risk factors such as cholesterol level and blood pressure?
- The studies within this thesis suggested that the individual effect of both tamoxifen and AIs on the risk of stroke are similar to the effect on other non-venous CVDs. However, due to conflicting results in previous research, there is still uncertainty on the association between

both endocrine therapies individually and the risk of stroke. Further large population-based studies that focus on the effect of these drugs on the risk of stroke (overall, and both haemorrhagic and ischaemic stroke individually) would allow a deeper understanding of these associations.

- There was little statistical power to detect an association for several specific CVD outcomes although the broad populations from which the study populations originated were large. It may be possible to further explore some of the outcomes for which there was statistical little power if there was a possibility for collaborative international work combining data from several sources. Greater statistical power would also mean that effect modification and the cumulative effect of treatment could also be assessed in more detail.
- A data source that allows for adjustment of all potential confounders would enable potential residual confounding problems to be addressed. A potential dataset that would address the current problem of being unable to adjust for other cancer therapies in the UK data would be the CPRD and HES linked to the Systemic Anti-Cancer Therapy (SACT) dataset, which has recently become available to researchers. SACT collects information routinely reported by NHS trusts on the treatment of malignant disease in secondary care in England. Submission of SACT data was only mandatory from 2014, and these data have only recently been linked to the CPRD. It was therefore not possible to use these data for the purpose of analysis in this thesis, but may be possible in the future once time since mandatory follow up increases.

8.8 CONCLUSIONS

This thesis focused on the effect of endocrine therapy use on the risk of a range of CVD outcomes. Based on the totality of information presented from the studies in this thesis, along with previous RCT and observational studies, there is convincing evidence of an increased risk of several non-venous CVDs in AI compared with to tamoxifen users in post-menopausal women diagnosed with ER+/PR+ breast cancer, with varying effect sizes for different clinically specific outcomes. However, the associations are likely due to protective effects of tamoxifen on these outcomes, rather than any toxic effects of AIs. There was also evidence of AI use being associated with a small decreased risk of several non-venous CVD events, but a causal effect cannot be assumed due to no current understanding of a biological mechanism, and a possible overestimation of any protective effect due to residual confounding. The studies within this thesis also confirmed the known effect of tamoxifen use (compared with both AI use and non-use of any endocrine therapy) on increased risk of VTE,

which was previously evident from the accumulation of evidence in the systematic review. No evidence of effect was observed between AI use and the risk of VTE. Overall, the effects of tamoxifen use on the risk of CVD are becoming clearer. However, as more postmenopausal women are prescribed AIs rather than tamoxifen, additional large-scale population based studies are needed to understand if any observed cardio-protective effects of AI are causal, as well as further understanding the risk-benefit balance of endocrine therapies with respect to both cancer and cardiovascular outcomes.

8.9 CONTRIBUTIONS

The following section outlines my contributions to each chapter of the thesis

- Chapter 1 – All work was carried out by me.
- Chapter 2 – I did the initial literature search, as well as the paper selection and data extraction. This process was replicated by two co-authors. I wrote the first draft, then all authors contributed to further drafts and approved the final manuscript.
- Chapter 3 – All work was carried out by me.
- Chapter 4 – All work was carried out by me, which included data extraction, data manipulation, and exploratory analyses.
- Chapter 5 – All work was carried out by me, supported by advisors, which included ethics approval, approval of protocol from ISAC, data extraction, data manipulation, and analyses.
- Chapter 6 – I led on the study design and protocol development, supported by advisors in the UK and US. Sharon Peacock Hinton, an employee at UNC, then produced the specified dataset from the raw data, with guidance from myself. I then carried out the data manipulation and analyses.
- Chapter 7 – All work was carried out by me.
- Chapter 8 – All work was carried out by me.

REFERENCES

1. Harris, J., et al., *Diseases of the Breast: fifth edition*. 2014, Wolters Kluwer Health.
2. National Cancer Institute, *Cancer staging.* National Cancer Institute: Diagnosis and Staging, available at: https://www.cancer.gov/about-cancer/diagnosis-staging/staging, accessed on 1/10/2018, 2015.
3. National Cancer Institute, *Cancer prognosis.* National Cancer Institute: Diagnosis and Staging, available at: https://www.cancer.gov/about-cancer/diagnosis-staging/prognosis, accessed on 1/10/2018, 2015.
4. Cancer Genome Atlas, N., *Comprehensive molecular portraits of human breast tumours.* Nature, 2012. **490**(7418): p. 61-70.
5. Dieci, M.V., et al., *Rare breast cancer subtypes: histological, molecular, and clinical peculiarities.* Oncologist, 2014. **19**(8): p. 805-13.
6. Eusebi, V., et al., *Long-term follow-up of in situ carcinoma of the breast.* Semin Diagn Pathol, 1994. **11**(3): p. 223-35.
7. Howlader, N., et al., *US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status.* J Natl Cancer Inst, 2014. **106**(5).
8. Blows, F.M., et al., *Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies.* PLoS Med, 2010. **7**(5): p. e1000279.
9. Haque, R., et al., *Impact of breast cancer subtypes and treatment on survival: an analysis spanning two decades.* Cancer Epidemiol Biomarkers Prev, 2012. **21**(10): p. 1848-55.
10. Bianchini, G., et al., *Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease.* Nat Rev Clin Oncol, 2016. **13**(11): p. 674-690.
11. London Cancer Alliance, *LCA Breast Cancer Clinical Guidelines.* LCA guidelines, 2016.
12. National Institute for Health and Care Excellence, *Early and locally advanced breast cancer: diagnosis and management, available at: https://www.nice.org.uk/guidance/ng101/chapter/Recommendations#ftn.footnote_5, accessed on: 15/11/2018.* 2018.
13. Early Breast Cancer Trialists' Collaborative, G., *Long-term outcomes for neoadjuvant versus adjuvant chemotherapy in early breast cancer: meta-analysis of individual patient data from ten randomised trials.* Lancet Oncol, 2018. **19**(1): p. 27-39.
14. Mauri, D., N. Pavlidis, and J.P. Ioannidis, *Neoadjuvant versus adjuvant systemic treatment in breast cancer: a meta-analysis.* J Natl Cancer Inst, 2005. **97**(3): p. 188-94.
15. McGale, P., et al., *Effect of radiotherapy after mastectomy and axillary surgery on 10-year recurrence and 20-year breast cancer mortality: meta-analysis of individual patient data for 8135 women in 22 randomised trials.* Lancet, 2014. **383**(9935): p. 2127-35.
16. National Institute for Health and Care Excellence, *Early and locally advanced breast cancer: diagnosis and management.* Available at: https://www.nice.org.uk/guidance/ng101, 2018.
17. Yang, L.H., et al., *Survival benefit of tamoxifen in estrogen receptor-negative and progesterone receptor-positive low grade breast cancer patients.* J Breast Cancer, 2012. **15**(3): p. 288-95.
18. Early Breast Cancer Trialists' Collaborative, G., et al., *Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials.* Lancet, 2011. **378**(9793): p. 771-84.
19. Al-Mubarak, M., et al., *Extended adjuvant tamoxifen for early breast cancer: a meta-analysis.* PLoS One, 2014. **9**(2): p. e88238.

20. Lonning, P.E., *Pharmacology of new aromatase inhibitors.* Breast 1996. **2**: p. 202-208.

21. National Institute for Health and Care Excellence. *Hormonal therapies for the adjuvant treatment of early oestrogen-receptor-positive breast cancer* 2006 14/04/2016]; Available from: https://www.nice.org.uk/guidance/TA112/documents/final-appraisal-determination2.

22. Dowsett, M., et al., *Meta-analysis of breast cancer outcomes in adjuvant trials of aromatase inhibitors versus tamoxifen.* J Clin Oncol, 2010. **28**(3): p. 509-18.

23. Goldvaser, H., et al., *Efficacy of extended adjuvant therapy with aromatase inhibitors in early breast cancer among common clinicopathologically-defined subgroups: A systematic review and meta-analysis.* Cancer Treat Rev, 2017. **60**: p. 53-59.

24. Ferlay, J.S., I. Ervik, M. Dikshit, R. Eser, S. Mathers, C. Rebelo, M. Parkin, DM. Forman, D. Bray, F. *GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11, available at http://globocan.iarc.fr/, accessed on 01/10/2018*. 2013.

25. Cancer Research UK. *Breast cancer survival statistics*. 2014 [cited 2018 01/10].

26. American Cancer Society, *Cancer Facts & Figures 2018*, A.C. Society, Editor. 2018: Atlanta.

27. Brown, K.F., et al., *The fraction of cancer attributable to modifiable risk factors in England, Wales, Scotland, Northern Ireland, and the United Kingdom in 2015.* Br J Cancer, 2018. **118**(8): p. 1130-1141.

28. Bhaskaran, K., et al., *Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5.24 million UK adults.* Lancet, 2014. **384**(9945): p. 755-65.

29. Hamajima, N., et al., *Alcohol, tobacco and breast cancer--collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease.* Br J Cancer, 2002. **87**(11): p. 1234-45.

30. Gaudet, M.M., et al., *Active smoking and breast cancer risk: original cohort data and meta-analysis.* J Natl Cancer Inst, 2013. **105**(8): p. 515-25.

31. Collaborative Group on Hormonal Factors in Breast, C., *Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies.* Lancet, 1996. **347**(9017): p. 1713-27.

32. Chowdhury, R., et al., *Breastfeeding and maternal health outcomes: a systematic review and meta-analysis.* Acta Paediatr, 2015. **104**(467): p. 96-113.

33. Collaborative Group on Hormonal Factors in Breast, C., *Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease.* Lancet, 2001. **358**(9291): p. 1389-99.

34. Antoniou, A., et al., *Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies.* Am J Hum Genet, 2003. **72**(5): p. 1117-30.

35. Lambertini, M., et al., *Reproductive behaviors and risk of developing breast cancer according to tumor subtype: A systematic review and meta-analysis of epidemiological studies.* Cancer Treat Rev, 2016. **49**: p. 65-76.

36. Wachtel, T.J., *Geriatric Clinical Advisor*. Vol. 1. 2017: Mosby.

37. S.J. Enna, D.B.B., *xPharm: The Comprehensive Pharmacology Reference*. 2008: Elsevier.

38. National Institute for Health and Care Excellence, *Chest pain of recent onset: assessment and diagnosis, available at: https://www.nice.org.uk/guidance/cg95/chapter/Recommendations, accessed on: 16/11/2018.* 2016.

39. National institute for Health and Care Excellence, *Unstable angina and NSTEMI: early management, available at: https://www.nice.org.uk/guidance/cg94, accessed on: 16/11/2018.* 2013.

40. Thygesen, K., et al., *Fourth universal definition of myocardial infarction (2018).* Eur Heart J, 2018.

41. Task Force on Myocardial Revascularization of the European Society of, C., et al., *Guidelines on myocardial revascularization.* Eur J Cardiothorac Surg, 2010. **38 Suppl**: p. S1-S52.

42. Ilkhanoff, L. and J.J. Goldberger, *Out-of-hospital cardiac arrest: getting beyond the tip of the iceberg.* Circulation, 2012. **126**(7): p. 793-6.

43. Rubart, M. and D.P. Zipes, *Mechanisms of sudden cardiac death.* J Clin Invest, 2005. **115**(9): p. 2305-15.

44. Morley, R.L., et al., *Peripheral artery disease.* BMJ, 2018. **360**: p. j5842.

45. Aboyans, V., et al., *2017 ESC Guidelines on the Diagnosis and Treatment of Peripheral Arterial Diseases, in collaboration with the European Society for Vascular Surgery (ESVS).* Rev Esp Cardiol (Engl Ed), 2018. **71**(2): p. 111.

46. Sacco, R.L., et al., *An updated definition of stroke for the 21st century: a statement for healthcare professionals from the American Heart Association/American Stroke Association.* Stroke, 2013. **44**(7): p. 2064-89.

47. Tse, G., *Mechanisms of cardiac arrhythmias.* J Arrhythm, 2016. **32**(2): p. 75-81.

48. Federmann, M. and O.M. Hess, *Differentiation between systolic and diastolic dysfunction.* Eur Heart J, 1994. **15 Suppl D**: p. 2-6.

49. Ponikowski, P., et al., *2016 ESC Guidelines for the Diagnosis and Treatment of Acute and Chronic Heart Failure.* Rev Esp Cardiol (Engl Ed), 2016. **69**(12): p. 1167.

50. Adler, Y., et al., *2015 ESC Guidelines for the diagnosis and management of pericardial diseases: The Task Force for the Diagnosis and Management of Pericardial Diseases of the European Society of Cardiology (ESC)Endorsed by: The European Association for Cardio-Thoracic Surgery (EACTS).* Eur Heart J, 2015. **36**(42): p. 2921-2964.

51. Zeng, Y.I., et al., *Pathophysiology of valvular heart disease.* Exp Ther Med, 2016. **11**(4): p. 1184-1188.

52. Baumgartner, H., et al., *2017 ESC/EACTS Guidelines for the Management of Valvular Heart Disease.* Rev Esp Cardiol (Engl Ed), 2018. **71**(2): p. 110.

53. Nicolaides, A.N., et al., *The origin of deep vein thrombosis: a venographic study.* Br J Radiol, 1971. **44**(525): p. 653-63.

54. Mazzolai, L., et al., *Diagnosis and management of acute deep vein thrombosis: a joint consensus document from the European society of cardiology working groups of aorta and peripheral circulation and pulmonary circulation and right ventricular function.* Eur Heart J, 2017.

55. Lapner, S.T. and C. Kearon, *Diagnosis and management of pulmonary embolism.* BMJ, 2013. **346**: p. f757.

56. Riedel, M., *Pulmonary embolic disease, In: Gibson GJ, Geddes DM, Costabel U, et al, eds. Respiratory medicine*. 2003, London: Saunders.

57. Konstantinides, S.V., *2014 ESC Guidelines on the diagnosis and management of acute pulmonary embolism.* Eur Heart J, 2014. **35**(45): p. 3145-6.

58. Brown, N.J. and D.E. Vaughan, *Angiotensin-converting enzyme inhibitors.* Circulation, 1998. **97**(14): p. 1411-20.

59. Endo, A., *The discovery and development of HMG-CoA reductase inhibitors.* J Lipid Res, 1992. **33**(11): p. 1569-82.

60. Burnier, M. and H.R. Brunner, *Angiotensin II receptor antagonists.* Lancet, 2000. **355**(9204): p. 637-45.

61. Frishman, W., *Current Cardiovascular Drugs, 4th Edition*. 2005: Current Medicine Group.

62. Katz, A.M., *Pharmacology and mechanisms of action of calcium-channel blockers.* J Clin Hypertens, 1986. **2**(3 Suppl): p. 28S-37S.

63. World Health Organisation. *Cardiovascular Diseases, available at: http://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds), accessed on 4/10/2018*. WHO 2017.

64. British Heart Foundation, *Heart Statistics, available at: https://www.bhf.org.uk/what-we-do/our-research/heart-statistics, accessed on 4/10/2018.* 2018.

65. Benjamin, E.J., et al., *Heart Disease and Stroke Statistics-2018 Update: A Report From the American Heart Association.* Circulation, 2018. **137**(12): p. e67-e492.

66. Bhatnagar, P., et al., *Trends in the epidemiology of cardiovascular disease in the UK.* Heart, 2016. **102**(24): p. 1945-1952.

67. Ezzati, M., et al., *Estimates of global and regional potential health gains from reducing multiple major risk factors.* Lancet, 2003. **362**(9380): p. 271-80.

68. Shah, A.D., et al., *Type 2 diabetes and incidence of cardiovascular diseases: a cohort study in 1.9 million people.* Lancet Diabetes Endocrinol, 2015. **3**(2): p. 105-13.

69. Lloyd-Jones, D.M., et al., *Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring.* JAMA, 2004. **291**(18): p. 2204-11.

70. Chapman, J.A., et al., *Competing causes of death from a randomized trial of extended adjuvant endocrine therapy for breast cancer.* Journal of the National Cancer Institute, 2008. **100**(4): p. 252-60.

71. Hanrahan, E.O., et al., *Overall survival and cause-specific mortality of patients with stage T1a,bN0M0 breast carcinoma.* J Clin Oncol, 2007. **25**(31): p. 4952-60.

72. Bradshaw, P.T., et al., *Cardiovascular Disease Mortality Among Breast Cancer Survivors.* Epidemiology, 2016. **27**(1): p. 6-13.

73. Rasmussen-Torvik, L.J., et al., *Ideal cardiovascular health is inversely associated with incident cancer: the Atherosclerosis Risk In Communities study.* Circulation, 2013. **127**(12): p. 1270-5.

74. Mehta, L.S., et al., *Cardiovascular Disease and Breast Cancer: Where These Entities Intersect: A Scientific Statement From the American Heart Association.* Circulation, 2018. **137**(8): p. e30-e66.

75. Swain, S.M., F.S. Whaley, and M.S. Ewer, *Congestive heart failure in patients treated with doxorubicin: a retrospective analysis of three trials.* Cancer, 2003. **97**(11): p. 2869-79.

76. Hochster, H., C. Wasserheit, and J. Speyer, *Cardiotoxicity and cardioprotection during chemotherapy.* Current Opinion in Oncology, 1995. **7**(4): p. 304-9.

77. Guglin, M., et al., *Introducing a new entity: chemotherapy-induced arrhythmia.* Europace, 2009. **11**(12): p. 1579-86.

78. Rowinsky, E.K., et al., *Clinical toxicities encountered with paclitaxel (Taxol).* Semin Oncol, 1993. **20**(4 Suppl 3): p. 1-15.

79. Meyer, C.C., et al., *Symptomatic cardiotoxicity associated with 5-fluorouracil.* Pharmacotherapy, 1997. **17**(4): p. 729-36.

80. Van Cutsem, E., et al., *Incidence of cardiotoxicity with the oral fluoropyrimidine capecitabine is typical of that reported with 5-fluorouracil.* Ann Oncol, 2002. **13**(3): p. 484-5.

81. Kosmas, C., et al., *Cardiotoxicity of fluoropyrimidines in different schedules of administration: a prospective study.* Journal of Cancer Research & Clinical Oncology, 2008. **134**(1): p. 75-82.

82. Layoun, M.E., et al., *Fluoropyrimidine-Induced Cardiotoxicity: Manifestations, Mechanisms, and Management.* Curr Oncol Rep, 2016. **18**(6): p. 35.

83. Ewer, M.S., et al., *Reversibility of trastuzumab-related cardiotoxicity: new insights based on clinical course and response to medical treatment.* Journal of Clinical Oncology, 2005. **23**(31): p. 7820-6.

84. Clarke, M., et al., *Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence and 15-year survival: an overview of the randomised trials.* Lancet, 2005. **366**(9503): p. 2087-106.

85. Henson, K.E., et al., *Radiation-related mortality from heart disease and lung cancer more than 20 years after radiotherapy for breast cancer.* Br J Cancer, 2013. **108**(1): p. 179-82.

86. Ryden, L., et al., *Aromatase inhibitors alone or sequentially combined with tamoxifen in postmenopausal early breast cancer compared with tamoxifen or placebo - Meta-analyses on efficacy and adverse events based on randomized clinical trials.* Breast, 2016. **26**: p. 106-14.

87. Khosrow-Khavar, F., et al., *Cardiotoxicity of aromatase inhibitors and tamoxifen in postmenopausal women with breast cancer: a systematic review and meta-analysis of randomized controlled trials.* Ann Oncol, 2017. **28**(3): p. 487-496.

88. Aydiner, A., *Meta-analysis of breast cancer outcome and toxicity in adjuvant trials of aromatase inhibitors in postmenopausal women.* Breast, 2013. **22**(2): p. 121-9.

89. Amir, E., et al., *Toxicity of adjuvant endocrine therapy in postmenopausal breast cancer patients: a systematic review and meta-analysis.* Journal of the National Cancer Institute, 2011. **103**(17): p. 1299-309.

90. Cuppone, F., et al., *Do adjuvant aromatase inhibitors increase the cardiovascular risk in postmenopausal women with early breast cancer? Meta-analysis of randomized trials.* Cancer, 2008. **112**(2): p. 260-7.

91. Braithwaite, R.S., et al., *Meta-analysis of vascular and neoplastic events associated with tamoxifen.* J Gen Intern Med, 2003. **18**(11): p. 937-47.

92. Dewar, J.A., et al., *Long term effects of tamoxifen on blood lipid values in breast cancer.* BMJ, 1992. **305**(6847): p. 225-6.

93. Esteva, F.J. and G.N. Hortobagyi, *Comparative assessment of lipid effects of endocrine therapy for breast cancer: implications for cardiovascular disease prevention in postmenopausal women.* Breast, 2006. **15**(3): p. 301-12.

94. Grey, A.B., et al., *The effect of the anti-estrogen tamoxifen on cardiovascular risk factors in normal postmenopausal women.* Journal of Clinical Endocrinology & Metabolism, 1995. **80**(11): p. 3191-5.

95. Morales, M., et al., *Effects of tamoxifen on serum lipid and apolipoprotein levels in postmenopausal patients with breast cancer.* Breast Cancer Res Treat, 1996. **40**(3): p. 265-70.

96. Love, R.R., T.S. Surawicz, and E.C. Williams, *Antithrombin III level, fibrinogen level, and platelet count changes with adjuvant tamoxifen therapy.* Archives of Internal Medicine, 1992. **152**(2): p. 317-20.

97. Jordan, V.C., N.F. Fritz, and D.C. Tormey, *Long-term adjuvant therapy with tamoxifen: effects on sex hormone binding globulin and antithrombin III.* Cancer Research, 1987. **47**(16): p. 4517-9.

98. Herrett, E., et al., *Data Resource Profile: Clinical Practice Research Datalink (CPRD).* Int J Epidemiol, 2015. **44**(3): p. 827-36.

99. Campbell, J.D., D.J. Eaton, S.C. Gallagher, A.M. Williams, T.J., *Is the CPRD GOLD Population Comparable to the U.K. Population?* Pharmacoepidemiology and Drug Safety, 2013. **22**(Suppl 1): p. 280.

100. Herrett, E., et al., *Validation and validity of diagnoses in the General Practice Research Database: a systematic review.* Br J Clin Pharmacol, 2010. **69**(1): p. 4-14.

101. Boggon, R., et al., *Cancer recording and mortality in the General Practice Research Database and linked cancer registries.* Pharmacoepidemiol Drug Saf, 2013. **22**(2): p. 168-75.

102. Booth, H.P., A.T. Prevost, and M.C. Gulliford, *Validity of smoking prevalence estimates from primary care electronic health records compared with national population survey data for England, 2007 to 2011.* Pharmacoepidemiol Drug Saf, 2013. **22**(12): p. 1357-61.

103. Bhaskaran, K., et al., *Representativeness and optimal use of body mass index (BMI) in the UK Clinical Practice Research Datalink (CPRD).* BMJ Open, 2013. **3**(9): p. e003389.

104. Clinical Practice Research Datalink. *CPRD Linked Data, available at: https://www.cprd.com/linked-data#HES%20Admitted%20Patient%20Care%20data, accessed on 17/10/2018*. 2018.

105. Clinical Practice Research Datalink, *Hospital Episode Statistics (HES) Admitted Patient Care and CPRD primary care data documentation (Set 16)*, Medicines and Healthcare products Regulatory Agency, Editor. 2018.

106. Lewis, J.D., et al., *Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research.* Pharmacoepidemiol Drug Saf, 2007. **16**(4): p. 393-401.

107. Swerdlow, A.J., et al., *The Breakthrough Generations Study: design of a long-term UK cohort study to investigate breast cancer aetiology.* Br J Cancer, 2011. **105**(7): p. 911-7.

108. National Cancer Institute. *Overview of the SEER program*. 30/8/2018].

109. Nattinger, A.B., T.L. McAuliffe, and M.M. Schapira, *Generalizability of the surveillance, epidemiology, and end results registry population: factors relevant to epidemiologic and health care research.* J Clin Epidemiol, 1997. **50**(8): p. 939-45.

110. Zippin, C., D. Lum, and B.F. Hankey, *Completeness of hospital cancer case reporting from the SEER Program of the National Cancer Institute.* Cancer, 1995. **76**(11): p. 2343-50.

111. Warren, J.L., et al., *Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population.* Med Care, 2002. **40**(8 Suppl): p. IV-3-18.

112. Engels, E.A., et al., *Use of surveillance, epidemiology, and end results-medicare data to conduct case-control studies of cancer among the US elderly.* Am J Epidemiol, 2011. **174**(7): p. 860-70.

113. Centers for Medicare and Medicaid Services. *Cost Report Audit & Reimbursement, available at: https://www.cms.gov/Medicare/Compliance-and-Audits/Part-A-Cost-Report-Audit-and-Reimbursement/Index.html, accessed on 18/10/2018*.

114. Kucharska-Newton, A.M., et al., *Identification of Heart Failure Events in Medicare Claims: The Atherosclerosis Risk in Communities (ARIC) Study.* J Card Fail, 2016. **22**(1): p. 48-55.

115. Hlatky, M.A., et al., *Use of Medicare data to identify coronary heart disease outcomes in the Women's Health Initiative.* Circ Cardiovasc Qual Outcomes, 2014. **7**(1): p. 157-62.

116. StataCorp, *Stata Statistical Software: Release 15*. 2017, StataCorp LLC: College Station, TX.

117. Denaxas, S.C., et al., *Data resource profile: cardiovascular disease research using linked bespoke studies and electronic health records (CALIBER).* Int J Epidemiol, 2012. **41**(6): p. 1625-38.

118. Mathur, R., et al., *Completeness and usability of ethnicity data in UK-based primary care and hospital databases.* J Public Health (Oxf), 2014. **36**(4): p. 684-92.

119. Castelli, W.P., *Epidemiology of coronary heart disease: the Framingham study.* Am J Med, 1984. **76**(2A): p. 4-12.

120. Organisation for Economic Cooperation and Develeoment (OECD), *Obesity Update, available at: https://www.oecd.org/els/health-systems/Obesity-Update-2017.pdf, accessed on 31/10/2018*. 2017.

121. Public Health England, C.R.U., *National Cancer Registration and Analysis Service Short Report: Chemotherapy, Radiotherapy and Surgical Tumour Resections in England: 2013-2014 (V2), available at: http://www.ncin.org.uk/publications/reports/, accessed on 29/10/2018*. 2017.

122. Dickstein, K., et al., *ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2008: the Task Force for the diagnosis and treatment of acute and chronic heart failure 2008 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association of the ESC (HFA) and endorsed by the European Society of Intensive Care Medicine (ESICM).* Eur J Heart Fail, 2008. **10**(10): p. 933-89.

123. Rapsomaniki, E., et al., *Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks, healthy life-years lost, and age-specific associations in 1.25 million people.* Lancet, 2014. **383**(9932): p. 1899-911.

124. Bliss, J.M., et al., *Disease-related outcomes with long-term follow-up: an updated analysis of the intergroup exemestane study.* J Clin Oncol, 2012. **30**(7): p. 709-17.

125. Bradbury, B.D., et al., *Tamoxifen-treated breast carcinoma patients and the risk of acute myocardial infarction and newly-diagnosed angina.* Cancer, 2005. **103**(6): p. 1114-21.

126. Hernandez, R.K., et al., *Tamoxifen treatment in Danish breast cancer patients and 5-year risk of arterial atherosclerotic events: a null association.* Cancer Epidemiology, Biomarkers & Prevention, 2008. **17**(9): p. 2509-11.

127. Goss, P.E., et al., *Randomized trial of letrozole following tamoxifen as extended adjuvant therapy in receptor-positive breast cancer: updated findings from NCIC CTG MA.17.* Journal of the National Cancer Institute, 2005. **97**(17): p. 1262-71.

128. Arimidex, T.A.o.i.C.T.G., et al., *Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 100-month analysis of the ATAC trial.* Lancet Oncology, 2008. **9**(1): p. 45-53.

129. Jakesz, R., et al., *Switching of postmenopausal women with endocrine-responsive early breast cancer to anastrozole after 2 years' adjuvant tamoxifen: combined results of ABCSG trial 8 and ARNO 95 trial.* Lancet, 2005. **366**(9484): p. 455-62.

130. Coombes, R.C., et al., *Survival and safety of exemestane versus tamoxifen after 2-3 years' tamoxifen treatment (Intergroup Exemestane Study): a randomised controlled trial.* Lancet, 2007. **369**(9561): p. 559-70.

131. Pagani, O., et al., *Adjuvant exemestane with ovarian suppression in premenopausal breast cancer.* N Engl J Med, 2014. **371**(2): p. 107-18.

132. Abdel-Qadir, H., et al., *The risk of myocardial infarction with aromatase inhibitors relative to tamoxifen in post-menopausal women with early stage breast cancer.* Eur J Cancer, 2016. **68**: p. 11-21.

133. McDonald, C.C., et al., *Cardiac and vascular morbidity in women receiving adjuvant tamoxifen for breast cancer in a randomised trial. The Scottish Cancer Trials Breast Group.* BMJ, 1995. **311**(7011): p. 977-80.

134. van de Velde, C.J., et al., *Adjuvant tamoxifen and exemestane in early breast cancer (TEAM): a randomised phase 3 trial.* Lancet, 2011. **377**(9762): p. 321-31.

135. Rutqvist, L.E. and A. Mattsson, *Cardiac and thromboembolic morbidity among postmenopausal women with early-stage breast cancer in a randomized trial of adjuvant*

tamoxifen. The Stockholm Breast Cancer Study Group. Journal of the National Cancer Institute, 1993. **85**(17): p. 1398-406.

136. Yang, T.L., et al., *Association of tamoxifen use and reduced cardiovascular events among asian females with breast cancer.* Circulation Journal, 2014. **78**(1): p. 135-40.

137. Ligibel, J.A., et al., *Risk of myocardial infarction, stroke, and fracture in a cohort of community-based breast cancer patients.* Breast Cancer Research & Treatment, 2012. **131**(2): p. 589-97.

138. Geiger, A.M., W. Chen, and L. Bernstein, *Myocardial infarction risk and tamoxifen therapy for breast cancer.* British Journal of Cancer, 2005. **92**(9): p. 1614-20.

139. Haque, R., et al., *Cardiovascular Disease After Aromatase Inhibitor Use.* JAMA Oncol, 2016.

140. Colleoni, M., et al., *Analyses adjusting for selective crossover show improved overall survival with adjuvant letrozole compared with tamoxifen in the BIG 1-98 study.* J Clin Oncol, 2011. **29**(9): p. 1117-24.

141. Abo-Touk, N.A., H.A. Sakr, and A. Abd El-Lattef, *Switching to letrozole versus continued tamoxifen therapy in treatment of postmenopausal women with early breast cancer.* J Egypt Natl Canc Inst, 2010. **22**(1): p. 79-85.

142. Kaufmann, M., et al., *Improved overall survival in postmenopausal women with early breast cancer after anastrozole initiated after treatment with tamoxifen compared with continued tamoxifen: the ARNO 95 Study.* J Clin Oncol, 2007. **25**(19): p. 2664-70.

143. Geiger, A.M., et al., *Stroke risk and tamoxifen therapy for breast cancer.* Journal of the National Cancer Institute, 2004. **96**(20): p. 1528-36.

144. Sogaard, K.K., et al., *Acute Pericarditis and Cancer Risk: A Matched Cohort Study Using Linked UK Primary and Secondary Care Data.* J Am Heart Assoc, 2018. **7**(16): p. e009428.

145. Boccardo, F., et al., *Switching to anastrozole versus continued tamoxifen treatment of early breast cancer. Updated results of the Italian tamoxifen anastrozole (ITA) trial.* Ann Oncol, 2006. **17 Suppl 7**: p. vii10-4.

146. Fisher, B., et al., *Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study.* Journal of the National Cancer Institute, 1998. **90**(18): p. 1371-88.

147. Chen, T.W., et al., *No increased venous thromboembolism risk in Asian breast cancer patients receiving adjuvant tamoxifen.* Breast Cancer Research & Treatment, 2014. **148**(1): p. 135-42.

148. Fisher, B., et al., *Tamoxifen in treatment of intraductal breast cancer: National Surgical Adjuvant Breast and Bowel Project B-24 randomised controlled trial.* Lancet, 1999. **353**(9169): p. 1993-2000.

149. Hernandez, R.K., et al., *Tamoxifen treatment and risk of deep venous thrombosis and pulmonary embolism: a Danish population-based cohort study.* Cancer, 2009. **115**(19): p. 4442-9.

150. Meier, C.R. and H. Jick, *Tamoxifen and risk of idiopathic venous thromboembolism.* British Journal of Clinical Pharmacology, 1998. **45**(6): p. 608-12.

151. Arimidex, T.A.o.i.C.T.G., et al., *Comprehensive side-effect profile of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: long-term safety analysis of the ATAC trial.* Lancet Oncol, 2006. **7**(8): p. 633-43.

152. Coates, A.S., et al., *Five years of letrozole compared with tamoxifen as initial adjuvant therapy for postmenopausal women with endocrine-responsive early breast cancer: update of study BIG 1-98.* Journal of Clinical Oncology, 2007. **25**(5): p. 486-92.

153. Chan, P.S., et al., *Recent trends in survival from out-of-hospital cardiac arrest in the United States.* Circulation, 2014. **130**(21): p. 1876-82.

154. British Medical Association, *2016/17 General Medical Services (GMS) contract Quality and Outcomes Framework (QOF)*, N. England, Editor. 2016: http://bit.ly/2Q1SZSp, accessed on 27/11/2018.

155. White, I.R. and J.B. Carlin, *Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values.* Stat Med, 2010. **29**(28): p. 2920-31.

156. Dunn, O., *Multiple comparison among means.* Journal of the American Statistical Association, 1961. **56**(293): p. 52-64.

157. Armstrong, R.A., *When to use the Bonferroni correction.* Ophthalmic Physiol Opt, 2014. **34**(5): p. 502-8.

158. Foglietta, J., et al., *Cardiotoxicity of Aromatase Inhibitors in Breast Cancer Patients.* Clin Breast Cancer, 2017. **17**(1): p. 11-17.

159. Early Breast Cancer Trialists' Collaborative, G., et al., *Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials.* Lancet, 2015. **386**(10001): p. 1341-52.

APPENDICES

CHAPTER 2

Appendix 2.1 – Overview of previous meta-analyses

Author	Year	Inclusion criteria	Main comparison	Main Cardiovascular Outcomes	Meta-analysis result RR (95% CI)	Notes
Khosrow-Khavar	2016	Phase III RCTs examining third generation AIs and tamoxifen among post-menopausal women with a diagnosis of breast cancer, with CVD or cerebrovascular adverse events reported.	AI vs Tam	CVD events	1.19 (1.07-1.34)	Result from AI vs Tam monotherapy. Concluded that the cardio-protective effects of tamoxifen accounted for the increase in CVD risk. Also explored sequenced therapy.
				Cerebrovascular events	0.96 (0.61-1.51)	
Ryden	2016	RCTs with long-term (at least 5 years) follow-up data of AI compared with tamoxifen or placebo with either efficacy (DFS and OS) or side effect outcomes	AI vs Tam	CVD events	1.13 (0.96-1.33)	Result from AI vs Tam monotherapy. Only one study in AI vs Tam analysis. Also explored sequenced therapy, and looked at time on and off treatment
Aydiner	2013	RCTs that included postmenopausal women that had undergone surgery for estrogen-sensitive early breast cancer, and examined the comparative effects of AIs and tamoxifen (either as monotherapy, sequenced therapy, or extended therapy) in relation to efficacy outcomes	AI vs Tam	CVD events	1.23 (0.95-1.60)	Result from AI vs Tam monotherapy. Also explored sequenced therapy.
				Thromboembolic events	0.61 (0.47-0.80)	
Amir	2011	Phase III RCTs that compared AIs with tamoxifen as initial adjuvant therapy in postmenopausal women with early stage breast cancer. Only trials that had treatment durations longer than 5 years were included.	AI vs Tam	CVD events (including MI, angina, and cardiac failure)	1.26 (1.10-1.43)	Result includes direct AI vs Tam, Tam to AI vs AI alone, and tam to AI vs AI alone
				Cerebrovascular events (including cerebrovascular accident and transient ischemic attack)	1.01 (0.81-1.26)	
				Venous Thrombosis (any venous thromboembolic episode)	0.55 (0.46-0.64)	
Cuppone	2007	Phase III RCTs that explored the cardiovascular risk of adjuvant AI compared with tamoxifen as an early switch strategy (after 2-3 years' tamoxifen) or as an upfront strategy (starting at the time of surgery and planned for 3 years. All trails must have included women who were previously untreated and had undergone surgical resection for early breast cancer.	AI vs Tam	CVD events	1.30 (1.07-1.60)	Result Includes both upfront and early switch comparisons of AI and tamoxifen
				Thromboembolic events	0.53 (0.42-0.65)	
				Cerebrovascular events	0.84 (0.68-1.05)	
Braithwaite	2003	Breast cancer treatment RCTs that explored the effect of tamoxifen on vascular outcomes	Tam vs No tam/ placebo	MI	0.74 (0.47-1.16)	Also explored some outcomes in trials of post-menopausal women, breast cancer reduction trials, and trials with tamoxifen as only treatment
				Stroke	1.48 (1.07-2.04)	

Appendix 2.2 - Systematic review search strategy

Medline	
Breast Cancer	
MeSH terms	breast neoplasms or carcinoma, ductal, breast or carcinoma, lobular or inflammatory breast neoplasms or unilateral breast neoplasms or triple negative breast neoplasms
Keywords	breast cancer or breast neoplasm* or breast tumour or breast adenocarcinoma or breast carcinogenesis or breast carcinoma or breast sarcoma
Endocrine Therapy	
MeSH terms	tamoxifen or aromatase inhibitors
Keywords	tamoxifen or aromatase inhibitor* or anastrazole or exemestane or letrozole or endocrine therapy
Cardiovascular Disease	
MeSH terms	cardiovascular diseases or heart diseases or cardiotoxicity or coronary artery disease or cardiomyopathies or heart arrest or heart failure or heart failure, diastolic or heart failure, systolic or heart valve diseases or aortic valve insufficiency or aortic valve stenosis or mitral valve insufficiency or mitral valve stenosis or pulmonary valve insufficiency or pulmonary valve stenosis or tricuspid valve insufficiency or tricuspid valve stenosis or angina pectoris or angina, unstable or angina, stable or myocardial infarction or stroke or venous thromboembolism or pulmonary embolism or pericarditis or peripheral vascular disease or arrhythmias, cardiac
Keywords	cardiovascular* or CVD or cardiac or cardiotoxi* or heart disease* or coronary artery dis* or revascular* or coronary bypass or artery bypass or aorta bypass or cardiomyopathy* or cardiopulmonary arrest* or cardiac arrest* or heart arrest* or heart failure or valvular*disease or valve disease or valve stenosis or valve insufficiency or angina* or heart infarc* or myocardial infarc* or heart attack or coronary infarc* or stroke or tia or transient ischemic attack or cerebrovascular accident or venous thromboembolism or deep*thrombo* or thromboem* or pulmonary embolism or pericarditis or peripheral vascular or peripheral art* or arrhythmia* or fibrillation or heart*flutter
Limits	
	• English language • Humans • 1960 –Current year
Embase	
Breast Cancer	
Indexed terms	breast cancer or breast tumour or basal like breast cancer or breast adenocarcinoma or breast carcinogenesis or breast carcinoma or breast sarcoma or estrogen receptor positive breast cancer or inflammatory breast cancer or triple negative breast cancer
Keywords	breast cancer or breast neoplasm* or breast neoplasm or breast tumour or breast adenocarcinoma or breast carcinogenesis or breast carcinoma or breast sarcoma
Endocrine Therapy	
Indexed terms	aromatase inhibitor or anastrozole or exemestane or letrozole or tamoxifen
Keywords	chemotherapy or anthracycline or daunorubicin or doxorubicin or epirubicin or cyclophosphamide or fluorouracil or methotrexate or taxoid* or taxane* or paclitaxel or docetaxel or tamoxifen or aromatase inhibitor* or anastrazole or exemestane or letrozole or endocrine therapy or trastuzumab or Herceptin or breast cancer treatment
Cardiovascular Disease	
Indexed terms	heart disease/ or cardiovascular disease/ or cardiotoxicity/ or heart arrhythmia/ or heart atrium arrhythmia/ or heart ventricle arrhythmia/ or atrial fibrillation/ or heart atrium flutter/ or heart ventricle arrhythmia/ or heart ventricle flutter/ or heart ventricle fibrillation/ or heart fibrillation/ or heart failure/ or acute heart failure/ or congestive heart failure/ or diastolic heart failure/ or systolic heart failure/ or heart ventricle failure/ or heart left ventricle failure/ or heart right ventricle failure/ or Ischemic cardiomyopathy/ or cardiomyopathy/ or congestive cardiomyopathy/ angina pectoris/ or stable angina pectoris/ or unstable angina pectoris/ or heart infarction/ or acute heart infarction/ or heart atrium infarction/ or pericarditis/ or valvular heart disease/ or

	aorta valve disease/ or mitral valve disease/ or pulmonary valve disease/ or tricuspid valve disease/ or aorta valve stenosis/ or mitral valve stenosis/ or heart valve stenosis/ or pulmonary valve stenosis/ or tricuspid valve stenosis/ or revascularization/ or heart arrest/ or cardiopulmonary arrest/ or cerebrovascular accident/ or venous thromboembolism/ or deep vein thrombosis/ or thromboembolism/ or embolism/ or vein thrombosis/ or peripheral vascular disease/
Keywords	cardiovascular* or CVD or cardiac or cardiotoxi* or heart disease* or coronary artery dis* or arrhythmia* or fibrillation or heart*flutter or heart failure or cardiomyopathy or angina or heart*infarc* or myocardial infarc* or heart attack or coronary infarc* or pericarditis or valvular*disease or valve disease or valve stenosis or valve insufficiency or revascular* or coronary bypass or artery bypass or aorta bypass or cardiopulmonary arrest* or cardiac arrest* or heart arrest* or cerebrovascular accident or stroke or tia or transient ischaemic attack or venous thromboembolism or deep*thrombo* or thromboem* or pulmonary embolism or peripheral vascular or peripheral art*
Limits	
	• English language • Human • Embase • 1960 –Current year • Article or review

Appendix 2.3 - Bias assessment criteria – cohort studies

Exposure	**Low** • Minimum exposure period or need for several prescriptions before classified • Exposure ascertained through prescription or pharmacy records **High** • Exposure ascertainment not clearly defined, or defined by patient or physician recall • Future information used to inform exposure status at baseline • Potential for exposure misclassification due to no information of exposure prior to index • No minimum exposure period or need for several prescriptions • Non exposed or referent group from a different population to exposed
Outcome Assessment	**Low** • Well defined diagnosis using hospital records, GP diagnosis, or similar methods • Method of outcome ascertainment has been clearly validated **High** • Unclear method of diagnosis, or diagnosis defined by patient or physician recall • Potential for differential misclassification due to different methods of outcome ascertainment being used for different exposure groups
Adjustments	**Low** • IPTW adjustment for CVD risk factors, CVD related treatment, cancer severity, major non-CVD comorbidities, other cancer treatments • Adjustment for most or all of the risk factors outlined above at baseline **High** • Minimal adjustment for one or two of the risk factors outlined above at baseline • No adjustment
Missing data	**Low** • None or low percentage of missing data, or appropriate missing data technique used such as multiple imputation **High** • Substantial amount of missing data (>20%) with no methods applied to deal with missingness • A missing category fitted to deal with missing data
Censoring	**Low** • No censoring/loss to follow up • Appropriate method of adjustment or sensitivity analysis if censoring or loss to follow up present • Censoring unlikely to have impact on results **High** • No adjustment or additional analysis where censoring/loss to follow up may cause bias

Appendix 2.4 - Bias assessment criteria – case-control studies

Case Definition	**Low** • Well defined diagnosis using hospital records, GP diagnosis, or similar methods • Method of case definition ascertainment has been clearly validated **High** • Unclear method of diagnosis, or diagnosis defined by patient or physician recall • Potential for differential misclassification due to different methods of case ascertainment between exposure groups • Likely that outcome can occur at time that is not appropriate to risk period relative to exposure
Control Selection	**Low** • Controls comparable to and chosen from the same population as cases **High** • Controls systematically different to cases due to being selected from a different population, or have very different characteristics that have not been adjusted for
Exposure Assessment	**Low** • Exposure ascertained through prescription or pharmacy records • Same method for exposure ascertainment used for cases and controls **High** • Exposure ascertainment not clearly defined, or defined by patient or physician recall • Risk of misclassification due to incomplete records on past exposure • Potential for misclassification of exposure based on outcome, or different methods used for exposure ascertainment between cases and controls
Adjustments	**Low** • Detailed adjustment for CVD risk factors, CVD related treatment, cancer severity, major non-CVD comorbidities, other cancer treatments • Adjustment for most or all of the risk factors outlined above, in less detail **High** • Minimal adjustment for one or two of the risk factors outlined above • No adjustment • Risk factors ascertained through recall by patient or physician
Missing data	**Low** • None or low percentage of missing data, or appropriate missing data technique used **High** • Substantial amount of missing data (>20%) with no methods applied to deal with missingness • A missing category fitted to deal with missing data

Appendix 2.5 – Overview of included studies

Author	Meier	Geiger	Bradbury
Year	1998	2004	2005
Title	Tamoxifen and risk of idiopathic venous thromboembolism	Stroke risk and tamoxifen therapy for breast cancer	Tamoxifen-treated breast carcinoma patients and the risk of acute myocardial infarction and newly-diagnosed angina
Country	UK	USA	UK
Study Type	Observational	Observational	Observational
Data source	GPRD	Kaiser Permanente Southern California	GPRD
Study Design	Case control	Case control	Nested case control
Age	<70 (at time of outcome)	All patients	35-80 years old
Inclusions	Women who had a computer-recorded diagnosis of breast cancer in or after 1980 and who were hospitalised for a first-time diagnosis of deep vein thrombosis or pulmonary embolism between January 1, 1991 and December 31, 1996. For each case, up to 10 control women with breast cancer were randomly selected, matched on age (within two years), duration of breast cancer (same year of breast cancer) and calendar year of VTE (same index date). Women were ineligible to be controls if they had—according to the computerized medical record—recurrent or metastatic breast cancer, died within 6 month after the index date, or underwent mastectomy, chemotherapy, radiotherapy, trauma, or major surgery within 6 months prior to the index date.	All women with a first invasive breast cancer diagnosed at KPSC between January 1, 1980, and July 1, 2000.	Women with a first-time diagnosis of breast carcinoma who were treated with tamoxifen or with bladder carcinoma, colorectal carcinoma, or non-melanoma skin cancer between January 1, 1991 and December 31, 1999. Women with other cancers (bladder, colorectal, and non-melanoma skin cancer), were selected to provide an unexposed population, because most women with breast carcinoma in the GPRD were treated with tamoxifen, and to increase the comparability of the exposure reference group to the tamoxifen-exposed group with respect to ongoing medical surveillance.
Exclusions	Any other malignancies besides breast cancer, a history of VTE or thrombophlebitis, stroke, angina pectoris, myocardial infarction, diabetes mellitus, chronic renal disease, hypertension, hyperlipidaemia, intermittent claudication, systemic lupus erythematosus, epilepsy, connective tissue disorders or cystic fibrosis. Furthermore, all potential cases were excluded if they underwent mastectomy, chemotherapy, radiotherapy, trauma (i.e. accident, bone fracture) or major	Patients with a subsequent primary cancer diagnosis (other than a second primary breast cancer, cervical cancer in situ, or basal or squamous cell skin cancer) before their stroke diagnoses were excluded from the study because the other cancer could alter their breast cancer treatment or their stroke risk. Patients with thromboembolic disease diagnoses other than stroke (i.e., myocardial infarction, venous	Women were excluded if they had a history of cancer, MI, angina pectoris, congestive heart failure, or HIV/acquired immunodeficiency syndrome before the study entry date. Women with known HIV infection were excluded because HIV infection may complicate cancer therapy, including adjuvant therapy. Women were required to have at least 1 year of recorded follow-up after their study entry date to assure

	surgery (i.e. abdominal surgery, hip replacement) within 6 months prior to the index date, who had recurrent or metastatic breast cancer, or who were in their terminal phase and died within 6 months after the index date (subjects who died from pulmonary embolism were included).	thromboembolism, or pulmonary embolism) were excluded.	adequate follow-up.
Intervention arm	Tamoxifen (any, currently exposed in VTE case-control analysis at index date) (n=133)	Any tamoxifen (in stroke case-control analysis) (n=286)	Current tamoxifen. Women who received 2 or more tamoxifen prescriptions within 1 year of their index date were considered current users (n=49)
Reference arm	Never or past tamoxifen (in VTE case-control analysis at index date) (n=64)	Unexposed to tamoxifen (in stroke case-control analysis). Unlikely to have been prescribed AIs due to study period being before approval of AIs (n=246)	Unexposed to tamoxifen. Unlikely to have been prescribed AIs due to study period being before approval of AI (n=158)
Primary end point	VTE	Stroke (hospitalisation)	Ischaemic heart disease diagnosis in primary care
Follow up time	Mean follow up 49.2 months (range 12-144)	Mean at-risk period 68.4 months (standard deviation 54 months)	N/A
Statistical methods (if available for CVD outcome)	A matched analysis was conducted by using conditional logistic regression models, and relative risk estimates (odds ratios) of developing VTE with regard to current and past use were obtained, using never users as reference group.	Case patients were compared with their individually matched control subjects using univariate and multivariable conditional logistic regression methods. Crude and adjusted odds ratios were estimated, and 95% confidence intervals were calculated. These analyses were limited to case patients who had their first stroke after their breast cancer diagnosis and their matched control subjects.	The risk of IHD was assessed for current tamoxifen users and according to the dose-response measures among all cases combined and among cases stratified by diagnosis of angina or MI. Odds ratios and 95% CIs were estimated using conditional logistic regression modelling
Adjustments	Cases and controls matched on age (within two years), duration of breast cancer (same year of breast cancer) and calendar year of VTE (same index date). Then analyses adjusted for BMI (< 30, 30+ kg m−2, unknown), smoking status (never, ex, current, unknown), and hysterectomy status (yes, no)	Menopausal status (pre- or perimenopausal, naturally postmenopausal, or menopausal because of surgery); history of hypertension (no, yes but not requiring medication, and yes requiring medication); history of diabetes (no, yes but not requiring medication, and yes requiring medication); chemotherapy (yes, no)	Cases and controls were matched on the date of the case's IHD diagnosis, age (1 year), and study entry date (6 months). Analyses were further adjusted for BMI (kg/m2), treated hypertension, use of hormone replacement therapy, and smoking status. Information concerning these risk factors was ascertained from the data base on or before the index date
Relative risk taken from paper, or calculated from raw numbers	Paper	Paper	Paper
CVDs outcome(s)	Thromboembolic events	Stroke	Angina, MI

Author	Geiger	Hernandez	Ligibel
Year	2005	2009	2012
Title	Myocardial infarction risk and tamoxifen therapy for breast cancer	Tamoxifen treatment and risk of deep venous thrombosis and pulmonary embolism: a Danish population-based cohort study	Risk of myocardial infarction, stroke, and fracture in a cohort of community-based breast cancer patients
Country	USA	Denmark	USA
Study Type	Observational	Observational	Observational
Data source	Kaiser Permanente Southern California	Danish Registries	HealthCore Integrated Research Database
Study Design	Case control	Cohort	Cohort
Age	All patients	45-69 years old	Post-menopausal
Inclusions	All women with a first invasive breast cancer diagnosed at KPSC between January 1, 1980, and July 1, 2000.	Women eligible for the study were diagnosed with International Union Against Cancer stage I or stage II oestrogen receptor-positive breast cancer between 1990 and 2004 at ages 45 to 69 years, as reported to the Danish Breast Cancer Cooperative Group (DBCG) clinical database	Women who were enrolled a minimum of 6–12 months before the first of at least 2 diagnosis codes for breast cancer during 2001–2007 and women with no diagnosis codes for breast cancer who were used as controls
Exclusions	Patients with another cancer diagnosis (other than second primary breast cancer, cervical cancer in situ or basal or squamous cell skin cancer) or thromboembolic disease (stroke, venous thromboembolism or pulmonary embolism) occurring before their MI were excluded	Women with no existing cardiovascular disease (defined using ICD-8 and ICD-10 codes) as of the date of breast cancer surgery	Metastatic cancer
Intervention arm	Any tamoxifen (in MI case-control analysis) (n=216)	Any tamoxifen during follow up (n=8232)	Currently exposed to Tamoxifen (n=4710) or AI (n=9067). Patients who were simultaneously prescribed both drugs contributed analysis time to both the tamoxifen and AI group.
Reference arm	Unexposed to tamoxifen (in case-control analysis). Unlikely to have been prescribed AIs due to study period being before approval of AIs (n=165)	Unexposed to tamoxifen. Unlikely to have been prescribed AIs due to study period being before approval of AIs (n=8057)	Not currently exposed to either tamoxifen or AI therapy (n=29497)
Primary end point	Myocardial infarction (hospitalisation)	DVT/PE (ICD-8 and -10 codes 45,099; 45,100; DI260; DI269; DI269A; DI801; DI802; DI802B; DI803; and DI803E)	Myocardial infarction, ischemic stroke, and fractures

Follow up time	Mean at-risk period 64.4 months (standard deviation 58.8 months)	Median follow up 48 months (range 0-174)	Median follow up 30 months for breast cancer patients and 33.5 months for non-breast-cancer patients
Statistical methods (if applicable/available for CVD outcome)	Case patients were compared with their individually matched control subjects using univariate and multivariable conditional logistic regression methods. Crude and adjusted odds ratios were estimated, and 95% confidence intervals were calculated. These analyses were limited to case patients who had their first stroke after their breast cancer diagnosis and their matched control subjects.	Follow-up was initiated 3 months after the surgery date. Follow-up ended on December 31, 2005. Risks of events were analysed individually by year for the first 5 years of follow-up, and then cumulatively for Years 1 to 5. RRs and 95% confidence intervals were calculated as estimates of the association between tamoxifen therapy and incident thromboembolic events. Cox proportional hazards models were used to estimate crude HRs and adjusted HRs controlling for confounding, for years 1 to 5 individually, and for Years 5 to 10 taken together. the proportional hazards assumption was tested by adding a covariate to the model to represent the interaction between exposure and the log of survival time	Propensity score matching was used. Cox proportional hazards models with time varying treatment variables were used to assess whether treatment with AIs or tamoxifen was associated with MI and stroke among women with breast cancer and to assess the association of breast cancer with the outcomes of interest. The time-varying treatment variables allowed women to contribute information to the treatment group when on treatment and to the control group when not on treatment; women who received both AIs and tamoxifen contributed to both groups. For each outcome event, women were followed from the time of their first diagnosis code only until the occurrence of the event or the censoring of their observation.
Adjustments	Menopausal status (pre- or perimenopausal, naturally postmenopausal, or menopausal because of surgery); history of hypertension (no, yes but not requiring medication, and yes requiring medication); history of diabetes (no, yes but not requiring medication, and yes requiring medication); chemotherapy (yes, no)	Age, surgical procedures (other than breast cancer surgery), metastatic tumours other than breast cancer, radiotherapy, chemotherapy, diabetes, stroke, chronic obstructive pulmonary disease, and heart failure were assessed at baseline	Age, census region, index year, Charlson index, number of drug classes used, statin use at baseline, PPI use at baseline, insurance produce, urban/rural residence, median household income in zip code, % in high school education in zip, % blacks in zip, % Hispanics in zip.
Relative risk taken from paper, or calculated from raw numbers	Paper	Paper	Paper
CVD outcome(s)	MI	Thromboembolic events	Stroke, MI

Author	Chen	Yang	Abdel-Qadir
Year	2014	2014	2016
Title	No increased venous thromboembolism risk in Asian breast cancer patients receiving adjuvant tamoxifen	Association of tamoxifen use and reduced cardiovascular events among Asian females with breast cancer	The risk of myocardial infarction with aromatase inhibitors relative to tamoxifen in post-menopausal women with early stage breast cancer
Country	Taiwan	Taiwan	Canada
Study Type	Observational	Observational	Observational
Data source	Taiwan Cancer Registry Database	NHI Research Database Taiwan	Canadian administrative databases
Study Design	Cohort	Cohort	Cohort
Age	All patients	All patients	Post-menopausal
Inclusions	Diagnosed with stage I, II, or III breast cancer, according to the American Joint Committee on Cancer staging system (sixth version) criteria between January 1, 2004 and December 31, 2009; and received curative breast cancer surgery (lumpectomy or mastectomy) within 1 year after diagnosis	Patients who were newly diagnosed with breast cancer	Women with a first diagnosis of stage I-III breast cancer r between January 1, 2005 and December 31, 2010, who were dispensed a prescription for tamoxifen, or an aromatase inhibitor (i.e. anastrozole, letrozole or exemestane) within 1 year of cancer diagnosis, along with a second prescription dispensed within 1.5-times the number of days of the preceding prescription's supply
Exclusions	Women with history of other types of cancer or multiple primary invasive breast cancer; the presence of lymphoma (ICD-O-3 morphology code, 9590-9989), Kaposi's sarcoma (ICD-O-3 morphology code, 9140), and phyllodes tumour (IDC-O-3 morphology code, 9020) of the breast; received tamoxifen treatment prior to the operation date; and death within 28 days after the operation.	Patients with pre-existing cardiovascular disease, such as coronary artery disease, ischemic stroke, haemorrhagic stroke or peripheral artery disease were excluded	Women were excluded if they were treated with tamoxifen or AIs in the year preceding breast cancer diagnosis. Women were also excluded if they had substantial exposure to both tamoxifen and an aromatase inhibitor. This was defined as >10% of the days during which either tamoxifen or an aromatase inhibitor was prescribed. Accordingly, women were only included if they were exposed to one of the drug categories for 90% of the days during which a study drug was dispensed.
Intervention arm	At least one tamoxifen prescription after the index date (n=17874)	at least one tamoxifen prescription associated with the breast cancer diagnosis (n=2056)	Aromatase inhibitors at index date (n=7049). Patients were analysed based on the drug category they were predominantly exposed to.
Reference arm	No tamoxifen prescriptions after the index date (n=10155). No information about AI prescriptions.	No tamoxifen (n=1634). No information about AI prescriptions.	Tamoxifen at index date (n=1941)
Primary end point	Deep vein thrombosis/Pulmonary embolism. DVT estimates used in systematic review results	AMI, ischemic stroke, haemorrhagic stroke and total cardiovascular events	Myocardial infarction (hospitalisation)
Follow up time	Median follow up 48 months (range 0 -96)	Mean follow up 82.8 months	Mean follow up 39.9 months

Statistical methods (if applicable/available for CVD outcome)	Outcomes were compared using the Cox proportional hazard model for estimating hazard ratios and 95 % confidence intervals.	Survival analysis was assessed using Kaplan-Meier analysis, with the significance based on the log-rank test. The survival time was calculated from the date of enrolment to the development of AMI, ischemic stroke or haemorrhagic stroke. Multiple regression analysis was carried out using Cox proportional hazard regression analysis to evaluate the effect of tamoxifen use on determining the occurrence of AMI, stroke, or total cardiovascular events.	Time-to-event analyses were performed for MI, using tamoxifen as the reference treatment. Cumulative incidence function curves were used to estimate the cumulative incidence of MI over time after accounting for the competing risk of death. This allowed us to estimate the incidence of MI, given that some subjects will die before the occurrence of a cardiac event. IPTW using the propensity score was used to reduce the effects of measured confounding variables when estimating the effect of AIs versus tamoxifen. The PS model was estimated using a logistic regression model with receipt of AIs as the dependent variable and all covariates as the independent or explanatory variables. The variables that were chosen included markers of cardiovascular disease, cardiovascular risk factors, cancer severity, major non-cardiovascular co-morbidities, health care utilisation, factors that increase risk of adverse cardiac events with breast cancer (left-sided disease, chemotherapy, trastuzumab, radiation), as well as medications that could impact risk of cardiovascular disease. Truncated weights were used to minimise undue influence from atypical individuals with very high weights. The distribution of measured baseline covariates was compared between treatment groups in the sample weighted by the inverse probability of treatment using standardised differences. Variables were determined to be well balanced if the standardised difference was
Adjustments	Congestive heart failure, rheumatic disease, renal disease, diabetes mellitus, hypertension, Charlson comorbidity index. Patients had to have been diagnosed with the comorbidity within 1 year prior to the index date. Information was obtained from the NHI database, and all of the diagnoses were identified from either a single report in the inpatient medicinal claims file or from no less than two reports in the outpatient medicinal claims files	Diabetes mellitus, cardiac arrhythmia, hyperlipidaemia, congestive heart failure, and chronic obstructive pulmonary disease, 365 days before the date of diagnosis of breast cancer. The diagnosis code of any comorbidity must have appeared at least twice and lasted longer than 30 days before officially being regarded as a comorbidity. Medications before enrolment were also reviewed within the database, which included angiotensin-converting enzyme inhibitors, β-adrenergic antagonists, calcium-channel blockers, diuretics, statins, antiplatelet agents (aspirin or clopidogrel) and thiazides.	Factors used in IPTW were: age, income quintile, rural residence, year of cohort entry, breast cancer side, chemotherapy, radiation, trastuzumab, CVD (other than MI), diabetes, dyslipidaemia, hypertension, venous thromboembolism, fracture, renal disease, dialysis, prior malignancy, Charlson index, primary care visits in past year, specialist visits in past year, total physician visits in past year, medications dispensed in past year, ACE inhibitor, ARB, aspirin, thienopyridines, beta-blockers, calcium channel blockers, digoxin, aldosterone antagonists, diuretics, statins, oral hypoglycaemics, insulin, vitamin K antagonists, low molecular weight heparin, nitrates, NSAIDs
Relative risk taken from paper, or calculated from raw numbers	Paper	Paper	Paper
CVD outcome(s)	Thromboembolic events	Stroke, MI	MI

Author	Haque	Rutqvist	McDonald
Year	2016	1993	1995
Title	Cardiovascular Disease After Aromatase Inhibitor Use	Cardiac and thromboembolic morbidity among postmenopausal women with early-stage breast cancer in a randomized trial of adjuvant tamoxifen. The Stockholm Breast Cancer Study Group	Cardiac and vascular morbidity in women receiving adjuvant tamoxifen for breast cancer in a randomised trial. The Scottish Cancer Trials Breast Group
Country	USA	Sweden	Scotland
Study Type	Observational	RCT	RCT
Data source	Kaiser Permanente Southern California		
Study Design	Cohort		
Age	Post-menopausal	Post-menopausal	<80 years old
Inclusions	Women with a first diagnosis of primary breast cancer between 1991 and 2010 and observed them through December 2011. For eligibility, women had to have pharmacy benefits, and have oestrogen– or progesterone receptor–positive breast cancer.	Histologically verified invasive breast cancer, and no previous history of cancer	Early invasive breast cancer suitable for mastectomy
Exclusions	Prior CVD (cardiac ischemia (acute myocardial infarction and angina), stroke, heart failure and cardiomyopathy, and other events (dysrhythmia, valvular dysfunction, and pericarditis))	Inoperable local disease or distant metastasis at the time of primary diagnosis, other concurrent cancers, medical contraindications to the therapy, and operation which deviated from the protocol	T4, N2, N3 or Mi lesions, more than one palpable malignant lesion or bilateral breast cancer, In-situ carcinoma (including Paget's disease) without proof of underlying or associated invasive carcinoma, Patients who were or wished to become pregnant, Those unwilling to discontinue unrelated hormone therapy including the contraceptive pill, Previous malignant disease other than successfully treated squamous or basal cell carcinoma of skin, Previous systemic therapy for breast cancer, Any cause likely to compromise adequate review, premenopausal women with proven involvement of axillary lymph nodes who were enrolled in a trial comparing ovarian ablation with chemotherapy as adjuvant treatment
Intervention arm	Current aromatase inhibitors only (n=3807)	Tamoxifen (40mg) daily for 2 years (n=203)	20mg tamoxifen daily for 5 years (n=661)
Reference arm	Current tamoxifen only (n=4207)	No tamoxifen (n=219). Unlikely to have been prescribed AIs due to study period being before approval of AIs.	No tamoxifen (n=651). Unlikely to have been prescribed AIs due to study period being before approval of AIs.
Primary end point	CVD events (cardiac ischemia (acute myocardial infarction and angina), stroke, heart failure and cardiomyopathy,	Cardiac and thromboembolic morbidity	Adverse events, including CVD

	and other events (dysrhythmia, valvular dysfunction, and pericarditis)		
Follow up time	72886 person-years of follow-up	Median follow up 60 months	5 years of tamoxifen, then could be randomized to receive more tamoxifen after this
Statistical methods (if applicable/available for CVD outcome)	Follow-up commenced on the breast cancer diagnosis date and ended on the date of one of the study end points (first CVD diagnosis date, death, termination of health plan membership, or study's end [December 31, 2011]), whichever occurred first. Hazard ratios (HRs) and 95% CIs were estimated using Cox proportional hazards models with time-dependent medication use variables	All analyses were on the basis of "intention to treat." All patient data were analysed according to the allocated treatment regardless of whether the patient actually received that treatment. No patient randomly assigned to treatment was excluded from analysis. Number of outcomes reported for adverse events.	Cox proportional hazards with censoring at date of systemic relapse, death, or at follow up to 31 December 1992.
Adjustments	Age at diagnosis, diagnosis year, breast cancer stage, race/ethnicity (from the SEER registry), geocoded median household income, body mass index, medical centre, tumour characteristics, and primary cancer treatment (surgery, radiotherapy, and chemotherapy). Comorbidities, captured in the year before breast cancer diagnosis, included hypertension, diabetes mellitus, and the Charlson comorbidity index score. Data on pharmacy use related to CVD therapy and/or prevention were also extracted. These drug covariates were coded as binary (ever or never)		
Relative risk taken from paper, or calculated from raw numbers	Paper	Calculated	Calculated (HR calculated, but using incorrect reference group)
CVD outcome(s)	MI	Stroke, Heart failure	MI, Thromboembolic events

Author	Fisher	Fisher	Jakesz
Year	1999	2001	2005
Title	Tamoxifen in treatment of intra-ductal breast cancer: National Surgical Adjuvant Breast and Bowel Project B-24 randomised controlled trial	Tamoxifen and chemotherapy for axillary node-negative, oestrogen receptor-negative breast cancer: findings from National Surgical Adjuvant Breast and Bowel Project B-23	Switching of postmenopausal women with endocrine-responsive early breast cancer to anastrozole after 2 years' adjuvant tamoxifen: combined results of ABCSG trial 8 and ARNO 95 trial
Country	USA	USA and Canada	Europe
Study Type	RCT	RCT	RCT
Data source			
Study Design			
Age	All patients	All patients	Post-menopausal
Inclusions	Women with DCIS were eligible for inclusion if their life expectancy was at least 10 years. Women with tumours that also consisted of DCIS and lobular carcinoma in situ (LCIS) were eligible. Women had to undergo lumpectomy.	primary operable, histologically node-negative, ER-negative breast cancer and a life expectancy of at least 10 years	Eligible patients were postmenopausal women aged 80 years or younger (ABCSG trial 8) or 75 years or younger (ARNO 95) with histologically verified, locally radically treated invasive or minimally invasive breast cancer without previous chemotherapy, hormone therapy, or radiotherapy, and absence of organ metastases. Women must have had 2yrs of tamoxifen (20mg) daily.
Exclusions	Women who had previously been diagnosed with cancer, except for those who had had in-situ carcinoma of the cervix or squamous-cell or basal-cell carcinoma of the skin, were not eligible.	No information given	indeterminate menopausal status (or menopausal status maintained by medication), presence of secondary malignant disease, tumour infiltration of skin or breast muscle (T4 tumours), and presence of other concomitant serious medical conditions— eg. those involving bone marrow function, the central nervous system, uncompensated cardiac insufficiency, or uncontrolled local or systemic infection.
Intervention arm	Radiation therapy followed by tamoxifen (10mg) twice daily for 5yrs (n=891)	CMF and tamoxifen (10 mg) twice a day (n=498)	Anastrozole (1mg) daily for remainder of 5-year endocrine treatment following 2 years of tamoxifen (n=1602). Follow up began after initial tamoxifen.
Reference arm	Radiation therapy followed by placebo (n=890). Unlikely to have been prescribed AIs due to study period being before	CMF and placebo (n=499). Unlikely to have been prescribed AIs due to study period being before approval	Tamoxifen (20mg) daily for remainder of 5yr endocrine treatment following 2 initial years of tamoxifen (n=1597).

	approval of AIs.	of AIs.	Follow up began after initial tamoxifen.
Primary end point	Disease free survival	Disease free survival	Disease free survival
Follow up time	Median follow up 74 months (range = 57-93 months)	Mean follow up 65 months (range 10 to 102 months)	Median follow-up 28 months (95% CI: 26–30)
Statistical methods (if applicable /available for CVD outcome)	Number of outcomes reported for adverse events.	Number of outcomes reported for adverse events.	Adverse events were only counted once per patient, and are described with absolute frequencies and proportions. Differences in the adverse event rates were estimated with exact odds ratios (OR) and corresponding 95% CIs. Exact ORs stratified by country were calculated for the five types of serious adverse events available for Austrian and German patients (myocardial infarct, embolism, thromboses, fractures, and endometrial cancer).
Adjustments			
Relative risk taken from paper, or calculated from raw numbers	Calculated	Calculated	Paper
CVD outcome(s)	Stroke, MI, Thromboembolic events	Thromboembolic events	Thromboembolic events, MI

Author	Goss	Boccardo	Coombes
Year	2005	2006	2007
Title	Randomized Trial of Letrozole Following Tamoxifen as Extended Adjuvant Therapy in Receptor-Positive Breast Cancer: Updated Findings from NCIC CTG MA.17	Switching to anastrozole versus continued tamoxifen treatment of early breast cancer. Updated results of the Italian tamoxifen anastrozole (ITA) trial	Survival and safety of exemestane versus tamoxifen after 2-3 years' tamoxifen treatment (Intergroup Exemestane Study): a randomised controlled trial
Country	Canada	Italy	UK
Study Type	RCT	RCT	RCT
Data source			
Study Design			
Age	Post-menopausal	Post-menopausal	All patients
Inclusions	Previous adjuvant tamoxifen therapy lasting 4.5 – 6 years; histologically confirmed primary breast cancer; a tumour that was positive for oestrogen receptor, progesterone receptor, or both (defined by a level of 10 fmol/mg protein or a positive result on immunohistochemical analysis of ER or PR); discontinuation of tamoxifen therapy less than 3 months before enrolment; an Eastern Cooperative Oncology Group performance status of 0, 1, or 2 (scored on a scale of 0 to 4, with lower scores indicating better function); a life expectancy of more than 5 years; and postmenopausal status. Women were defined as being postmenopausal if they were at least 50 years of age at the start of adjuvant tamoxifen therapy, were younger than 50 years at the start of tamoxifen therapy but postmenopausal at the initiation of tamoxifen therapy, were younger than 50 years at the start of tamoxifen therapy but had undergone bilateral oophorectomy, were premenopausal and younger than 50 years of age at the start of tamoxifen therapy but became amenorrheic during chemotherapy or treatment with tamoxifen, or were any age but had postmenopausal levels of luteinizing hormone or follicle-stimulating hormone prior to study enrolment. Women with unknown hormone receptor status were eligible, provided an effort was made to determine the receptor status of the primary tumour.	Histologically confirmed primary breast cancer, tumour oestrogen receptor positivity , positive axillary nodes, and no evidence of recurrent or metastatic disease, who were receiving adjuvant treatment with tamoxifen for the last 2–3 years	Patients were eligible if they had histologically confirmed, completely resected unilateral invasive breast carcinoma that was positive for oestrogen receptors or that was of unknown receptor status. Patients were postmenopausal and had received adjuvant tamoxifen therapy for at least two years but not more than three years and one month. Patients were required to have adequate hematologic, renal, and liver function at the time of randomization. b

Exclusions	No information given	Patients with a history or presence of any other cancer (except adequately treated skin cancer or carcinoma-in-situ of the cervix) and patients with any condition that may jeopardize their compliance to treatment or follow-up	The presence of a tumour with known negative oestrogen-receptor status; evidence of local relapse or a distant metastasis since the time of diagnosis; a clinically significant skeletal, cardiac, or endocrine disorder; and the use of hormone-replacement therapy within four weeks before randomization. Patients were also excluded if they had clinical evidence of severe osteoporosis or a history of a previous neoplasm other than carcinoma in situ of the cervix or basal-cell skin carcinoma or if they were taking concomitant anticoagulant agents, a selective oestrogen-receptor modulator other than tamoxifen, or any other form of hormonal therapy.
Intervention arm	Letrozole (2.5mg) daily for 5 years, following previous adjuvant tamoxifen for 4.5-6 years. Follow up began after initial tamoxifen.	1mg of anastrozole daily for remainder of 5yr endocrine treatment following 2-3yrs of tamoxifen (n=223). Follow up began after initial tamoxifen.	Exemestane (25mg) daily for remainder of 5yr endocrine treatment following 2-3 years of tamoxifen (n=2320). Follow up began after initial tamoxifen.
Reference arm	Placebo daily for 5 years, following previous adjuvant tamoxifen for 4.5-6 years. Follow up began after initial tamoxifen.	20mg of tamoxifen daily for the remainder of their 5yr endocrine treatment following 2-3 initial yrs of tamoxifen (n=225). Follow up began after initial tamoxifen.	Tamoxifen (20 or 30mg) daily for the remainder of their of 5yr endocrine treatment following 2-3 initial years of tamoxifen (n=2338). Follow up began after initial tamoxifen.
Primary end point	Disease free survival	Disease free survival	Disease free survival
Follow up time	Median follow-up 30 months (range 1.5 - 61.4 months)	Median follow up 64 months (range = 12-92 months)	Median follow up 55.7 months (range = 0-89.7 months)
Statistical methods (if available for CVD outcome)	Number of outcomes reported for adverse events.	Number of outcomes reported for adverse events.	Number of outcomes reported for adverse events.
Adjustments			
Relative risk taken from paper, or calculated from raw numbers	Calculated	Calculated	Calculated
CVD outcome(s)	Angina, Stoke, MI, Thromboembolic events	Thromboembolic events	MI, PVD

Author	Kaufmann	Forbes	Abo-Touk
Year	2007	2008	2010
Title	Improved survival in postmenopausal women with early breast cancer after anastrozole initiated after treatment with tamoxifen compared with continued tamoxifen: The ARNO 95 study	Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 100-month analysis of the ATAC trial	Switching to Letrozole Versus Continued Tamoxifen Therapy in Treatment of Postmenopausal Women with Early Breast Cancer
Country	Germany	International	Egypt
Study Type	RCT	RCT	RCT
Data source			
Study Design			
Age	Post-menopausal	Post-menopausal	Post-menopausal
Inclusions	Women with histologically verified, grade 1 to 3 invasive breast cancer (pT1-3, node negative, or up to nine tumour-infiltrated lymph nodes [pN0-2] and no distant metastases), who had undergone primary surgery (with or without radiotherapy) and had received 2 years of continuous adjuvant tamoxifen (20 or 30 mg/d) with disease recurrence	women with histologically proven operable invasive breast cancer who had completed primary surgery and chemotherapy (where given), and were candidates to receive hormonal adjuvant therapy. Patients with negative or unknown hormone-receptor status were included because hormone-receptor-negative patients were thought to derive benefit from adjuvant therapy with a hormonal agent.	Histologically confirmed operable invasive early breast carcinoma with positive oestrogen, or progesterone receptors, or both. Primary surgery was modified radical mastectomy or breast conserving surgery with axillary lymph-node dissection with resulting clear margins. There was no evidence of metastatic or recurrent disease; previous or concurrent cancer. Adequate hematologic, renal and hepatic functions were required
Exclusions	No information given	Patients were ineligible if there was any clinical evidence of metastatic disease; if chemotherapy was started more than 8 weeks after surgery or completed more than 8 weeks before starting randomised treatment (neo-adjuvant chemotherapy was not allowed) or, in patients not receiving chemotherapy, if primary surgery was completed more than 8 weeks before starting randomised treatment; or if they had received hormonal therapy for breast-cancer prevention or for adjuvant treatment of breast cancer (except if tamoxifen treatment was started before surgery and received for less than 29 days, or if hormonal therapy was received before surgery in the context of a formal trial previously approved by the Steering Committee). Patients were not eligible if they were unwilling to stop any hormonal drug including HRT; if they had a previous history of invasive	No information given

		malignant disease (breast cancer at any time, other malignant disorders within the past 10 years excluding squamous or basal-cell carcinoma of the skin or carcinoma in situ of the cervix, adequately cone biopsied); or if the patient had any severe concomitant disease which would place the patient at unusual risk or confound the results of the trial. Patients were included o	
Intervention arm	Tamoxifen (20 or 30mg) daily for 2 years followed by anastrozole (1mg) daily for another 3 years (n=445). Follow up began after initial tamoxifen.	Anastrozole only after surgery (n=3125)	Tamoxifen (20mg) daily for 2 years followed by letrozole (2.5mg) daily for another 3 years. Follow up began after initial tamoxifen.
Reference arm	Tamoxifen (20 or 30mg) daily for the remainder of their 5yr endocrine treatment following 2 years of initial tamoxifen (n=452). Follow up began after initial tamoxifen.	Tamoxifen only after surgery (n=3116)	Tamoxifen (20mg) daily for 3 years following 2 years initial tamoxifen. Follow up began after initial tamoxifen.
Primary end point	Disease free survival	Disease free survival	Disease free survival
Follow up time	Median follow up 30.1 months	Median follow up 100 months (range 0–126)	Median follow up 41 months (range 15 to 62 months)
Statistical methods (if available for CVD outcome)	Number of outcomes reported for adverse events.	Side-effects were summarised according to the hormone treatment first received. Except for other cancers, side-effect events were accrued up to 14 days after stopping treatment. Information on new primary cancers was collected during and after trial treatment (before and after recurrence), but only summarised up to the point of recurrence. The comparisons of pre-specified adverse events were based on a simple comparison of proportions, and Fisher's exact two-sided p values were used when necessary.	Number of outcomes reported for adverse events.
Adjustments			
Relative risk taken from paper, or calculated from raw numbers	Calculated	Calculated	Calculated
CVD outcome(s)	Stroke	Stroke, MI, Thromboembolic events	Stroke

Author	Colleoni	van de Velde	Bliss
Year	2011	2011	2012
Title	Analyses Adjusting for Selective Crossover Show Improved Overall Survival With Adjuvant Letrozole Compared With Tamoxifen in the BIG 1-98 Study	Adjuvant tamoxifen and exemestane in early breast cancer (TEAM): a randomised phase 3 trial	Disease-related outcomes with long-term follow-up: an updated analysis of the intergroup exemestane study
Country	International	Europe	International
Study Type	RCT	RCT	RCT
Data source			
Study Design			
Age	Post-menopausal	Post-menopausal	Post-menopausal
Inclusions	Patients were eligible for the study if they had tumours that were positive for oestrogen receptors, progesterone receptors, or both. Primary surgery with resulting clear margins and adequate hematologic, renal, and hepatic function were required.	Histologically confirmed breast adenocarcinoma and locally assessed oestrogen-receptor-positive or progesterone-receptor-positive disease who had completed local treatment administered with curative intent. Other eligibility criteria were invasive tumours of all sizes, with or without involvement of the lymph nodes (N0 to N3) and no evidence of metastatic disease.	ER–positive/ER-unknown primary invasive breast cancer who remained disease-free and on treatment after 2 to 3 years of tamoxifen, with adequate hematologic, renal, and liver function at the time of randomization
Exclusions	Evidence of metastatic disease; previous or concurrent cancer other than adequately treated non-invasive breast or cervical cancer or basal-cell or squamous-cell carcinoma of the skin within 5 years before randomization; receipt of adjuvant antioestrogen therapy for the primary breast cancer for at least 1 month; and treatment with systemic investigational drugs within 30 days before randomization or topical investigational drugs within 7 days before randomization.	Patients were excluded if they had substantial cardiac disease, other malignant diseases, or illnesses interfering with participation in the study. Further details have been previously reported.	Presence of a tumour with known negative oestrogen-receptor status; evidence of local relapse or a distant metastasis since the time of diagnosis; a clinically significant skeletal, cardiac, or endocrine disorder; and the use of hormone-replacement therapy within four weeks before randomization. Patients were also excluded if they had clinical evidence of severe osteoporosis or a history of a previous neoplasm other than carcinoma in situ of the cervix or basal-cell skin carcinoma or if they were taking concomitant anticoagulant agents, a selective oestrogen-receptor modulator other than tamoxifen, or any other form of hormonal therapy. The protocol required adequate treatment of primary disease, including postoperative radiotherapy in patients who had been treated with breast-preserving surgery. Neo-adjuvant chemotherapy was permitted according to a consistent policy within each centre. Patients were required to have started chemotherapy

			within three months after diagnosis and to have begun receiving tamoxifen and radiotherapy within three months after the completion of chemotherapy.
Intervention arm	1) Letrozole (2.5mg) daily for 2 years followed by tamoxifen (25mg) daily for 3 years (n=1540), 2) Tamoxifen (25mg) daily for 2 years followed by letrozole (2.5mg) daily for 3 years (n=1548). Follow up began after initial 2 year treatment..	Tamoxifen (20 mg) daily for 2-3yrs followed by Exemestane (25 mg) daily for the remainder of the 5yrs (n=4868). Follow up for the whole period.	25mg of exemestane daily for remainder of 5yr endocrine treatment following 2-3 years of tamoxifen (n=2105). Follow up began after initial tamoxifen.
Reference arm	1) Only tamoxifen (25mg) daily for 5 years (n=1548), 2) Only letrozole (2.5mg) daily for 5 years (n=1546). Follow up began after initial 2 year treatment.	Exemestane (25 mg) daily for 5yrs (n=4898). Follow up for the whole period.	20 or 30mg of tamoxifen daily for the remainder of their of 5yr endocrine treatment following 2-3 initial years of tamoxifen (n=2036). Follow up began after initial tamoxifen.
Primary end point	Disease free survival	Disease free survival	Disease free survival
Follow up time	Media follow up 74 months	Median follow-up 5·1 years	Median follow up 91 months (IQR=83-99.2 months)
Statistical methods (if available for CVD outcome)	Selective crossover then number of outcomes reported for adverse events.	Number of outcomes reported for adverse events.	Kaplan-Meier plots, log-rank tests, and Cox proportional hazards models were used.
Adjustments			
Relative risk taken from paper, or calculated from raw numbers	Calculated	Calculated	Paper
CVD outcome(s)	Stroke, Heart failure, Thromboembolic events	Arrhythmia, Heart failure, MI, Thromboembolic events	Angina, Thromboembolic events

Author	Pagani	Hernandez
Year	2014	2008
Title	Adjuvant Exemestane with Ovarian Suppression in Premenopausal Breast Cancer	Tamoxifen Treatment in Danish Breast Cancer Patients and 5-Year Risk of Arterial Atherosclerotic Events: A Null Association
Country	International	Denmark
Study Type	RCT	Observational
Data source		Danish Registries
Study Design		Cohort
Age	Pre-menopausal	45-69 years old
Inclusions	Histologically proven operable breast cancer confined to the breast and ipsilateral axilla, with the exception of internal-mammary-node involvement detected by means of sentinel-node biopsy, and tumour that expressed oestrogen or progesterone receptors in at least 10% of the cells, as assessed with the use of immunohistochemical testing. Patients with synchronous bilateral hormone-receptor–positive breast cancer were eligible. Patients had undergone either a total mastectomy with subsequent optional radiotherapy or breast-conserving surgery with subsequent radiotherapy. Either axillary dissection or a negative sentinel-node biopsy was required. Macrometastasis in a sentinel node required axillary dissection or irradiation.	Women eligible for the study were diagnosed with International Union Against Cancer stage I or stage II oestrogen receptor-positive breast cancer between 1990 and 2004 at ages 45 to 69 years, as reported to the Danish Breast Cancer Cooperative Group (DBCG) clinical database
Exclusions		Women with no existing cardiovascular disease (defined using ICD-8 and ICD-10 codes) as of the date of breast cancer surgery
Intervention arm	In the TEXT study, 5 yrs of exemestane (25mg daily) plus triptorelin. In the SOFT study 5 yrs of exemestane plus ovarian suppression.	Any tamoxifen during follow up (n=8232)

Reference arm	In the TEXT study 5 yrs of tamoxifen (20mg daily) plus triptorelin. In the SOFT study 5 yrs of tamoxifen plus ovarian suppression.	Unexposed to tamoxifen. Unlikely to have been prescribed AIs due to study period being before approval of AIs (n=8057)
Primary end point	Disease free survival	Angina, MI, HF, stroke
Follow up time	Median follow up of 68 months	Not reported
Statistical methods (if available for CVD outcome) Number of outcomes reported for adverse events.		Follow-up was initiated 3 months after the surgery date. Follow-up ended on December 31, 2005. Risks of events were analysed individually by year for the first 5 years of follow-up, and then cumulatively for Years 1 to 5. RRs and 95% confidence intervals were calculated as estimates of the association between tamoxifen therapy and incident CVD events. Cox proportional hazards models were used to estimate crude HRs and adjusted HRs controlling for confounding, for years 1 to 5 individually, and for Years 5 to 10 taken together. The proportional hazards assumption was tested by adding a covariate to the model to represent the interaction between exposure and the log of survival time
Adjustments		Age group, diabetes, renal disease, hypertension, chronic obstructive pulmonary disease, radiation therapy, and chemotherapy
Relative risk taken from paper, or calculated from raw numbers	Calculated	Paper
CVD outcome(s)	Stroke, MI, Thromboembolic events	Angina, stroke, MI, heart failure

Appendix 2.6 – Bias assessment of RCTs

Study	Random sequence generation	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Other sources of bias
Bliss 2012	1 - permuted blocks	0	1 - double blinded	1 - All patients included	1 - CVD events were coded according to criteria specified by an independent cardiologist	1 - No other risks of bias
Boccardo 2006	0	0	0	1 - All patients included	0	1 - No other risks of bias
Coombes 2007	1 - permuted blocks	0	1 - double blinded	1 - 95% had full follow up	0	1 - No other risks of bias
Fisher 1999	0	0	0	1 - All patients included	0	1 - No other risks of bias
Fisher 2001	1- biased coin	0	0	1 - 98% had full follow up	0	1 - No other risks of bias
Forbes 2008	1 - randomisation by computer	1 - central allocation	0	0	0	1 - No other risks of bias
Jakesz 2005	1 - randomisation by computer	1 - central allocation	2 - open label trials	0	0	1 - No other risks of bias
Kaufmann 2007	1 - randomisation by computer	1 - central allocation	0	1 - 6 patients discontinued	0	1 - No other risks of bias
McDonald 1995	0	0	0	0	0	1 - No other risks of bias
Colleoni 2011	1 - permuted blocks	0	1 - double blinded	1 - All patients included	0	1 - No other risks of bias
Rutqvist 1993	0	0	0	1 - Outcome data taken from registries	0	1 - No other risks of bias
van de Velde 2001	1 - randomisation by computer	1 - Only statistician and steering committee had access to unmasked data	2 - open label trial	0	0	1 - No other risks of bias
Abo-Touk 2010	1 - simple randomisation method	0	0	1 - All patients included	0	1 - No other risks of bias
Goss 2005	1 - minimisation method	0	1 - double blinded	1 - All patients included	0	1 - No other risks of bias
Pagani 2014	1 - permuted blacks	0	2 - open label trials	1 - All patients included	0	1 - No other risks of bias

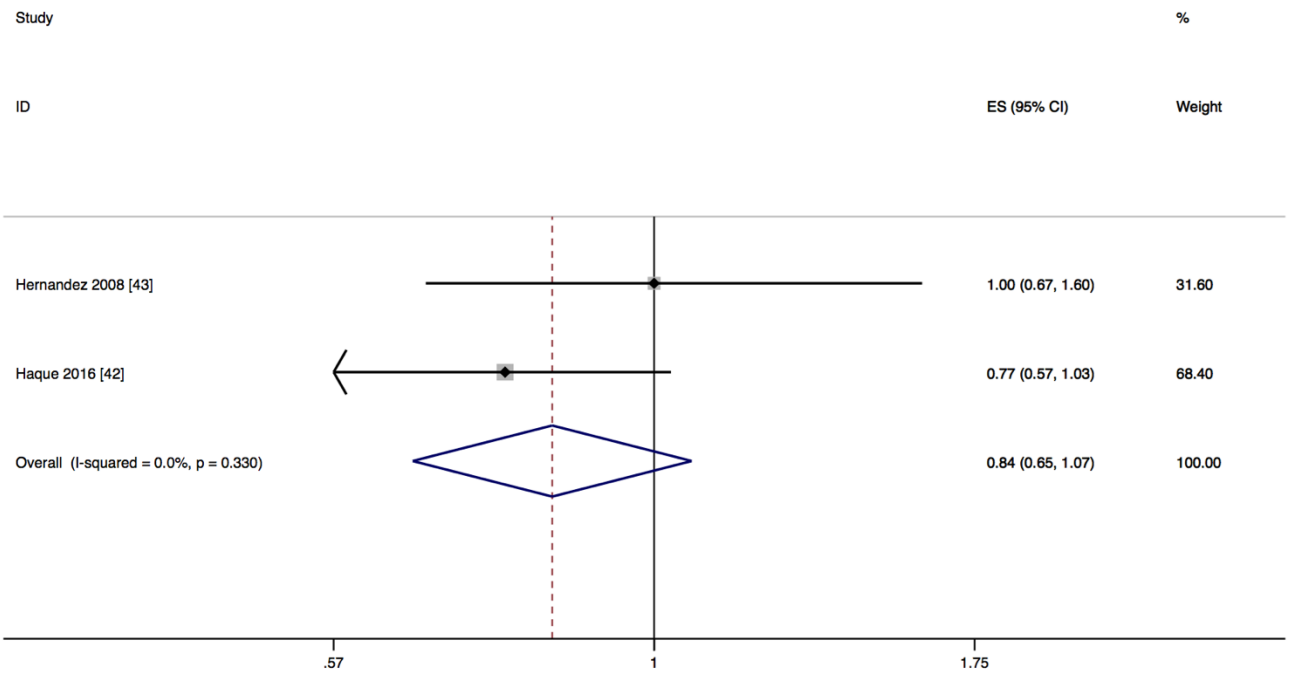
Bias assessment categories
0 - No information given
1 - Low risk of bias
2 - High risk of bias

Appendix 2.7 – Bias assessment of observational studies

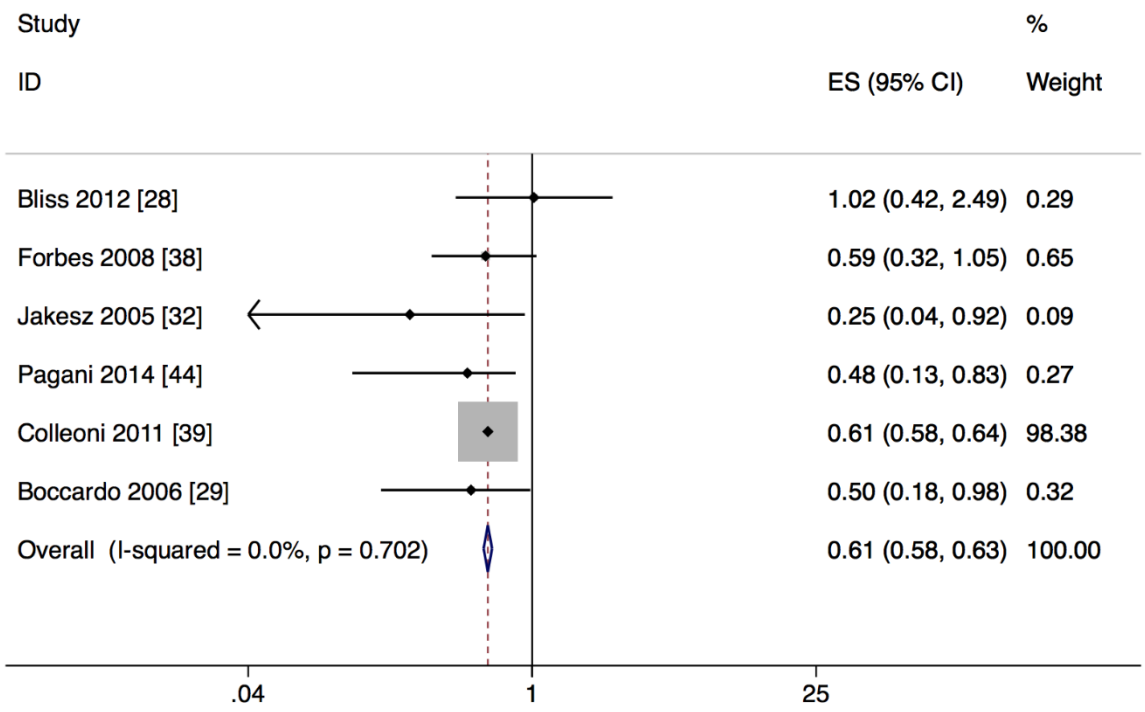
Study	Exposure definition	Outcome/case definition	Control selection	Confounding	Missing Data	Censoring
Abdel-Qadir 2016	2 - Only included women who were exposed to either AI or Tam for >90% of days dispensed	2 - Used hospital records, but not clear what method used to define outcome	N/A	1 - Adjustment for a wide range of confounders using IPTW	0 - No information given	1 - Only censored at end of study
Chen 2014	2 - Patients only needed one tamoxifen prescription to be defined as exposed	1 - Use hospital records and outlines ICD-9 codes used	N/A	2 - No adjustment for CVD related treatment, cancer severity, or other cancer treatments	0 - No information given	1 - Only censored at end of study or death
Haque 2016	2 - Patients only needed one tamoxifen prescription to be defined as exposed	1 - Identified by medical records and validated by clinician	N/A	1 - Adjustment for wide range of covariates and used IPTW	1 - Missing data on BMI, but sensitivity analyses performed to assess the impact of this	1 - Censored at death or termination of health plan membership
Hernandez 2008	0 - Not enough information given	1 - Outcome defined by ICD 8 and 10 codes	N/A	1 - Adjustment for wide range of covariates	0 - No information given	1 - Only censored at outcome or end of follow-up
Hernandez 2009	0 - Not enough information given	1 - Outcome defined by ICD 8 and 10 codes	N/A	1 - Adjustment for wide range of covariates	2 - A lot of missing BMI data and no explanation of how it is dealt with	1 - Only censored at outcome or end of follow-up
Ligibel 2012	2 - Ascertained through pharmacy data, but no information on how exposure begins	1 - Outcome identified through hospital records	N/A	2 - No adjustment for cancer severity or other cancer related treatments	0 - No information given	1 - Only censored at outcome or end of follow-up
Yang 2014	2 -Not clear when patient defined as exposed, and all breast cancer patients could be unexposed, even with ER- BC	1 - based on ICD-9 from medical records	N/A	2 - No adjustment for cancer severity or other cancer treatment	0 - No information given	0 - No information given
Bradbury 2005	2 - Patients are taken from a study population that include bladder, colorectal, and non-melanoma skin cancer patients, who would not be prescribed tamoxifen	2 - Used GP records, but no indication of terms used to define the case	1 - Three controls matched on date of IHD diagnosis, age, and study entry date	2 - No adjustment for CVD related treatment, cancer severity, or other cancer treatments	1 - Minimal missing data	N/A
Geiger 2004	1 - Exposure abstracted from medical records for cases and controls	1 - Hospital records used	1 -Two controls matched on age and members of the same health maintenance organisation during their at-risk period	2 - All risk factors adjusted for, but breast cancer therapies, smoking and some medical therapies through patient recall	2 - Missing category fitted to deal with missing data	N/A
Geiger 2005	1 - Exposure abstracted from medical records for cases and controls	1 - Hospital records used	1 -Two controls matched on age and members of the same health maintenance organisation during their at-risk period	2 - All risk factors adjusted for, but breast cancer therapies, smoking and some medical therapies through patient recall	2 - Missing category fitted to deal with missing data	N/A
Meier 1998	1 - Ascertained through computerised medical records	1 - Based on hospital records	1 - Cancer free controls from GPRD population	2 - No adjustment for cancer severity or other cancer treatment	2 - Missing data category fitted	N/A

Bias assessment categories
0 - No information given
1 - Low risk of bias
2 - High risk of bias

Appendix 2.8 - Meta-analysis of observational studies examining the risk of heart failure in tamoxifen users compared with non-users



Appendix 2.9 - Meta-analysis of RCTs examining the risk of thromboembolic events in AI users compared with tamoxifen



CHAPTER 3

Appendix 3.1 – Patient and practice level data acceptability in CPRD (provided by CPRRD)

Acceptable Patients

Patients are labelled as 'acceptable' for use in research by a process that identifies and excludes patients with non-continuous follow up or patients with poor data recording that raises suspicion as to the validity of the that patients record. Patient data are checked, for the following issues:

- An empty or invalid first registration date
- An empty or invalid current registration date
- Absence of a record for a year of birth
- A first registration date prior to their birth year
- A current registration date prior to their birth year
- A transferred out reason with no transferred out date
- A transferred out date with no transferred out reason
- A transferred out date prior to their first registration date
- A transferred out date prior to their current registration date
- A current registration date prior to their first registration date
- A gender other than Female/Male/Indeterminate
- An age of greater than 115 at end of follow up
- Recorded health care episodes in years prior to birth year
- All recorded health care episodes have empty or invalid event dates
- Registration status of temporary patients

If any of these conditions are true then the patient is labelled unacceptable, and is not recommended for use in research.

Up to standard date

The overall quality of data in practices is mediated by use of an 'up to standard' (UTS) date, which is deemed as the date at which data in the practice is considered to have continuous high quality data fit for use in research. This is mediated by an analysis on the total data in the practice, which is refreshed every time a new collection for a practice is processed into the database. It is based on two central concepts: assurance of continuity in data recording (gap analysis), and avoidance of use

of data for which transferred out and dead patients have been removed (death recording).

Gap Analysis

To detect whether there are any meaningful gaps in the data it is necessary to look in more detail at single day gaps as well as longer gaps. A single day alone may reflect a situation where nothing was recorded that day at the practice, i.e. the practice was not open, such as on a bank holiday. A longer gap may reflect a situation where the practice did not offer a service and patients may have been treated elsewhere. If a meaningful gap is found, the earliest date after which there is no significant gap is identified.

Death Recording

It is expected that a standard number of deaths will be recorded at a practice over time. Assessment of gaps in death recording is performed taking the size of the practice into account. A safety margin is built in to account for both geographical and seasonal variation in death rates. If a meaningful gap is found, the earliest date after which there is no significant gap is identified.

The UTS date is set to the latest of these dates for each practice. The CPRD recommend that analyses are performed on data following the practice UTS date.

CHAPTER 4

Appendix 4.1 – CVD code list inclusion and exclusion criteria

Outcome	Inclusion	Exclusion
Angina	angina*, stenocard*, coronary*artery*spasm*, spasm*cor*artery*, card*syndrome*x*, cor*syndrome*x*, preinfarct* impending infarct* acute coronary insuff*	fh*, no fh*, score*, vincent*, strepto*, herp*, abdominal*, ludwig*, bullosa haemor*, infarc*, therapy*
MI	*stemi*, *st*elevation*, *steami*, myocard*infarct, infarct*myocard*, acute*infarct, infarct*acute*, acute*mi*, mi*acute*, heart*infarct*, infarct*heart*, subendocard*infarct*, infarct*subendocard*, transmur*infarct*, infarct*transmur*, card*infarct*, infarct*card*, heart*attack*, attack*heart*, coronary*attack*, attack*coronary*, coronary*thrombosis*, thrombosis*coronary*, myocard*necro*, necro*myocard*, heart*necro*, necro*heart*, coron*necro*, necro*coron*, mural*thrombosis*, q*wave*	family*, fh*, leukaemia*, systemic*, chemical*, tuberculosis*, alcohol*, colitis*, epidid*, treatment programme*, memory*, endophthal*, vesiculitis*, meningoc*, porphyria*, spinal*, anaemia*, compression*, meloidosis*, psychotic*, pericarditis*, miss*, dsmiv*, minority*, mibg*, abort*, abdomin*, uraemic*, nephritic*, test of motor impairment*, no *, normal*, scor*, stress study*, mibi study, quality indicator*, military*, disseminated demyel*, aminoglycoside*
Revascularisation	cor*by*pass*, by*pass*cor*, cardio*by*pass*, by*pass*cardio*, artery*by*pass*, by*pass*artery*, aorta*by*pass*, by*pass*aorta*, vein*by*pass*, by*pass*vein*, revascular*, coron*graft*, graft*coron*, percut*cor*int*, saphen*graft*, graft*saphen*	*fh*, family*history*, presence*, rejection*, impotence*, cornea*, *revision*, complication*, planned, occlusion
SCA	cardiac*arrest*, arrest*cardiac*, electro*dissoc*, dissoc*electro*, cardio*arrest*, arrest*cardio*, circ*arrest*, arrest*circ*, resusc*cardio*, cardio*resusc*, asys*, defib*, vent*fib*, vent*tach*, cardiac massage*	fh*, family*history*, ivf, pregna*, viral*, not*, tachycardia*, flutter*, fibrillation*, renewal*, resiting*, cardiovert*defib*, pacemaker*
PVD	extremit*, leg*, limb*, iliac*, femoral*, pedis*, tibial*, popliteal*, periphera*l, extremit*, peripheral*vascular*, peripheral*arter*, intermitt*claudicat*, peripheral*vascular*, peripheral*arter*, intermitt*claudicat*, thromboangiitis*obliter*, acrocyanosis*, acroparaesthes*, erythrocyanosis*, nothnagel*, schultze*, buerger*, erythromelalgia*, claudication*, peripher*angiopath*, aortoiliac*obstruction*, arter*ulcer*, peripher*ischaem*, leg*gangrene*, leg*ischaem*, ischaem*leg*, ischaem*ulcer*foot*, ischaem*foot*, ischaem*toe*, gangrene*foot*, peripheral*circul*disorder, ischaem*toe*, gangrene*toe*, failure*peripheral*circul*, peripher*circul*fail*, extrem*arter*atheroma*, foot*gangrene*, toe*gangrene*, peripheral*gangrene*, atherosclerosis, peripheral arterial*, peripheral vascular*, goldblatt*, athero*gangrene*, monkeberg*, endarterect*, angioplasty*, arteriosclerosis*, renal*stenosis*, arter*atherom*, aorto*iliac*disease*, arteriosclerot*vascular*diseas*	fh:, family*, symptom*, excluded*, no *, prevent*, advice*, screen*, buergers exercises*, arm or leg*, pulmonary*, cerebral*, coronary*, cor art*, complications of care*, pulmon art*, false*, pulmonary*, neurogenic*, vein*, lymph node*, congenital*, anomaly*, liver*, *valsalva*, radiological*, anomalies*, arter*venous*, inflammat*, syphili*, spasm*

Stroke	stroke*, carotid*steno*, steno*carotid*, basil*steno*, steno*basil*, verteb*steno*, steno*verteb*, cereb*steno*, steno*cereb*, carotid*occlu*, occlu*carotid*, basil*occlu*, occlu*basil*, verteb*occlu*, occlu*verteb*, cereb*occlu*, occlu*cereb*, carotid*thrombo*, thrombo*carotid*, cereb*thrombo*, thrombo*cereb*, verteb*thrombo*, thrombo*verteb*, basil*thrombo*, thrombo*basil*, basil*infarct*, infarct*basil*, basal*infarct*, infarct*basal*, brain*infarct*, infarct*brain*, pontine*infarct*, infarct*pontine*, cereb*infarct*, infarct*cereb*, infarct*lobe*, lobe*infarct*, cereb*insuff*, insuff*cereb*, basil*insuff*, insuff*basil*, vertebral*insuff*, insuff*vertebral*, carotid*insuff*, insuff*carotid*, cereb*syndrom*, syndrom*cereb*, basil*syndrom*, syndrom*basil*, vertebral*syndrom*, syndrom*vertebral*, carotid*syndrom*, syndrom*carotid*, cva*, clot*brain*, brain*clot*, cereb*accident*, accident*cereb*, vascular*accident*, accident*vascular*, cereb*ischaem*, ischaem*cereb*, cereb*embol*, embol*cereb*, cran*embol*, embol*cran*, haemorr*brain*, brain*haemorr*, haemorr*cerebr*, cereb*haemorr*, haemorr*stroke*, stroke*haemorr*, basal*haemorr*, haemorr*basal*, bulbar*haemorr*, haemorr*bulbar*, capsule*haemorr*, haemorr*capsule*, pontine*haemorr*, haemorr*pontine*, cort*haemorr*, haemorr*cort*, cran*haemorr*, haemorr*cran*, dural*haematoma*, haematoma*dural*, dural*haemorr*, haemorr*dural*, brain*bleed*, bleed*brain*, cereb*bleed*, bleed*cereb*, brain*aneurys*, aneurys*brain*, cereb*aneurys*, aneurys*cereb*	All codes were checked individually
Arrhythima	arrhythmia*, dysrhythmia*, bradyarrhytmia*, irregular*heartbeat*, tachycardia*, bradycardia*, atrial*flut*, cardiac*fibrillation*, atrial*fibrillation*, premature*atrial*contrad*, premature*ventric*contrad*, *junctional*rhythm*, artioventricular*block*	fh*, family*history*, fetal*, pregnan*, neonatal*, hypsarrhyth*, excepted*, exception*, no *, cardiomyopathy*, excluded*, administration*
HF	*heart*fail*, fail*heart*, card*fail*, fail*card*, pulmon*oedema*, oedema*pulmon*, left*vent*fail*, fail*left*vent*, lvf*, card*asthma*, asthma*card*, right*vent*fail*, fail*right*vent*, rvf*, congesti*card*, card*congesti*, heart*congesti*, congesti*heart*, congesti*pulm*, pulm*congesti*, cor pulmonale*, new york heart assoc*class*, lvd*, vent*dys, rvd*, dys*vent*, malignant*hypert*, ccf*, cardiomyopath*, arrhythmogenic*dysplasia*, right*ventricular*dysplaisa*	transplant*, mechanical*, post*operation*, complication*care*, no evidence*, neonatal*, postoperative*, screen*, information given*, advice*, hyperthermia*, asthma*, preferred*, discharge*, not available*, not indicated*, declined*, except*, weak heart*, fh*, family*history*, familial*, newborn*, alcoholic*, tachy*
Pericarditis	pericarditis*, serous*pericardi*, purulent*pericardi*, fibrinous*pericardi*, caseous*pericardi*, hemhorr*pericardi*, dressler*	fh*, family*history*, dressler*
VHD	aortic*valve*stenosis*, mitral*valve*stenosis*, tricuspid*valve*stenosis*, pulmonary*valve*stenosis*, aortic*insuff*, mitral*insuff*, tricuspid*insuff*, pulmonary*insuff*, aortic*regurgit*, mitral*regurgit*, tricuspid*regurgit*, pulmonary*regurgit*, valvular*heart*, valv*dysplasia*	fh, family*history*, surgery*, shock*
DVT	deep*vein*thrombo*, thrombosis*venous*, thrombus*veins*, deep*venous*thrombosis*, thrombophleb*iliac*, thrombophleb*intracranial sinus*" "*thrombophleb*cavernous sinus*, thrombo*venous sinus*" "*venous sinus*thrombo*, thrombophleb*vein*, portal vein*thrombo*, hepatic vein*thrombo*, thrombo*vena cava*, thromb*superior mesenteric vein*	fh*, family*history*, no *, probability of*, probability score*, risk of*, screening*, prevention*, care pathway*, services admin*, saphenous*, superficial*, budd - *
PE	pulmonary*embol*	fh*, family*history*, no *, probability of*, probability score*, risk of*

Appendix 4.2 – Angina code list

Read Code	Description	Definite/possible/history of
G33..00	angina pectoris	Definite
G330.00	angina decubitus	Definite
G330000	nocturnal angina	Definite
G330z00	angina decubitus nos	Definite
G33z.00	angina pectoris nos	Definite
G33z100	stenocardia	Definite
G33z200	syncope anginosa	Definite
G33z300	angina on effort	Definite
G33z600	new onset angina	Definite
G33z700	stable angina	Definite
G33zz00	angina pectoris nos	Definite
G34y000	chronic coronary insufficiency	Definite
Gyu3000	[x]other forms of angina pectoris	Definite
187..00	frequency of angina	Possible
661M000	angina self-management plan agreed	Possible
661N000	angina self-management plan review	Possible
662K400	angina self management plan commenced	Possible
8IEY.00	referral to angina plan self-management programme declined	Possible
8T04.00	referral to angina plan self-management programme	Possible
G311.11	crescendo angina	Definite
G311.13	unstable angina	Definite
G311.14	angina at rest	Definite
G311100	unstable angina	Definite
G311200	angina at rest	Definite
G31y000	acute coronary insufficiency	Definite
G33z000	status anginosus	Definite
G311300	refractory angina	Definite
G37..00	cardiac syndrome x	Definite
G331.00	prinzmetal's angina	Definite
G331.11	variant angina pectoris	Definite
G332.00	coronary artery spasm	Definite
14A5.00	h/o: angina pectoris	History
14AJ.00	h/o: angina in last year	History
662K.00	angina control	History
662K000	angina control - good	History
662K100	angina control - poor	History
662K200	angina control - improving	History
662K300	angina control - worsening	History
662K500	angina self management plan completed	History
662Kz00	angina control nos	History
G311400	worsening angina	History

Appendix 4.3 – MI code list

Read Code	Description	Definite/possible/history of
323..00	ecg: myocardial infarction	Definite
3235.00	ecg: subendocardial infarct	Definite
323Z.00	ecg: myocardial infarct nos	Definite
889A.00	diab mellit insulin-glucose infus acute myocardial infarct	Definite
G30..00	acute myocardial infarction	Definite
G30..11	attack - heart	Definite
G30..13	cardiac rupture following myocardial infarction (mi)	Definite
G30..14	heart attack	Definite
G30..15	mi - acute myocardial infarction	Definite
G30..17	silent myocardial infarction	Definite
G300.00	acute anterolateral infarction	Definite
G301.00	other specified anterior myocardial infarction	Definite
G301000	acute anteroapical infarction	Definite
G301100	acute anteroseptal infarction	Definite
G301z00	anterior myocardial infarction nos	Definite
G302.00	acute inferolateral infarction	Definite
G303.00	acute inferoposterior infarction	Definite
G304.00	posterior myocardial infarction nos	Definite
G305.00	lateral myocardial infarction nos	Definite
G306.00	true posterior myocardial infarction	Definite
G307.00	acute subendocardial infarction	Definite
G307000	acute non-q wave infarction	Definite
G308.00	inferior myocardial infarction nos	Definite
G309.00	acute q-wave infarct	Definite
G30B.00	acute posterolateral myocardial infarction	Definite
G30X.00	acute transmural myocardial infarction of unspecif site	Definite
G30y.00	other acute myocardial infarction	Definite
G30y000	acute atrial infarction	Definite
G30y100	acute papillary muscle infarction	Definite
G30y200	acute septal infarction	Definite
G30yz00	other acute myocardial infarction nos	Definite
G30z.00	acute myocardial infarction nos	Definite
G35..00	subsequent myocardial infarction	Definite
G350.00	subsequent myocardial infarction of anterior wall	Definite
G351.00	subsequent myocardial infarction of inferior wall	Definite
G353.00	subsequent myocardial infarction of other sites	Definite
G35X.00	subsequent myocardial infarction of unspecified site	Definite
G38..00	postoperative myocardial infarction	Definite
G380.00	postoperative transmural myocardial infarction anterior wall	Definite
G381.00	postoperative transmural myocardial infarction inferior wall	Definite
G383.00	postoperative transmural myocardial infarction unspec site	Definite
G384.00	postoperative subendocardial myocardial infarction	Definite
G38z.00	postoperative myocardial infarction, unspecified	Definite
Gyu3400	[x]acute transmural myocardial infarction of unspecif site	Definite
Gyu3500	[x]subsequent myocardial infarction of other sites	Definite
Gyu3600	[x]subsequent myocardial infarction of unspecified site	Definite
SP08V00	very mild acute rejection of renal transplant	Definite
G30..12	coronary thrombosis	Possible
G30..16	thrombosis - coronary	Possible
G30A.00	mural thrombosis	Possible
G31y100	microinfarction of heart	Possible
ZV71900	[v]observation for suspected myocardial infarction	Possible
G30X000	acute st segment elevation myocardial infarction	Definite
G307100	acute non-st segment elevation myocardial infarction	Definite

14A3.00	h/o: myocardial infarct <60	History
14A4.00	h/o: myocardial infarct >60	History
14AH.00	h/o: myocardial infarction in last year	History
14AT.00	history of myocardial infarction	History
3232.00	ecg: old myocardial infarction	History
G310.00	postmyocardial infarction syndrome	History
G310.11	dressler's syndrome	History
G32..00	old myocardial infarction	History
G32..11	healed myocardial infarction	History
G32..12	personal history of myocardial infarction	History
G33z500	post infarct angina	History
G36..00	certain current complication follow acute myocardial infarct	History
G360.00	haemopericardium/current comp folow acut myocard infarct	History
G361.00	atrial septal defect/curr comp folow acut myocardal infarct	History
G362.00	ventric septal defect/curr comp fol acut myocardal infarctn	History
G364.00	ruptur chordae tendinae/curr comp fol acute myocard infarct	History
G365.00	rupture papillary muscle/curr comp fol acute myocard infarct	History
G366.00	thrombosis atrium,auric append&vent/curr comp foll acute mi	History
G501.00	post infarction pericarditis	History

Appendix 4.4 – Revascularisation code list

Read Code	Description	Definite/possible/history of
790H300	revascularisation of wall of heart	Definite
792..11	coronary artery bypass graft operations	Definite
7920.00	saphenous vein graft replacement of coronary artery	Definite
7920.11	saphenous vein graft bypass of coronary artery	Definite
7920000	saphenous vein graft replacement of one coronary artery	Definite
7920100	saphenous vein graft replacement of two coronary arteries	Definite
7920200	saphenous vein graft replacement of three coronary arteries	Definite
7920300	saphenous vein graft replacement of four+ coronary arteries	Definite
7920y00	saphenous vein graft replacement of coronary artery os	Definite
7920z00	saphenous vein graft replacement coronary artery nos	Definite
7921.00	other autograft replacement of coronary artery	Definite
7921.11	other autograft bypass of coronary artery	Definite
7921000	autograft replacement of one coronary artery nec	Definite
7921100	autograft replacement of two coronary arteries nec	Definite
7921200	autograft replacement of three coronary arteries nec	Definite
7921300	autograft replacement of four of more coronary arteries nec	Definite
7921y00	other autograft replacement of coronary artery os	Definite
7921z00	other autograft replacement of coronary artery nos	Definite
7922.00	allograft replacement of coronary artery	Definite
7922.11	allograft bypass of coronary artery	Definite
7922000	allograft replacement of one coronary artery	Definite
7922100	allograft replacement of two coronary arteries	Definite
7922200	allograft replacement of three coronary arteries	Definite
7922300	allograft replacement of four or more coronary arteries	Definite
7922y00	other specified allograft replacement of coronary artery	Definite
7922z00	allograft replacement of coronary artery nos	Definite
7923.11	prosthetic bypass of coronary artery	Definite
7925.11	creation of bypass from mammary artery to coronary artery	Definite
7928200	percut translum balloon angioplasty bypass graft coronary a	Definite
792D.00	other bypass of coronary artery	Definite
792Dy00	other specified other bypass of coronary artery	Definite
792Dz00	other bypass of coronary artery nos	Definite
792E.00	percutaneous coronary intervention	Definite
792E000	emergency percutaneous coronary intervention	Definite
7A10.00	extraanatomic bypass of aorta	Definite
7A10100	bypass aorta by anastomosis axillary to femoral artery nec	Definite
7A10400	bypass aorta anastomosis axillary artery bi femoral arteries	Definite
7A10y00	other specified extraanatomic bypass of aorta	Definite
7A10z00	extraanatomic bypass of aorta nos	Definite
7A12.00	other bypass of bifurcation of aorta	Definite
7A12000	emerg bypass bifurc aorta by anast aorta to femoral artery	Definite
7A12100	bypass bifurc aorta by anastom aorta to femoral artery nec	Definite
7A12300	bypass bifurcation aorta by anastom aorta to iliac artery	Definite
7A12y00	other specified other bypass of bifurcation of aorta	Definite
7A12z00	other bypass of bifurcation of aorta nos	Definite
7A15.00	other emergency bypass of segment of aorta	Definite
7A15000	emerg bypass ascending aorta by anastom aorta to aorta nec	Definite
7A15300	emerg bypass infrarenal aorta by anastom aorta to aorta nec	Definite
7A15400	emerg bypass abdominal aorta by anastom aorta to aorta nec	Definite
7A15y00	other specified other emergency bypass of segment of aorta	Definite
7A15z00	other emergency bypass of segment of aorta nos	Definite
7A16.00	other bypass of segment of aorta	Definite
7A16000	bypass of ascending aorta by anastomosis aorta to aorta nec	Definite
7A16100	bypass of thoracic aorta by anastomosis aorta to aorta nec	Definite

7A16300	bypass of infrarenal aorta by anastomosis aorta to aorta nec	Definite
7A16400	bypass of abdominal aorta by anastomosis aorta to aorta nec	Definite
7A16y00	other specified other bypass of segment of aorta	Definite
7A16z00	other bypass of segment of aorta nos	Definite
7A20100	intracranial bypass to carotid artery	Definite
7A20200	bypass to carotid artery nec	Definite
7A20700	intracranial bypass from carotid artery nec	Definite
7A26000	bypass of subclavian artery nec	Definite
7A26100	bypass of axillary artery nec	Definite
7A26700	bypass of brachial artery nec	Definite
7A30100	bypass of renal artery	Definite
7A33000	bypass of coeliac artery nec	Definite
7A33100	bypass of superior mesenteric artery nec	Definite
7A33200	bypass of inferior mesenteric artery nec	Definite
7A33H00	bypass of visceral branch of abdominal aorta nec	Definite
7A41.00	other bypass of iliac artery	Definite
7A41.11	other bypass of iliac artery by anastomosis	Definite
7A41100	bypass iliac artery by iliac/femoral artery anastomosis nec	Definite
7A41200	emerg bypass iliac artery by femoral/femoral art anast nec	Definite
7A41300	bypass iliac artery by femoral/femoral art anastomosis nec	Definite
7A41400	emerg bypass comm iliac art by aorta/com iliac art anast nec	Definite
7A41600	emerg bypass leg artery by aorta/com fem art anastomosis nec	Definite
7A41900	bypass common iliac artery by aorta/com iliac art anast nec	Definite
7A41B00	bypass leg artery by aorta/com femoral art anastomosis nec	Definite
7A41C00	bypass leg artery by aorta/deep femoral art anastomosis nec	Definite
7A41D00	bypass iliac artery by iliac/iliac artery anastomosis nec	Definite
7A41y00	other specified other bypass of iliac artery	Definite
7A41z00	other bypass of iliac artery nos	Definite
7A47.00	other emergency bypass of femoral artery or popliteal artery	Definite
7A47.12	other emergency bypass of common femoral artery	Definite
7A47.13	other emergency bypass of deep femoral artery	Definite
7A47.14	other emergency bypass of popliteal artery	Definite
7A47.15	other emergency bypass of superficial femoral artery	Definite
7A47.16	other emergency bypass of femoral artery	Definite
7A47200	emerg bypass femoral art by fem/pop a anast c vein graft nec	Definite
7A47300	emerg bypass pop art by pop/pop art anast c vein graft nec	Definite
7A47600	emerg bypass femoral art by fem/tib a anast c vein graft nec	Definite
7A47700	emerg bypass pop art by pop/tib art anast c vein graft nec	Definite
7A47B00	emerg bypass pop art by pop/peron art anast c vein graft nec	Definite
7A47C00	emerg bypass femoral artery by fem/fem art anastomosis nec	Definite
7A47D00	emerg bypass popliteal artery by pop/fem art anastomosis nec	Definite
7A47y00	other emergency bypass of femoral or popliteal artery os	Definite
7A47z00	other emergency bypass of femoral or popliteal artery nos	Definite
7A48.00	other bypass of femoral artery or popliteal artery	Definite
7A48.11	other bypass of femoral or popliteal artery by anastomosis	Definite
7A48.12	other bypass of common femoral artery	Definite
7A48.14	other bypass of femoral artery	Definite
7A48.15	other bypass of popliteal artery	Definite
7A48.16	other bypass of superficial femoral artery	Definite
7A48000	bypass femoral artery by fem/pop art anast c prosthesis nec	Definite
7A48100	bypass popliteal artery by pop/pop a anast c prosthesis nec	Definite
7A48200	bypass femoral artery by fem/pop art anast c vein graft nec	Definite
7A48300	bypass popliteal artery by pop/pop a anast c vein graft nec	Definite
7A48400	bypass femoral artery by fem/tib art anast c prosthesis nec	Definite
7A48500	bypass popliteal artery by pop/tib a anast c prosthesis nec	Definite
7A48600	bypass femoral artery by fem/tib art anast c vein graft nec	Definite
7A48700	bypass popliteal artery by pop/tib a anast c vein graft nec	Definite
7A48800	bypass femoral artery by fem/peron a anast c prosthesis nec	Definite
7A48900	bypass popliteal artery by pop/peron art anast c prosth nec	Definite
7A48A00	bypass femoral artery by fem/peron a anast c vein graft nec	Definite
7A48B00	bypass popliteal art by pop/peron art anast c vein graft nec	Definite

7A48C00	bypass femoral artery by femoral/femoral art anastomosis nec	Definite
7A48D00	bypass popliteal artery by pop/fem artery anastomosis nec	Definite
7A48y00	other bypass of femoral artery or popliteal artery os	Definite
7A48z00	other bypass of femoral artery or popliteal artery nos	Definite
7A64.00	other bypass operations on vein	Definite
7A64y00	other specified bypass operation on vein	Definite
7A64z00	bypass operation on vein nos	Definite
7A66000	crossover graft of saphenous vein	Definite
7A66011	palma crossover graft of saphenous vein	Definite
7M36000	cardiopulmonary bypass	Definite
7M36400	modified ultrafiltration adjunct cardiopulmonary bypass	Definite
8L40.00	coronary artery bypass graft operation planned	Possible
SP07600	coronary artery bypass graft occlusion	Definite

Appendix 4.5 - SCA code list

Read Code	Description	Definite/possible/history of
2241.00	o/e - collapse -cardiac arrest	Definite
7932111	cardiac massage - open	Definite
7937500	implantation of internal cardiac defibrillator	Definite
7L1H600	advanced cardiopulmonary resuscitation	Definite
853..00	cardiac massage - external	Definite
8531.00	closed cardiac massage alone	Definite
8532.00	closed cardiac massage+ventil.	Definite
8532.11	cardiopulmonary resuscitation	Definite
853Z.00	external cardiac massage nos	Definite
G575.00	cardiac arrest	Definite
G575.11	cardio-respiratory arrest	Definite
G575.12	asystole	Definite
G575000	cardiac arrest with successful resuscitation	Definite
G575200	electromechanical dissociation with successful resuscitation	Definite
G575300	electromechanical dissociation	Definite
G575z00	cardiac arrest, unspecified	Definite
SP11000	cardiac arrest as a complication of care	Definite
7937600	removal of internal cardiac defibrillator	History

Appendix 4.6 - PVD code list

Read Code	Description	Definite/possible/history of
Q31y400	perinatal acrocyanosis	Definite
Gyu7400	[x]other specified peripheral vascular diseases	Definite
G73z012	vascular claudication	Definite
G73..00	other peripheral vascular disease	Definite
1M11100	ischaemic foot pain when walking	Definite
16I..00	claudication distance	Definite
C109F00	non-insulin-dependent d m with peripheral angiopath	Definite
7A49000	endarterectomy and patch repair of femoral artery	Definite
7A4B000	percutaneous transluminal angioplasty of femoral artery	Definite
G73y100	peripheral angiopathic disease ec nos	Definite
R055011	[d]peripheral circulatory failure	Definite
Pyu2B00	[x]oth specified cong malform of peripheral vascular system	Definite
7A42011	endarterectomy and patch repair of common iliac artery	Definite
C107z00	diabetes mellitus nos with peripheral circulatory disorder	Definite
G73z011	claudication	Definite
7A56600	percutaneous transluminal placement peripheral stent artery	Definite
7A42000	endarterectomy and patch repair of iliac artery	Definite
G732000	gangrene of toe	Definite
M271400	mixed venous and arterial leg ulcer	Definite
G73y000	diabetic peripheral angiopathy	Definite
G73..11	peripheral ischaemic vascular disease	Definite
C109F11	type ii diabetes mellitus with peripheral angiopathy	Definite
C109F12	type 2 diabetes mellitus with peripheral angiopathy	Definite
C108G00	insulin dependent diab mell with peripheral angiopathy	Definite
G702.00	extremity artery atheroma	Definite
G731z00	thromboangiitis obliterans nos	Definite
7A42111	endarterectomy of common iliac artery nec	Definite
G742z00	peripheral arterial embolism and thrombosis nos	Definite
7A44000	percutaneous transluminal angioplasty of iliac artery	Definite
C10FF00	type 2 diabetes mellitus with peripheral angiopathy	Definite
7A49300	endarterectomy of popliteal artery nec	Definite
G73y511	nothnagel's vasomotor acroparaesthesia	Definite
M271300	arterial leg ulcer	Definite
G731.00	thromboangiitis obliterans	Definite
C107100	diabetes mellitus, adult, + peripheral circulatory disorder	Definite
7A4B100	percutaneous transluminal angioplasty of popliteal artery	Definite
G731000	buerger's disease	Definite
7A49100	endarterectomy and patch repair of popliteal artery	Definite
G734.00	peripheral arterial disease	Definite
C10FF11	type ii diabetes mellitus with peripheral angiopathy	Definite
M271000	ischaemic ulcer diabetic foot	Definite
M271.12	ischaemic leg ulcer	Definite
G740.12	aortoiliac obstruction	Definite
G73y800	erythromelalgia	Definite
G700.11	aorto-iliac disease	Definite
G73y400	acroparaesthesia - schultze's type	Definite
1M11000	ischaemic foot pain at rest	Definite
C107.00	diabetes mellitus with peripheral circulatory disorder	Definite
G733.00	ischaemic foot	Definite
G73..13	peripheral ischaemia	Definite
R054300	[d]widespread diabetic foot gangrene	Definite
G73..12	ischaemia of legs	Definite

38DJ.00	edinburgh claudication questionnaire	Definite
C107400	niddm with peripheral circulatory disorder	Definite
G73z.00	peripheral vascular disease nos	Definite
7A49200	endarterectomy of femoral artery nec	Definite
G73y600	acroparaesthesia - unspecified	Definite
G732.00	peripheral gangrene	Definite
C107300	iddm with peripheral circulatory disorder	Definite
G73y200	acrocyanosis	Definite
7A42100	endarterectomy of iliac artery nec	Definite
G73z000	intermittent claudication	Definite
R055000	[d]failure of peripheral circulation	Definite
G702z00	extremity artery atheroma nos	Definite
G73y700	erythrocyanosis	Definite
R054200	[d]gangrene of toe in diabetic	Definite
G73y.00	other specified peripheral vascular disease	Definite
2G63.00	ischaemic toe	Definite
G73yz00	other specified peripheral vascular disease nos	Definite
C107000	diabetes mellitus, juvenile +peripheral circulatory disorder	Definite
G73y500	acroparaesthesia - nothnagel's type	Definite
G732100	gangrene of foot	Definite
G73zz00	peripheral vascular disease nos	Definite
A3A0F00	gas gangrene-foot	Definite
C10EG00	type 1 diabetes mellitus with peripheral angiopathy	Definite
7A42012	iliac endarterectomy and patch	Definite
9hS0.00	excepted frm peripheral arterial dis qual ind: pt unsuitable	Possible
9N4h.00	dna - did not attend peripheral vascular disease clinic	Possible
8HlP.00	referred for peripheral artery disease assessment	Possible
9hS1.00	except frm peripheral arter dis qual indicat: inform dissent	Possible
9m10.00	peripheral vascular disease monitoring first letter	History
9m12.00	peripheral vascular disease monitoring third letter	History
14F7.00	h/o: arterial lower limb ulcer	History
9m1..00	peripheral vascular disease monitoring invitation	History
9m11.00	peripheral vascular disease monitoring second letter	History
662U.00	peripheral vascular disease monitoring	History
14NB.00	h/o: peripheral vascular disease procedure	History

Appendix 4.7 – Stroke code list

Read Code	Description	Definite/possible/history of
G63..11	infarction - precerebral	Definite
G63y000	cerebral infarct due to thrombosis of precerebral arteries	Definite
G63y100	cerebral infarction due to embolism of precerebral arteries	Definite
G64..11	cva - cerebral artery occlusion	Definite
G64..12	infarction - cerebral	Definite
G64..13	stroke due to cerebral arterial occlusion	Definite
G640000	cerebral infarction due to thrombosis of cerebral arteries	Definite
G641000	cerebral infarction due to embolism of cerebral arteries	Definite
G64z.00	cerebral infarction nos	Definite
G64z.11	brainstem infarction nos	Definite
G64z.12	cerebellar infarction	Definite
G64z000	brainstem infarction	Definite
G64z200	left sided cerebral infarction	Definite
G64z300	right sided cerebral infarction	Definite
G64z400	infarction of basal ganglia	Definite
G676000	cereb infarct due cerebral venous thrombosis, nonpyogenic	Definite
G6W..00	cereb infarct due unsp occlus/stenos precerebr arteries	Definite
G6X..00	cerebrl infarctn due/unspcf occlusn or sten/cerebrl artrs	Definite
Gyu6300	[x]cerebrl infarctn due/unspcf occlusn or sten/cerebrl artrs	Definite
Gyu6400	[x]other cerebral infarction	Definite
Gyu6G00	[x]cereb infarct due unsp occlus/stenos precerebr arteries	Definite
G62..00	other and unspecified intracranial haemorrhage	Definite
G62z.00	intracranial haemorrhage nos	Definite
C154211	adrenocortical haemorrhage	Definite
G61..00	intracerebral haemorrhage	Definite
G61..11	cva - cerebrovascular accid due to intracerebral haemorrhage	Definite
G61..12	stroke due to intracerebral haemorrhage	Definite
G610.00	cortical haemorrhage	Definite
G611.00	internal capsule haemorrhage	Definite
G612.00	basal nucleus haemorrhage	Definite
G613.00	cerebellar haemorrhage	Definite
G614.00	pontine haemorrhage	Definite
G616.00	external capsule haemorrhage	Definite
G617.00	intracerebral haemorrhage, intraventricular	Definite
G618.00	intracerebral haemorrhage, multiple localized	Definite
G619.00	lobar cerebral haemorrhage	Definite
G61X.00	intracerebral haemorrhage in hemisphere, unspecified	Definite
G61X000	left sided intracerebral haemorrhage, unspecified	Definite
G61X100	right sided intracerebral haemorrhage, unspecified	Definite
G61z.00	intracerebral haemorrhage nos	Definite
Gyu6200	[x]other intracerebral haemorrhage	Definite
Gyu6F00	[x]intracerebral haemorrhage in hemisphere, unspecified	Definite
G601.00	subarachnoid haemorrhage from carotid siphon and bifurcation	Definite
G602.00	subarachnoid haemorrhage from middle cerebral artery	Definite
G60X.00	subarachnoid haemorrh from intracranial artery, unspecif	Definite
Gyu6000	[x]subarachnoid haemorrhage from other intracranial arteries	Definite
Gyu6E00	[x]subarachnoid haemorrh from intracranial artery, unspecif	Definite
2Ba2200	scpe class predom patt c.3 infarct of middle cerebral artery	Definite
9N0p.00	seen in stroke clinic	Definite
Fyu5700	[x]other vascular syndroms/brain in cerebrovasculr diseases	Definite
G66..00	stroke and cerebrovascular accident unspecified	Definite

G66..11	cva unspecified	Definite
G66..12	stroke unspecified	Definite
G66..13	cva - cerebrovascular accident unspecified	Definite
G663.00	brain stem stroke syndrome	Definite
G664.00	cerebellar stroke syndrome	Definite
G665.00	pure motor lacunar syndrome	Definite
G667.00	left sided cva	Definite
G668.00	right sided cva	Definite
L440.11	cva - cerebrovascular accident in the puerperium	Definite
L440.12	stroke in the puerperium	Definite
1JA1000	suspected cerebrovascular accident	Possible
1JA1011	suspected stroke	Possible
661M700	stroke self-management plan agreed	Possible
8HBJ.00	stroke / transient ischaemic attack referral	Possible
8HHM.00	ref to multidisciplinary stroke function improvement service	Possible
8HTQ.00	referral to stroke clinic	Possible
8Hd6.00	admission to stroke unit	Possible
C315100	mitochond encephalopathy, lact acidosis & strokelike episode	Possible
7017000	evacuation of subdural haematoma	Definite
G621.00	subdural haemorrhage - nontraumatic	Definite
G622.00	subdural haematoma - nontraumatic	Definite
G623.00	subdural haemorrhage nos	Definite
7032000	evacuation of extradural haematoma	Definite
G620.00	extradural haemorrhage - nontraumatic	Definite
A94y600	rupture of syphilitic cerebral aneurysm	Definite
14A7.00	h/o: cva/stroke	History
14A7.11	h/o: cva	History
14A7.12	h/o: stroke	History
14AK.00	h/o: stroke in last year	History
1M4..00	central post-stroke pain	History
5C13.00	old cerebral infarction on imaging	History
661N700	stroke self-management plan review	History
662M.00	stroke monitoring	History
662M100	stroke 6 month review	History
662M200	stroke initial post discharge review	History
662e.00	stroke/cva annual review	History
662e.11	stroke annual review	History
662o.00	haemorrhagic stroke monitoring	History
7P24200	delivery of rehabilitation for stroke	History
9Om..00	stroke/transient ischaemic attack monitoring administration	History
9Om0.00	stroke/transient ischaemic attack monitoring first letter	History
9Om1.00	stroke/transient ischaemic attack monitoring second letter	History
9Om2.00	stroke/transient ischaemic attack monitoring third letter	History
9Om3.00	stroke/transient ischaemic attack monitoring verbal invitati	History
9Om4.00	stroke/transient ischaemic attack monitoring telephone invte	History
G64z100	wallenberg syndrome	History
G64z111	lateral medullary syndrome	History
G681.00	sequelae of intracerebral haemorrhage	History
G682.00	sequelae of other nontraumatic intracranial haemorrhage	History
G683.00	sequelae of cerebral infarction	History
G68X.00	sequelae of stroke,not specfd as h'morrhage or infarction	History
ZLEP.00	discharge from stroke serv	History
ZV12511	[v]personal history of stroke	History
ZV12512	[v]personal history of cerebrovascular accident (cva)	History

Appendix 4.8 – Arrhythmia code list

Read Code	Description	Definite/possible/history of
3272.00	ecg: atrial fibrillation	Definite
3273.00	ecg: atrial flutter	Definite
793M100	perc transluminal ablation of atrial wall for atrial flutter	Definite
793M300	perc translum ablat conduct sys heart for atrial flutter nec	Definite
8OAD.00	provision of written information about atrial fibrillation	Definite
G573.00	atrial fibrillation and flutter	Definite
G573000	atrial fibrillation	Definite
G573100	atrial flutter	Definite
G573200	paroxysmal atrial fibrillation	Definite
G573300	non-rheumatic atrial fibrillation	Definite
G573400	permanent atrial fibrillation	Definite
G573500	persistent atrial fibrillation	Definite
G573600	paroxysmal atrial flutter	Definite
G573z00	atrial fibrillation and flutter nos	Definite
7936A00	implant intravenous pacemaker for atrial fibrillation	Possible
8CMW200	atrial fibrillation care pathway	Possible
8HTy.00	referral to atrial fibrillation clinic	Possible
2426.00	o/e - pulse rate tachycardia	Definite
2426.11	o/e - tachycardia	Definite
3282.00	ecg: ventricular tachycardia	Definite
G570.00	paroxysmal supraventricular tachycardia	Definite
G570000	paroxysmal atrial tachycardia	Definite
G570100	paroxysmal atrioventricular tachycardia	Definite
G570200	paroxysmal junctional tachycardia	Definite
G570300	paroxysmal nodal tachycardia	Definite
G570z00	paroxysmal supraventricular tachycardia nos	Definite
G571.00	paroxysmal ventricular tachycardia	Definite
G571.11	ventricular tachycardia	Definite
G572.00	paroxysmal tachycardia unspecified	Definite
G572000	essential paroxysmal tachycardia	Definite
G572z00	paroxysmal tachycardia nos	Definite
G57y700	sinus tachycardia	Definite
G57y900	supraventricular tachycardia nos	Definite
R050.00	[d]tachycardia, unspecified	Definite
R050.12	[d]postural orthostatic tachycardia syndrome (pots)	Definite
2422.00	o/e - pulse rate - bradycardia	Definite
2422.11	o/e - bradycardia	Definite
G57y000	persistent sinus bradycardia	Definite
G57y100	severe sinus bradycardia	Definite
R059.00	[d]sinus bradycardia	Definite
R05W.00	[d] bradycardia, unspecified	Definite
Ryu0600	[x]bradycardia, unspecified	Definite
327..00	ecg: supraventricular arrhythmia	Definite
328..00	ecg: ventricular arrhythmia	Definite
328Z.00	ecg: ventricular arrhythmia nos	Definite
G57..00	cardiac dysrhythmias	Definite
G57..11	cardiac arrhythmias	Definite
G577.00	sinus arrhythmia	Definite
G57y.00	other cardiac dysrhythmias	Definite
G57yA00	re-entry ventricular arrhythmia	Definite
G57yz00	other cardiac dysrhythmia nos	Definite
G57z.00	cardiac dysrhythmia nos	Definite
Gyu5a00	[x]other specified cardiac arrhythmias	Definite
1J62.00	suspected arrhythmia	Possible

14AN.00	h/o: atrial fibrillation	History
14AP.00	history of ventricular tachycardia	History
14AQ.00	history of supraventricular tachycardia	History
14AR.00	history of atrial flutter	History
212R.00	atrial fibrillation resolved	History
662S.00	atrial fibrillation monitoring	History
6A9..00	atrial fibrillation annual review	History
9Os0.00	atrial fibrillation monitoring first letter	History
9Os1.00	atrial fibrillation monitoring second letter	History
9Os2.00	atrial fibrillation monitoring third letter	History
9Os3.00	atrial fibrillation monitoring verbal invite	History
9Os4.00	atrial fibrillation monitoring telephone invite	History

Appendix 4.9 – HF code list

Read Code	Description	Definite/possible/history of
G343.00	ischaemic cardiomyopathy	Definite
G554.00	other primary cardiomyopathies	Definite
G554000	congestive cardiomyopathy	Definite
G554011	congestive obstructive cardiomyopathy	Definite
G554400	primary dilated cardiomyopathy	Definite
G554z00	other primary cardiomyopathy nos	Definite
G55y.11	secondary dilated cardiomyopathy	Definite
G551.00	hypertrophic obstructive cardiomyopathy	Definite
G554300	hypertrophic non-obstructive cardiomyopathy	Definite
Gyu5M00	[x]other hypertrophic cardiomyopathy	Definite
Gyu5N00	[x]other restrictive cardiomyopathy	Definite
G559.00	arrhythmogenic right ventricular cardiomyopathy	Definite
F391B00	cardiomyopathy in duchenne muscular dystrophy	Definite
G55..00	cardiomyopathy	Definite
G552.00	obscure african cardiomyopathy	Definite
G554100	constrictive cardiomyopathy	Definite
G554500	takotsubo cardiomyopathy	Definite
G554511	stress cardiomyopathy	Definite
G557.00	nutritional and metabolic cardiomyopathies	Definite
G557z00	nutritional and metabolic cardiomyopathy nos	Definite
G558.00	cardiomyopathy in disease ec	Definite
G558000	cardiomyopathy in friedreich's ataxia	Definite
G558100	cardiomyopathy in myotonic dystrophy	Definite
G558200	dystrophic cardiomyopathy	Definite
G558400	amyloid cardiomyopathy	Definite
G558z00	cardiomyopathy in diseases ec, nos	Definite
G55y.00	secondary cardiomyopathy nos	Definite
G55y000	cardiomyopathy due to drugs and other external agents	Definite
G55z.00	cardiomyopathy nos	Definite
Gyu5P00	[x]other cardiomyopathies	Definite
Gyu5R00	[x]cardiomyopathy in metabolic diseases ce	Definite
L186500	cardiomyopathy in the puerperium	Definite
1736.00	paroxysmal nocturnal dyspnoea	Possible
1J60.00	suspected heart failure	Possible
1O1..00	heart failure confirmed	Definite
23E1.00	o/e - pulmonary oedema	Possible
2JZ..00	on optimal heart failure therapy	Possible
388D.00	new york heart assoc classification heart failure symptoms	Definite
585f.00	echocardiogram shows left ventricular systolic dysfunction	Definite
585g.00	echocardiogram shows left ventricular diastolic dysfunction	Definite
661M500	heart failure self-management plan agreed	Possible
662f.00	new york heart association classification - class i	Definite
662g.00	new york heart association classification - class ii	Definite
662h.00	new york heart association classification - class iii	Definite
662i.00	new york heart association classification - class iv	Definite
679W100	education about deteriorating heart failure	Possible
679X.00	heart failure education	Possible
8B29.00	cardiac failure therapy	Possible
8CL3.00	heart failure care plan discussed with patient	Possible
8CMW800	heart failure clinical pathway	Possible
8H2S.00	admit heart failure emergency	Definite

8HHz.00	referral to heart failure exercise programme	Possible
8HTL000	referral to rapid access heart failure clinic	Possible
8Hk0.00	referred to heart failure education group	Possible
9N0k.00	seen in heart failure clinic	Possible
9N2p.00	seen by community heart failure nurse	Possible
9N4s.00	did not attend practice nurse heart failure clinic	Possible
9N4w.00	did not attend heart failure clinic	Possible
9N6T.00	referred by heart failure nurse specialist	Possible
G400.00	acute cor pulmonale	Definite
G41z.11	chronic cor pulmonale	Definite
G58..00	heart failure	Definite
G58..11	cardiac failure	Definite
G580.00	congestive heart failure	Definite
G580.11	congestive cardiac failure	Definite
G580.12	right heart failure	Definite
G580.13	right ventricular failure	Definite
G580.14	biventricular failure	Definite
G580000	acute congestive heart failure	Definite
G580100	chronic congestive heart failure	Definite
G580200	decompensated cardiac failure	Definite
G580300	compensated cardiac failure	Definite
G581.00	left ventricular failure	Definite
G581.12	pulmonary oedema - acute	Possible
G581.13	impaired left ventricular function	Definite
G581000	acute left ventricular failure	Definite
G582.00	acute heart failure	Definite
G583.00	heart failure with normal ejection fraction	Definite
G583.11	hfnef - heart failure with normal ejection fraction	Definite
G583.12	heart failure with preserved ejection fraction	Definite
G584.00	right ventricular failure	Definite
G58z.00	heart failure nos	Definite
G58z.12	cardiac failure nos	Definite
G5yy900	left ventricular systolic dysfunction	Definite
G5yyA00	left ventricular diastolic dysfunction	Definite
G5yyB00	right ventricular diastolic dysfunction	Definite
G5yyD00	left ventricular cardiac dysfunction	Definite
G5yyE00	right ventricular systolic dysfunction	Definite
H465.00	chemical-induced pulmonary oedema	Possible
H54..00	pulmonary congestion and hypostasis	Possible
H541.00	pulmonary congestion	Possible
H541000	chronic pulmonary oedema	Possible
H541z00	pulmonary oedema nos	Possible
H54z.00	pulmonary congestion and hypostasis nos	Possible
H584.00	acute pulmonary oedema unspecified	Possible
H584z00	acute pulmonary oedema nos	Possible
R2y1000	[d]cardiorespiratory failure	Definite
ZRad.00	new york heart assoc classification heart failure symptoms	Definite
G580400	congestive heart failure due to valvular disease	Definite
G210.00	malignant hypertensive heart disease	Definite
G210000	malignant hypertensive heart disease without ccf	Definite
G210100	malignant hypertensive heart disease with ccf	Definite
G210z00	malignant hypertensive heart disease nos	Definite
G211100	benign hypertensive heart disease with ccf	Definite
G21z100	hypertensive heart disease nos with ccf	Definite
G230.00	malignant hypertensive heart and renal disease	Definite
G232.00	hypertensive heart&renal dis wth (congestive) heart failure	Definite
G234.00	hyperten heart&renal dis+both(congestv)heart and renal fail	Definite
G1yz100	rheumatic left ventricular failure	Definite

H461.00	acute pulmonary oedema due to chemical fumes	Possible
Q48y100	congenital cardiac failure	Definite
14A6.00	h/o: heart failure	History
14AM.00	h/o: heart failure in last year	History
183B.00	worsening pulmonary oedema	History
2126400	heart failure resolved	History
662T.00	congestive heart failure monitoring	History
662W.00	heart failure annual review	History
662p.00	heart failure 6 month review	History
8CMK.00	has heart failure management plan	History
8HBE.00	heart failure follow-up	History
9On..00	left ventricular dysfunction monitoring administration	History
9On0.00	left ventricular dysfunction monitoring first letter	History
9On1.00	left ventricular dysfunction monitoring second letter	History
9On2.00	left ventricular dysfunction monitoring third letter	History
9On3.00	left ventricular dysfunction monitoring verbal invite	History
9On4.00	left ventricular dysfunction monitoring telephone invite	History
9Or..00	heart failure monitoring administration	History
9Or0.00	heart failure review completed	History
9Or1.00	heart failure monitoring telephone invite	History
9Or2.00	heart failure monitoring verbal invite	History
9Or3.00	heart failure monitoring first letter	History
9Or4.00	heart failure monitoring second letter	History
9Or5.00	heart failure monitoring third letter	History

Appendix 4.10 – Pericarditis code list

Read Code	Description	Definite/possible/history of
AB40300	histoplasma capsulatum with pericarditis	Definite
AB41300	histoplasma duboisii with pericarditis	Definite
G501.00	post infarction pericarditis	Definite
G531.00	adhesive pericarditis	Definite
G531z00	adhesive pericarditis nos	Definite
G532.00	constrictive pericarditis	Definite
G532z00	constrictive pericarditis nos	Definite
Gyu5400	[x]pericarditis in other diseases classified elsewhere	Definite
N000400	systemic lupus erythematosus with pericarditis	Definite
A364100	meningococcal pericarditis	Definite
A742100	coxsackie pericarditis	Definite
A93y000	syphilitic pericarditis	Definite
A98y200	gonococcal pericarditis	Definite
G500000	acute pericarditis - coxsackie	Definite
G500100	acute pericarditis - meningococcal	Definite
G500300	acute pericarditis - tuberculous	Definite
G500311	tb - acute pericarditis	Definite
G500500	acute pericarditis - gonococcal	Definite
G50z111	viral pericarditis nos	Definite
G50z200	acute pericarditis - pneumococcal	Definite
G50z300	acute pericarditis - staphylococcal	Definite
G50z400	acute pericarditis - streptococcal	Definite
G010.00	acute rheumatic pericarditis	Definite
G50..00	acute pericarditis	Definite
G500.00	acute pericarditis in diseases ec	Definite
G500400	acute pericarditis - uraemic	Definite
G500z00	acute pericarditis in diseases ec nos	Definite
G50z.00	other and unspecified acute pericarditis	Definite
G50z000	acute pericarditis - unspecified	Definite
G50z100	acute idiopathic pericarditis	Definite
G50z500	acute purulent pericarditis unspecified	Definite
G50zz00	acute pericarditis nos	Definite
Gyu5000	[x]other forms of acute pericarditis	Definite
G10..00	chronic rheumatic pericarditis	Definite
G102.00	chronic rheumatic myopericarditis	Definite
G53yz11	chronic pericarditis	Definite

Appendix 4.11 – VHD code list

Read Code	Description	Definite/possible/history of
G121.00	rheumatic aortic insufficiency	Definite
G121.12	aortic regurgitation - rheumatic	Definite
G122.00	rheumatic aortic stenosis with insufficiency	Definite
G541011	aortic insufficiency, non-rheumatic	Definite
G541012	aortic regurgitation, non-rheumatic	Definite
G541211	aortic insufficiency alone, cause unspecified	Definite
G541212	aortic regurgitation alone, cause unspecified	Definite
G541400	aortic valve stenosis with insufficiency	Definite
P63..00	congenital aortic valve stenosis	Definite
P64..00	congenital aortic valve insufficiency	Definite
P640.00	congenital aortic valve insufficiency, unspecified	Definite
P64z.00	congenital aortic valve insufficiency nos	Definite
G111.00	rheumatic mitral insufficiency	Definite
G111.12	mitral regurgitation - rheumatic	Definite
G112.00	mitral stenosis with insufficiency	Definite
G112.13	mitral stenosis with regurgitation	Definite
G113.00	nonrheumatic mitral valve stenosis	Definite
G131.00	mitral stenosis and aortic insufficiency	Definite
G131.14	mitral stenosis and aortic regurgitation	Definite
G132.00	mitral insufficiency and aortic stenosis	Definite
G132.13	mitral regurgitation and aortic stenosis	Definite
G133.11	mitral and aortic insufficiency	Definite
G133.12	mitral and aortic regurgitation	Definite
G540.12	mitral valve insufficiency	Definite
G540.14	mitral valve regurgitation	Definite
G540.16	mitral regurgitation	Definite
P66..00	congenital mitral insufficiency	Definite
G141100	rheumatic pulmonary insufficiency	Definite
G141200	rheumatic pulmonary stenosis and insufficiency	Definite
G543011	pulmonary insufficiency, non-rheumatic	Definite
G543012	pulmonary regurgitation, non-rheumatic	Definite
G543213	pulmonary insufficiency, cause unspecified	Definite
G543215	pulmonary regurgitation, cause unspecified	Definite
G543400	pulmonary valve stenosis with insufficiency	Definite
H585.00	trauma and post-operative pulmonary insufficiency	Definite
H585200	pulmonary insufficiency following trauma	Definite
G140100	rheumatic tricuspid insufficiency	Definite
G140111	tricuspid regurgitation - rheumatic	Definite
G140200	rheumatic tricuspid stenosis and insufficiency	Definite
G14021X	rheumatic tricuspid stenosis and regurgitation	Definite
G140400	tricuspid insufficiency, cause unspecified	Definite
G140413	tricuspid regurgitation, cause unspecified	Definite
G140500	tricuspid stenosis and insufficiency, cause unspecified	Definite
G140514	tricuspid stenosis and regurgitation, cause unspecified	Definite
G542011	tricuspid insufficiency, non-rheumatic	Definite
G542012	tricuspid regurgitation, non-rheumatic	Definite
G542200	nonrheumatic tricuspid valve stenosis with insufficiency	Definite
G54z500	valvular heart disease	Definite

Appendix 4.12 – DVT code list

Read Code	Description	Definite/possible/history of
1JH..00	suspected deep vein thrombosis	Possible
8HTm.00	referral to deep vein thrombosis clinic	Possible
G801.11	deep vein thrombosis	Definite
G801.13	dvt - deep vein thrombosis	Definite
G801900	thrombophlebitis of the dorsalis pedis vein	Definite
G801D00	deep vein thrombosis of lower limb	Definite
G801G00	recurrent deep vein thrombosis	Definite
G801H00	unprovoked deep vein thrombosis	Definite
G801J00	provoked deep vein thrombosis	Definite
G80y.11	phlebitis and/or thrombophlebitis of iliac vein	Definite
G80y400	thrombophlebitis of the common iliac vein	Definite
G80y500	thrombophlebitis of the internal iliac vein	Definite
G80y600	thrombophlebitis of the external iliac vein	Definite
G80y700	thrombophlebitis of the iliac vein unspecified	Definite
G80y800	phlebitis and thrombophlebitis of the iliac vein nos	Definite
G81..00	portal vein thrombosis	Definite
G820.11	hepatic vein thrombosis	Definite
G822.00	embolism and thrombosis of the vena cava	Definite
G822000	thrombosis of inferior vena cava	Definite
J420200	thrombus of the superior mesenteric veins	Definite
SP12200	post operative deep vein thrombosis	Definite
G801.00	deep vein phlebitis and thrombophlebitis of the leg	Definite
G801.12	deep vein thrombosis, leg	Definite
G801600	thrombophlebitis of the femoral vein	Definite
G801700	thrombophlebitis of the popliteal vein	Definite
G801800	thrombophlebitis of the anterior tibial vein	Definite
G801A00	thrombophlebitis of the posterior tibial vein	Definite
G801B00	deep vein thrombophlebitis of the leg unspecified	Definite
G801C00	deep vein thrombosis of leg related to air travel	Definite
G801E00	deep vein thrombosis of leg related to intravenous drug use	Definite
G801F00	deep vein thrombosis of peroneal vein	Definite
G801z00	deep vein phlebitis and thrombophlebitis of the leg nos	Definite
F05..00	phlebitis and thrombophlebitis of intracranial sinuses	Definite
F051.00	thrombosis of central nervous system venous sinuses	Definite
F051z00	thrombosis of central nervous system venous sinus nos	Definite
F053.00	thrombophlebitis of central nervous system venous sinuses	Definite
F053000	thrombophlebitis of cavernous sinus	Definite
F053100	thrombophlebitis of superior longitudinal venous sinus	Definite
F05z.00	phlebitis or thrombophlebitis of cns venous sinus nos	Definite
G676.00	nonpyogenic venous sinus thrombosis	Definite
14A8100	h/o: deep vein thrombosis	History
ZV12800	[v] personal history deep vein thrombosis	History
ZV12811	[v] personal history dvt- deep vein thrombosis	History

Appendix 4.13 – Pulmonary Embolism code list

Read Code	Description	Definite/possible/history of
G401.00	pulmonary embolism	Definite
L43..00	obstetric pulmonary embolism	Definite
G401.12	pulmonary embolus	Definite
G401000	post operative pulmonary embolus	Definite
L430.00	obstetric air pulmonary embolism	Definite
7A09311	trendelenburg pulmonary embolectomy	Definite
L431.00	amniotic fluid pulmonary embolism	Definite
L432.00	obstetric blood-clot pulmonary embolism	Definite
L096400	pulmonary embolism following abortive pregnancy	Definite
L43..11	obstetric pulmonary embolus	Definite
L43z.00	obstetric pulmonary embolism nos	Definite
L431100	amniotic fluid pulmonary embolism - delivered	Definite
G401100	recurrent pulmonary embolism	Definite
L43z100	obstetric pulmonary embolism nos - delivered	Definite
L43zz00	obstetric pulmonary embolism nos	Definite
L43z000	obstetric pulmonary embolism nos, unspecified	Definite
L43yz00	other obstetric pulmonary embolism nos	Definite
1JC..00	suspected pulmonary embolism	Possible
14AC.00	h/o: pulmonary embolus	History
ZV12900	[v] personal history of pulmonary embolism	History

Appendix 4.14 – HES ICD-10 and OPCS-4 code list for CVD outcomes

Outcome	ICD-10/OPCS-4 codes
Angina	I20, I20.0, I20.1, I20.8, I20.9
MI	I21, I21.0, I21.1, I21.2, I21.3, I21.4, I21.9
Revascularisation	K234, K40, K401, K402, K403, K404, K408, K409, K41, K411, K412, K413, K414, K418, K419, K42, K421, K422, K423, K424, K428, K429, K46, K468, K469, K493, L16, L161, L162, L163, L168, L169, L20, L201, L202, L203, L204, L205, L206, L208, L209, L21, L211, L212, L213, L214, L215, L216, L218, L219, L292, L293, L296, L297, L371, L412, L451, L50, L501, L502, L503, L504, L505, L506, L508, L509, L51, L511, L512, L513, L514, L515, L516, L518, L519, L58, L581, L582, L583, L584, L585, L586, L587, L588, L589, L59, L591, L592, L593, L594, L595, L596, L597, L598, L599, L81, L818, L819, L831, Y731
SCA	I46, I46.0, I46.1, I46.9
PVD	I70.2, I70.20, I70.21, I73, I73.8, I73.9, I79.2
Stroke	I60, I60.0, I60.1, I60.2, I60.3, I60.4, I60.5, I60.6, I60.7, I60.8, I60.9, I61, I61.0, I61.1, I61.2, I61.3, I61.4, I61.5, I61.6, I61.8, I61.9, I62, I62.0, I62.1, I62.9, I63, I63.0, I63.1, I63.2, I63.3, I63.4, I63.5, I63.6, I63.8, I63.9, I64
Arrhythmia	I48, I49, I49.0, I49.1, I49.2, I49.3, I49.4, I49.5, I49.8, I49.9
HF	I25.5, I42, I42.0, I42.1, I42.2, I42.3, I42.4, I42.5, I42.6, I42.7, I42.8, I42.9, I43, I43.0, I43.1, I43.2, I43.8, I50, I50.0, I50.1, I50.9
Pericarditis	I30, I30.0, I30.1, I30.8, I30.9, I31, I31.0, I31.1, I31.2, I31.3, I31.8, I31.9, I32, I32.0, I32.1, I32.8
VHD	I05, I05.0, I05.1, I05.2, I05.8, I05.9, I06, I06.0, I06.1, I06.2, I06.8, I06.9, I07, I07.0, I07.1, I07.2, I07.8, I07.9, I08, I08.0, I08.1, I08.2, I08.3, I08.8, I08.9, I34, I34.0, I34.1, I34.2, I34.8, I34.9, I35, I35.0, I35.1, I35.2, I35.8, I35.9, I36, I36.0, I36.1, I36.2, I36.8, I36.9, I37, I37.0, I37.1, I37.2, I37.8, I37.9
DVT	I80.1, I80.2, I80.3
Pulmonary Embolism	I26, I26.0, I26.9

Appendix 4.15 - Concordance between events identified in CPRD and HES for the outcomes in which new codes were identified

Outcome	Before or after code list update	Women in study population	Incident events when use both CPRD and HES (% of study population)	Incident events identified form CPRD (% of total events)	Incident events identified form HES (% of total events)	Overlap between CPRD and HES (% of total events)	Median time between events in CPRD and HES (days)
Arrhythmia	Before	9060	542 (5.98)	344 (63.47)	359 (66.24)	184 (33.95)	102.5
	After	8445	679 (8.04)	507 (74.67)	331 (48.75)	188 (27.69)	203.5
Pericarditis	Before	9986	22 (0.22)	1 (4.55)	22 (100.00)	1 (4.55)	12
	After	9983	27 (0.27)	9 (33.33)	22 (81.48)	4 (14.81)	12.5
Valvular Heart Disease	Before	9755	223 (2.29)	70 (31.39)	166 (74.44)	18 (8.07)	311.5
	After	9729	232 (2.38)	107 (46.12)	157 (67.67)	39 (16.81)	214
Revascularisation	Before	9916	44 (0.44)	21 (47.73)	37 (84.09)	14 (31.82)	5
	After	9901	47 (0.47)	35 (74.47)	37 (78.72)	25 (53.19)	3

Appendix 4.16 - Proportion of primary diagnoses in HES events, and corresponding primary diagnoses if event was a secondary diagnosis

Outcome (ICD letter of outcome)	Total number of events in HES	Number of HES events that were the primary diagnosis (proportion of all HES events)	ICD chapter of primary diagnosis, if event was secondary diagnosis	Number of secondary diagnoses with same ICD chapter of primary diagnosis (proportion of all secondary diagnoses)
Angina (I)	167	42 (25.10)	A	2 (1.60)
			C	8 (6.40)
			D	2 (1.60)
			E	5 (4.00)
			G	1 (0.80)
			H	9 (7.20)
			I	36 (28.80)
			J	10 (8.00)
			K	11 (8.80)
			M	9 (7.20)
			N	9 (7.20)
			R	15 (12.00)
			S	7 (5.60)
			Z	1 (0.80)
MI (I)	132	79 (59.80)	A	1 (1.90)
			C	4 (7.50)
			D	1 (1.90)
			G	2 (3.80)
			I	19 (35.80)
			J	5 (9.40)
			K	3 (5.70)
			M	2 (3.80)
			R	11 (20.80)
			S	5 (9.40)
SCA (I)	38	10 (26.30)	A	1 (3.60)
			C	1 (3.60)
			D	1 (3.60)
			E	1 (3.60)
			G	1 (3.60)
			I	8 (28.60)
			J	7 (25.00)
			K	4 (14.30)
			N	1 (3.60)
			R	2 (7.10)
			T	1 (3.60)
PVD (I)	41	10 (24.40)	A	1 (3.20)
			C	1 (3.20)
			D	1 (3.20)
			E	1 (3.20)
			G	1 (3.20)
			H	2 (6.50)
			I	10 (32.30)
			J	1 (3.20)
			K	2 (6.50)
			L	1 (3.20)
			M	3 (9.70)
			N	1 (3.20)
			R	4 (12.90)
			T	1 (3.20)
			Z	1 (3.20)
Stroke (I)	197	137 (69.50)	C	5 (8.30)
			E	1 (1.70)
			G	5 (8.30)
			H	2 (3.30)
			I	16 (26.70)
			J	8 (13.30)
			M	4 (6.70)
			N	1 (1.70)
			R	10 (16.70)

			S	7 (11.70)
			T	1 (1.70)
Arrhythmia (I)	416	68 (16.30)	A	12 (3.40)
			C	26 (7.50)
			D	6 (1.70)
			E	7 (2.00)
			F	2 (0.60)
			G	3 (0.90)
			H	4 (1.10)
			I	80 (23.00)
			J	57 (16.40)
			K	31 (8.90)
			L	7 (2.00)
			M	19 (5.50)
			N	26 (7.50)
			R	43 (12.40)
			S	20 (5.70)
			T	4 (1.10)
			Z	1 (0.30)
HF (I)	322	75 (23.30)	A	10 (4.00)
			C	11 (4.50)
			E	7 (2.80)
			G	4 (1.60)
			H	2 (0.80)
			I	67 (27.10)
			J	46 (18.60)
			K	16 (6.50)
			L	5 (2.00)
			M	12 (4.90)
			N	17 (6.90)
			R	29 (11.70)
			S	16 (6.50)
			T	4 (1.60)
			Z	1 (0.40)
Pericarditis (I)	33	3 (9.10)	A	1 (3.30)
			C	8 (26.70)
			I	9 (30.00)
			J	6 (20.00)
			K	1 (3.30)
			N	1 (3.30)
			R	3 (10.00)
			T	1 (3.30)
VHD (I)	192	17 (8.90)	A	3 (1.70)
			C	4 (2.30)
			E	2 (1.10)
			F	1 (0.60)
			G	1 (0.60)
			H	6 (3.40)
			I	71 (40.60)
			J	24 (13.70)
			K	8 (4.60)
			L	2 (1.10)
			M	12 (6.90)
			N	4 (2.30)
			Q	1 (0.60)
			R	21 (12.00)
			S	11 (6.30)
			T	3 (1.70)
			Z	1 (0.60)
DVT (I)	100	57 (57.00)	A	1 (2.30)
			C	7 (16.30)
			D	1 (2.30)
			G	1 (2.30)
			I	12 (27.90)
			J	2 (4.70)
			K	4 (9.30)
			L	2 (4.70)
			M	4 (9.30)
			N	3 (7.00)

			R	4 (9.30)
			S	1 (2.30)
			Z	1 (2.30)
Pulmonary Embolism (I)	147	79 (53.70)	A	4 (5.90)
			B	1 (1.50)
			C	12 (17.60)
			D	2 (2.90)
			G	1 (1.50)
			I	10 (14.70)
			J	10 (14.70)
			K	5 (7.40)
			M	6 (8.80)
			N	3 (4.40)
			R	11 (16.20)
			S	2 (2.90)
			T	1 (1.50)

CHAPTER 5

Appendix 5.1 – Defining a breast cancer diagnosis

Code list is available at https://doi.org/10.17037/DATA.177. All women with a breast cancer diagnosis in the study period were identified in order to define the study population. A code list generated for a previous project was used, which included all breast cancer Read codes that fall under the C50 ICD-10 code. Clinical and referral files were then used to search for incident breast cancer diagnoses within the study period, that were more than a year after the women had entered the CPRD system (to ensure an incident event).

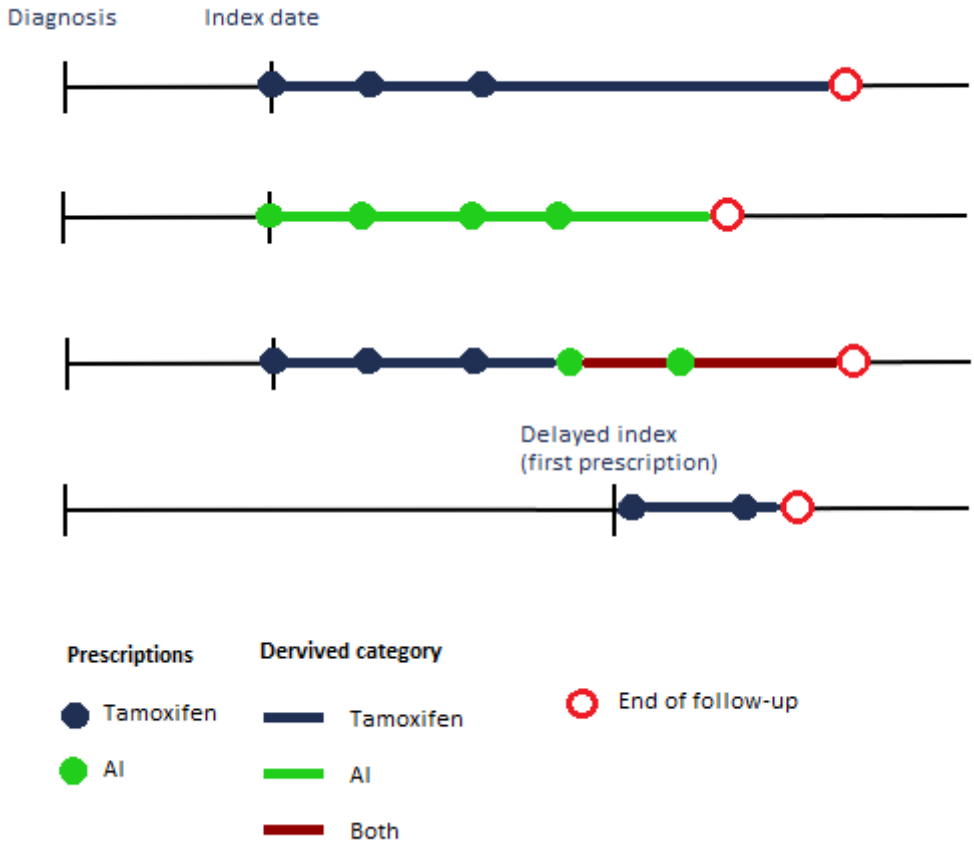
Appendix 5.2 – Identification of drug prescription codes

CPRD assigns product codes to drugs prescribed by GPs. Information on the product name, British National Formulary (BNF) header, drug substance, strength, formulation, and route by which drug is given are also provided in the CPRD therapy file.

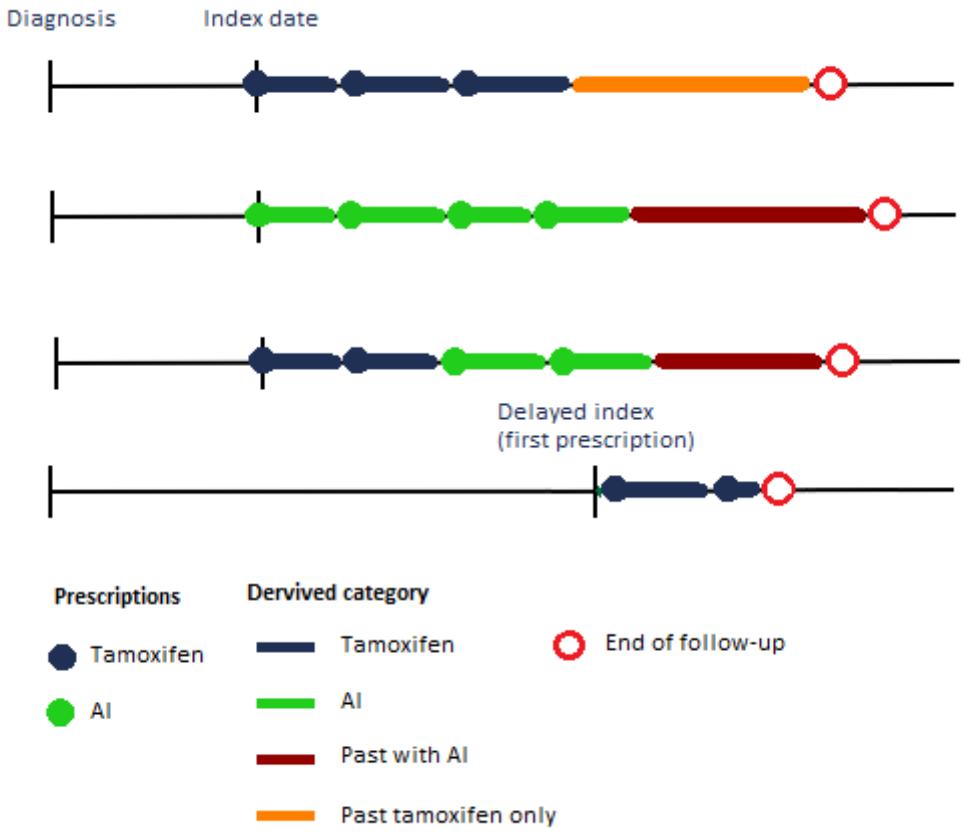
The following algorithm was used to define code lists for drug prescriptions:

- A list of inclusion search terms, which were synonyms of the drug and any related brand names, was agreed with the clinicians involved in the study
- The CPRD product code dictionary was then searched to identify any codes with one of the search terms in the product name or drug substance field
- Using a list of pre-specified terms and scanning the codes initially identified, a list of exclusion terms was created and applied
- Codes and their descriptors were manually reviewed to decide if they were appropriate for the final code list

Appendix 5.3 - Ever exposure to endocrine therapy categorisation



Appendix 5.4 - Current exposure to endocrine therapy categorisation



Appendix 5.5 – Defining length endocrine therapy prescriptions

Information on the quantity of drugs prescribed and the recommended number of pills to take per day (numeric daily dose, ndd) is in the therapy file of CPRD. From this, it is possible to calculate the intended length of prescription (ndd multiplied by quantity). However, some of the entries within these fields are unreliable. The following sections therefore explore the quantity and ndd of tamoxifen and AI prescriptions to identify unreliable entries and replace them with entries that are more realistic.

Refinement of tamoxifen prescriptions

There were 146,670 tamoxifen prescriptions given to those women within the linked CPRD and HES study population. The ndd of these prescriptions ranged from 0 to 20, with a median ndd of 1. As 99.95% of prescriptions had an ndd of 10 or less, and it is impossible to have an ndd of 0, all prescriptions with an ndd of 0 or greater than 10 were imputed with the median of 1.

The quantity of drugs prescribed ranged from 1 to 900, with a median of 30. It is unlikely that tamoxifen is prescribed in quantities of less than 7 (0.16% of prescriptions were given in a quantity of less than 7) and an extremely high quantity of drugs are very unlikely (99.45% of prescriptions were given in a quantity of 300 or less). All prescriptions with quantities of less than 7, or more than 300 were therefore imputed with the median quantity of 30 days.

Refinement of AI prescriptions

There were 244,776 AI prescriptions given to those women within the linked CPRD and HES study population. The ndd of these prescriptions ranged from 0 to 12, with a median ndd of 1. As 99.99% of prescriptions had an ndd of 10 or less, and it is impossible to have an ndd of 0, all prescriptions with an ndd of 0 or greater than 10 were imputed with the median of 1, similar to tamoxifen prescriptions above.

The quantity of drugs prescribed ranged from 1 to 924 for each prescription, with a median of 28. As with tamoxifen, it is unlikely that AIs are prescribed in quantities of less than 7 (0.03% of prescriptions were given in a quantity of less than 7), or in extremely high quantities (99.93% of prescriptions were given in a quantity of 90 or less). All prescriptions with quantities of less than 7, or more than 90 were therefore imputed with the median quantity of 28 days.

Appendix 5.6 – Identification of clinical diagnosis codes

CPRD

CPRD uses Read codes to identify clinical events that have been diagnosed in primary care. Read codes are a coded thesaurus of clinical terms that have been used in the NHS since 1985, which provide a standard vocabulary by which clinicians can record patient findings and procedures in health and social care IT systems across primary and secondary care. Creation of Read code lists representing a certain disease allows identification of patients in CPRD with a diagnosis of that disease by merging the code lists with the patients' raw data files.

A systematic approach was used to define code lists in order to identify clinical diagnoses. The relevant dictionaries of codes were searched using STATA do files,[1] so that all decisions on inclusion and exclusion criteria were recorded and easily replicated.

A code list for each clinical diagnosis was created using the following algorithm:

- A list of inclusion search terms, which were synonyms of the medical event, was agreed through discussion with the clinical collaborators involved in the study.
- The CPRD Read code dictionary was then searched to identify any codes with one of the search terms in the read term field, which is used to describe the Read code
- Using a list of pre-specified terms and scanning the codes initially identified, a list of exclusion terms was created and applied
- Codes and their descriptors were manually reviewed to decide if they were appropriate for the final code list

HES

HES uses ICD-10 codes to identify clinical events diagnosed in secondary care. ICD-10 codes are a comprehensive classification of causes of morbidity and mortality that is published by the World Health Organisation. The 10th revision of ICD codes was published in April 1995, and followed the 9th revision (ICD-9) that was published in 1975 and came into use in hospital health systems in 1979. Furthermore, OPCS-4 codes are used to classify interventions and procedures, and were originally published in 1987 by the Office of Population Censuses and Surveys, and was followed by a 4th revision in 1992.

CVD outcome events were identified in HES. Clinicians guided the creation of ICD-10 code lists relating to all CVD outcomes. OPCS-4 codes lists were also created to identify revascularisation procedures carried out in secondary care. These codes lists were then used to search for relevant diagnoses in the patients' raw HES data.

Appendix 5.7 – Defining covariates

All code lists are available at https://doi.org/10.17037/DATA.177.

Diagnoses

Code lists for clinical diagnoses were created using the algorithm for CPRD outlined in Appendix 6 above. Once code lists were created, the following steps were taken to define covariates in the study population.

Diabetes and rheumatoid arthritis

Read code lists were used to search CPRD clinical and referral files to identify any women in the study population with a diabetes or rheumatoid arthritis diagnosis prior to index date.

Chronic kidney disease

Records of an eGFR reading by GPs were identified in the CPRD additional files. CKD was then established by calculating eGFR using the Chronic Kidney Disease Epidemiology Collaboration equation.[2] Serum creatinine measurements were not routinely isotope-dilution mass spectrometry-standardised until 2013. It was therefore assumed that all creatinine results were unstandardised and multiplied results with a correction factor of 0.95 before calculating eGFR without regard to ethnicity.[3] To avoid selection bias, an absent CKD category was included for those with no recorded serum creatinine result. The recording closest to, but before index date was used.

Systolic and diastolic blood pressure

No code lists were created to identify blood pressure as all information was in the CPRD additional file.

CPRD Additional files were searched for instances of when blood pressure was recorded by the GP. The following algorithm to identified blood pressure records:

- Drop record if systolic or diastolic <30 or >250
- Drop if the record indicates it is a target blood pressure
- Drop a duplicate if they indicate the same reading on the same day
- If there are more than once sensible reading in the same day, calculate the mean

The following algorithm then assigned a systolic and diastolic blood pressure recording to all women with relevant records:

- Take the nearest status in the period -1y to +1month from index if available (best option)
- if not, then take nearest in the period +1month to +1y after index if available (second best option)
- if not, then take any nearest before -1y from index if available (third best option) if not, then take nearest after +1y from index (least best option)

Blood pressures were then categorised into the following categories.

- Systolic – low and ideal (<120), pre-high (>=120 & <140), high (>=140)
- Diastolic – low and ideal (<80), pre-high (>=80 & <90), high (>=90)

Drug Prescriptions

The algorithm presented in Appendix 2 was used to define drug prescription code lists.

Once code lists were created, the CPRD therapy files were searched to identify women in the study population with a prescription of any of statins, ACE inhibitors, calcium channel blockers, angiotensin II receptor blockers, anti-platelets, or HRT prescriptions prior to index date

Lifestyle measures

A similar algorithm to that outlined for diagnoses above was used to create codes lists to identify lifestyle measures. The following steps were then taken to define covariates in the study population.

Smoking status

A Read code list was created to search for all records of a smoking status in the CPRD clinical file. Further available data from the additional file were also extracted, including the patients smoking "status" (yes, no or ex) and the "number of cigarettes per day" smoked. The following algorithm was then used to assign a smoking status to all women with relevant records.

- Take the nearest status in the period -1y to +1month from index if available (best option)
- if not, then take nearest in the period +1month to +1y after index if available (second best option)

- if not, then take any nearest before -1y from index if available (third best option)
- if not, then take nearest after +1y from index (least best option)

Smoking status was then categorised into non-smoker, current smoker, and ex-smoker.

Alcohol status

A Read code list was created to identify all records of alcohol usage in the CPRD clinical file. Further available data from the additional file were also extracted, including the patient's alcohol drinking "status" (yes, no or ex) and the "units per week" consumed. The above algorithm used to define smoking status was also used to assign an alcohol status to all women with relevant records.

Alcohol status was then categorised into non-drinker, current drinker, ex-drinker

BMI

No read code list was created as all information was available in the CPRD additional files.

CPRD additional files were searched for records of women's height and weight. BMI is calculated using patients weight in kilograms / (height in meters)^2. The following algorithm, which was described in the paper by Bhaskaran et al. on BMI in the CPRD,[4] was used to identify for BMI records:

- drop if 3+ measurements on the same day
- if 2 measurements on same day: drop if >5cm (ht)/1kg (wt) difference, otherwise take the mean
- initial pass, drop weights less than 2kg, heights less than 2 feet
- drop records after end of follow-up but keeps those before start of follow-up
- later, drops weights < 20kg, heights less than 4 or more than 7 feet
- fill in missing heights using last observation carried forward or if no previous, first future height measurement
- calculate a version of BMI directly from height and weight
- drop BMIs <5 or >200 (but if CPRD and calculated version differ, and one is in the range 10-100, use the one within this range)

- in general, prioritises calculated BMI, and only uses CPRD version if cannot be calculated (as no height measurement available at all)

The above algorithm used to assign smoking status was then used to assign BMI status to all women with relevant records.

BMI was categorised according to adult BMI cut-offs defined by the WHO, as underweight or healthy weight (BMI <24.9), overweight (BMI 25+), or obese (BMI ≥30).[5]

Demographic measures

Age

As CPRD only supplies year of birth, the date of birth was set to 1st of July for all women. Age at index was then calculated and categorised into 54-59, 60-69, 70+ years.

Index of multiple deprivation (IMD)

CPRD supplies patient level IMD scores at the practice level for all patients. IMD combines a number of indicators, chosen to cover a range of economic, social and housing issues, into a single deprivation score.[6] This score was consolidated into quintiles, with a low quintile representing the least deprived, and a high quintile representing the most deprived.

Appendix 5.8 – Study design, covariate adjustment, and model selection alternatives

Study design

This study is a cohort study using prospectively collected data. It would also have been possible to set up a case control study within the CPRD. However, a cohort study enables the calculation of risk (which is informative for clinicians), allows modelling prospectively collected data and inclusion of time dependent variables (such as endocrine therapy treatment), and more suited to assessing causal relationships due to being less prone to bias (potential confounders measured at point of exposure in cohort studies instead of at outcome in non-nested case control studies). A cohort study was therefore more suited this research question and the data used in comparison with a case control study.

Covariate adjustment

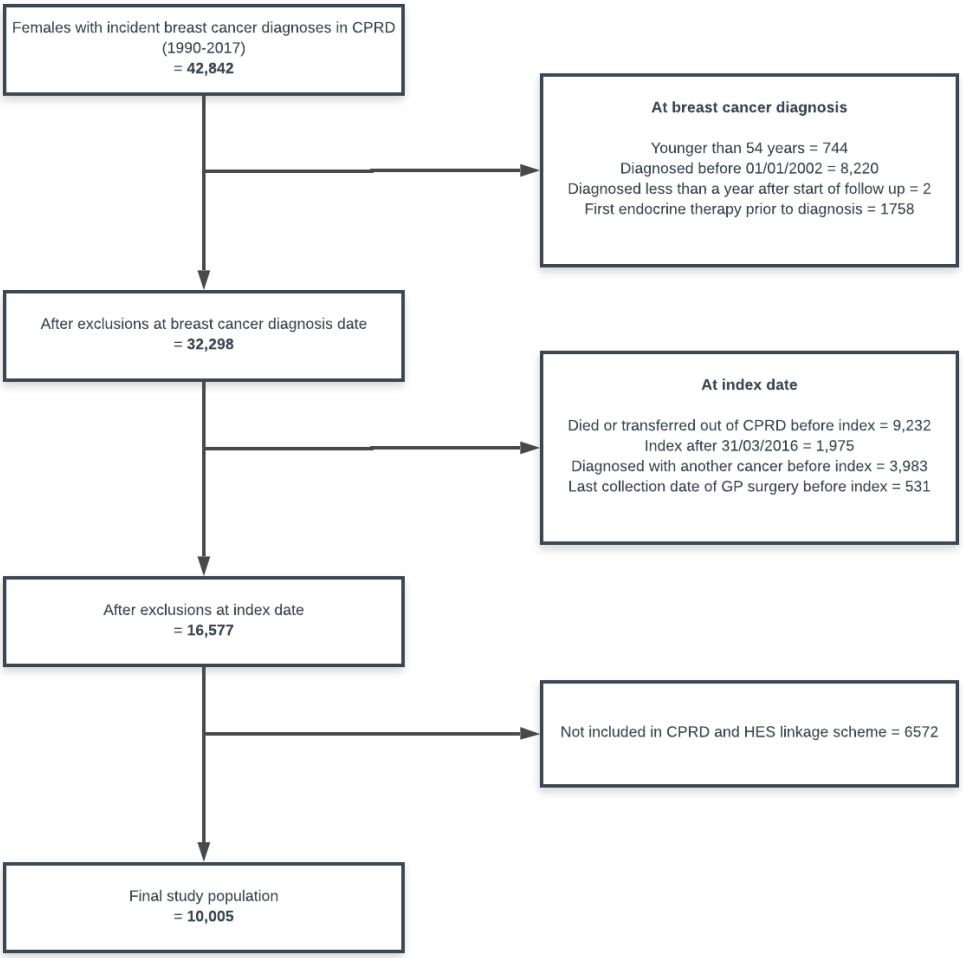
All potential confounders were measured at baseline and were adjusted for in Cox regression model. Another possible approach could have been propensity score matching. A propensity score is the conditional probability of being treated given the vector of observed covariates. The patients in each exposure arm can then be matched based on the propensity score, with the aim of obtaining an unbiased estimate of treatment effect adjusted for the impact of given confounding factors. However, the propensity score relies on two major assumptions: the observed variables do not affect the clinician's decision that a patient will be treated (e.g. lifestyle factors to not affect the choice of endocrine therapy); and there are no unmeasured confounders (all the covariates potentially related to treatment assignment are known). The propensity score was not used in this study because we did not have information on other cancer treatments (such as chemotherapy and trastuzumab), which could be potential confounders of the exposure outcome association, so the assumption of no unmeasured confounders was not satisfied.

Model selection

Cox regression was used to model the effect of endocrine therapies on the risk of a range of CVD outcomes. Cox regression is a method for investigating the effect of several variables upon the time a specified event takes to happen. It also allows time dependent and time fixed covariates, which were essential when assessing this association as women could change their exposure from AIs to tamoxifen, or vice versa, and their exposure was updated in the models to indicate this change.

Another regression strategy could have been Poisson regression. However, Poisson regression models count data, which is not suitable in this context as we were only modelling the effect of endocrine therapies on incident CVD events. Poisson regression would have been suitable if we further explored the effect of treatment on the risk of multiple CVD events. Furthermore, time-to-event techniques, such as survival analysis and Cox regression achieve greater efficiency and accuracy when modelling time varying covariates (such as exposure to endocrine therapy) in comparison with Poisson regression.

Appendix 5.9: STROBE flow diagram for inclusion in UK study



Appendix 5.10 - Crude rates, unadjusted HRs, and adjusted HRs by ever exposure to endocrine therapies for a range of clinical CVDs

Outcome	Ever exposure	Number of Events	Person years (1000s)	Crude Rate (per 1000 Person-Years)	Unadjusted HR (95% CI)	Adjusted HR (95% CI) *
Coronary Artery Disease	Tamoxifen	93	13.54	6.87 (5.60, 8.41)	1	1
	AI	131	12.87	10.18 (8.58, 12.08)	1.40 (1.06, 1.86)	1.29 (0.94, 1.76)
	Both	67	10.6	6.32 (4.98, 8.03)	0.95 (0.69, 1.31)	0.95 (0.68, 1.32)
Angina	Tamoxifen	56	13.73	4.08 (3.14, 5.30)	1	1
	AI	80	13.15	6.09 (4.89, 7.58)	1.43 (1.00, 2.05)	1.31 (0.88, 1.97)
	Both	31	10.73	2.89 (2.03, 4.11)	0.70 (0.44, 1.10)	0.69 (0.43, 1.10)
MI	Tamoxifen	32	14.6	2.19 (1.55, 3.10)	1	1
	AI	61	14.11	4.32 (3.36, 5.56)	1.73 (1.12, 2.68)	1.56 (0.96, 2.52)
	Both	39	11.24	3.47 (2.54, 4.75)	1.62 (1.01, 2.60)	1.62 (0.99, 2.63)
Revascularisation	Tamoxifen	15	14.72	1.02 (0.61, 1.69)	1	1
	AI	20	14.3	1.40 (0.90, 2.17)	1.36 (0.69, 2.68)	1.84 (0.85, 4.02)
	Both	12	11.33	1.06 (0.60, 1.86)	0.99 (0.46, 2.11)	1.01 (0.46, 2.22)
SCA	Tamoxifen	13	14.88	0.87 (0.51, 1.50)	1	1
	AI	21	14.53	1.45 (0.94, 2.22)	1.87 (0.83, 4.22)	1.65 (0.65, 4.19)
	Both	5	11.43	0.44 (0.18, 1.05)	0.73 (0.24, 2.19)	0.68 (0.22, 2.09)
PVD	Tamoxifen	35	14.64	2.39 (1.72, 3.33)	1	1
	AI	41	14.08	2.91 (2.14, 3.96)	1.25 (0.77, 2.01)	1.31 (0.76, 2.25)
	Both	22	11.23	1.96 (1.29, 2.97)	0.81 (0.45, 1.43)	0.86 (0.48, 1.57)
Stroke	Tamoxifen	91	14.45	6.30 (5.13, 7.73)	1	1
	AI	118	13.71	8.61 (7.19, 10.31)	1.22 (0.91, 1.63)	1.11 (0.81, 1.52)
	Both	88	10.98	8.02 (6.51, 9.88)	1.33 (0.98, 1.80)	1.25 (0.91, 1.71)
Arrhythmia	Tamoxifen	219	12.7	17.25 (15.11, 19.69)	1	1
	AI	287	11.52	24.90 (22.18, 27.96)	1.37 (1.14, 1.64)	1.37 (1.11, 1.68)
	Both	174	9.51	18.30 (15.77, 21.23)	1.07 (0.87, 1.32)	1.10 (0.89, 1.36)
HF	Tamoxifen	90	14.33	6.28 (5.11, 7.72)	1	1
	AI	178	13.59	13.10 (11.31, 15.17)	1.87 (1.43, 2.45)	1.68 (1.24, 2.26)
	Both	76	11.04	6.89 (5.50, 8.62)	1.12 (0.81, 1.55)	1.12 (0.80, 1.56)
Pericarditis	Tamoxifen	3	14.85	0.20 (0.07, 0.63)	1	1
	AI	14	14.53	0.96 (0.57, 1.63)	3.96 (1.12, 14.03)	3.25 (0.86, 12.23)
	Both	10	11.4	0.88 (0.47, 1.63)	4.29 (1.18, 15.61)	3.57 (0.95, 13.50)
VHD	Tamoxifen	66	14.54	4.54 (3.57, 5.78)	1	1
	AI	114	13.93	8.18 (6.81, 9.83)	1.58 (1.15, 2.17)	1.30 (0.92, 1.85)
	Both	52	11.14	4.67 (3.56, 6.13)	1.08 (0.74, 1.57)	0.98 (0.67, 1.43)
VTE	Tamoxifen	122	14.41	8.47 (7.09, 10.11)	1	1
	AI	116	13.4	8.66 (7.22, 10.38)	0.93 (0.71, 1.21)	0.82 (0.61, 1.10)
	Both	85	10.63	8.00 (6.47, 9.89)	0.91 (0.69, 1.22)	0.95 (0.71, 1.28)
DVT	Tamoxifen	83	14.52	5.72 (4.61, 7.09)	1	1
	AI	62	13.78	4.50 (3.51, 5.77)	0.69 (0.48, 0.98)	0.63 (0.42, 0.92)
	Both	59	10.83	5.45 (4.22, 7.03)	0.98 (0.69, 1.38)	1.04 (0.73, 1.49)
PE	Tamoxifen	48	14.75	3.25 (2.45, 4.32)	1	1
	AI	68	14.06	4.84 (3.81, 6.13)	1.39 (0.95, 2.04)	1.21 (0.79, 1.85)
	Both	33	11.19	2.95 (2.10, 4.15)	0.80 (0.51, 1.28)	0.78 (0.48, 1.25)

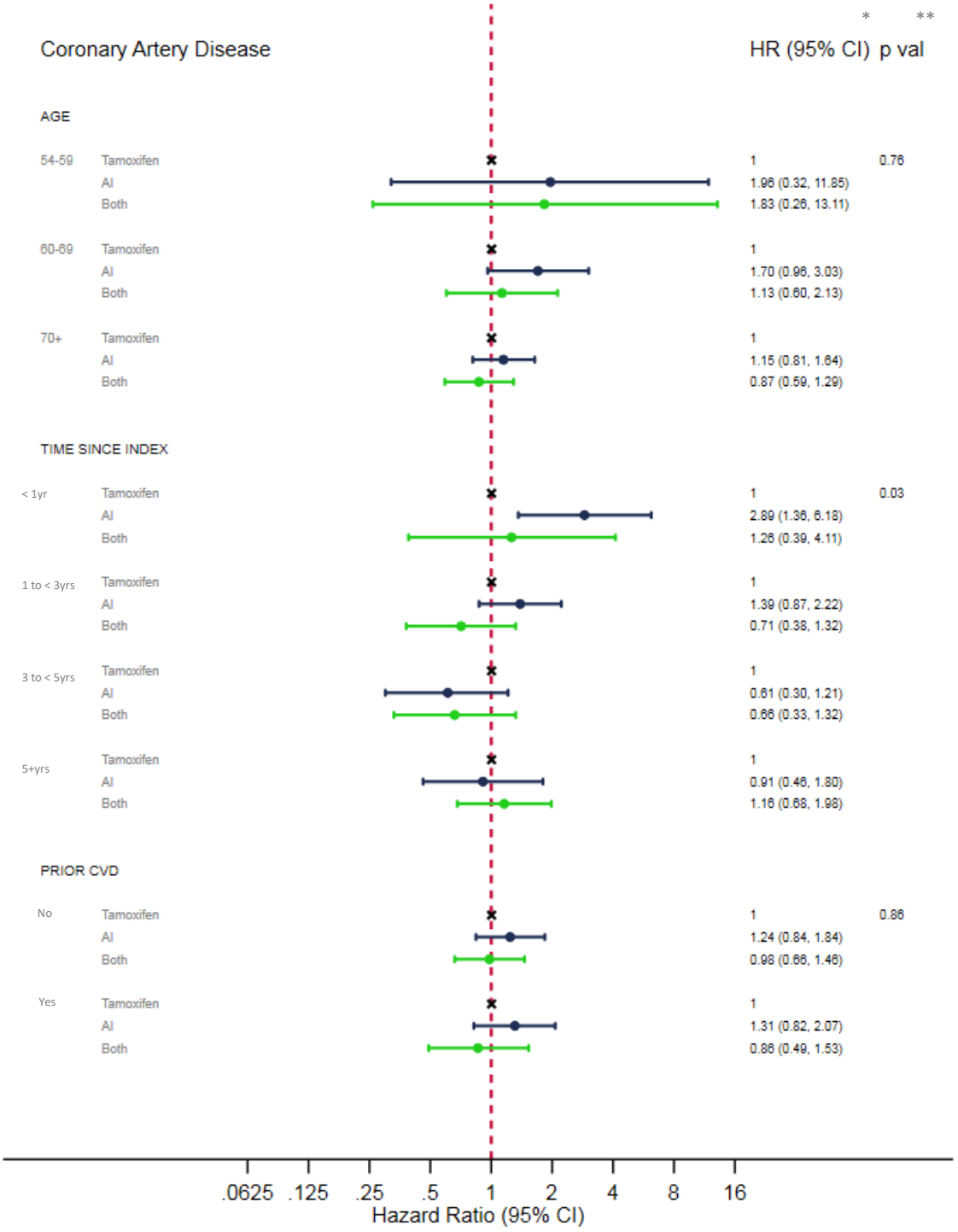
Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+);* smoking status (non-smoker, current smoker, ex-smoker)*;* BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker)*;* IMD score (level 1-5 based on GP level IMD data)*;* use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes*;* chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of VTE; history of non-venous CVD year of breast cancer diagnosis; time since index (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year.

Appendix 5.11 - Crude rates, unadjusted HRs, and adjusted HRs by current exposure to endocrine therapies for a range of clinical CVDs

Outcome	Ever exposure	Number of Events	Person years (1000s)	Crude Rate (per 1000 Person-Years)	Unadjusted HR (95% CI)	Adjusted HR (95% CI)*
Coronary Artery Disease	Tamoxifen	67	10.42	6.43 (5.06, 8.17)	1	1
	AI	137	15.57	8.80 (7.44, 10.41)	1.28 (0.94, 1.73)	1.19 (0.87, 1.64)
	Past with AI	49	6.49	7.55 (5.71, 9.99)	1.05 (0.72, 1.54)	1.09 (0.68, 1.74)
	Past Tam only	38	4.54	8.38 (6.10, 11.51)	1.10 (0.73, 1.67)	1.23 (0.74, 2.05)
Angina	Tamoxifen	40	10.55	3.79 (2.78, 5.17)	1	1
	AI	79	15.85	4.98 (4.00, 6.21)	1.21 (0.82, 1.80)	1.12 (0.74, 1.70)
	Past with AI	28	6.59	4.25 (2.94, 6.16)	0.99 (0.60, 1.63)	1.20 (0.66, 2.18)
	Past Tam only	20	4.61	4.34 (2.80, 6.73)	0.95 (0.54, 1.68)	1.25 (0.63, 2.44)
MI	Tamoxifen	27	11.16	2.42 (1.66, 3.53)	1	1
	AI	67	16.91	3.96 (3.12, 5.03)	1.39 (0.88, 2.18)	1.26 (0.78, 2.03)
	Past with AI	26	6.95	3.74 (2.55, 5.49)	1.28 (0.75, 2.20)	1.11 (0.57, 2.17)
	Past Tam only	12	4.92	2.44 (1.38, 4.29)	0.73 (0.36, 1.48)	0.67 (0.29, 1.52)
Revascularisation	Tamoxifen	9	11.28	0.80 (0.42, 1.53)	1	1
	AI	21	17.13	1.23 (0.80, 1.88)	1.47 (0.67, 3.24)	1.67 (0.72, 3.86)
	Past with AI	9	7.01	1.28 (0.67, 2.47)	1.38 (0.54, 3.49)	1.30 (0.41, 4.12)
	Past Tam only	8	4.94	1.62 (0.81, 3.24)	1.75 (0.67, 4.56)	1.61 (0.50, 5.18)
SCA	Tamoxifen	7	11.36	0.62 (0.29, 1.29)	1	1
	AI	20	17.35	1.15 (0.74, 1.79)	2.70 (0.90, 8.03)	2.28 (0.73, 7.11)
	Past with AI	5	7.1	0.70 (0.29, 1.69)	1.51 (0.37, 6.08)	1.53 (0.30, 7.80)
	Past Tam only	7	5.02	1.39 (0.66, 2.93)	3.01 (0.84, 10.73)	3.47 (0.70, 17.04)
PVD	Tamoxifen	26	11.19	2.32 (1.58, 3.41)	1	1
	AI	45	16.9	2.66 (1.99, 3.57)	1.27 (0.75, 2.18)	1.37 (0.78, 2.42)
	Past with AI	14	6.92	2.02 (1.20, 3.42)	0.92 (0.46, 1.86)	0.95 (0.40, 2.26)
	Past Tam only	13	4.95	2.63 (1.53, 4.53)	1.28 (0.63, 2.58)	1.47 (0.62, 3.49)
Stroke	Tamoxifen	63	11.08	5.68 (4.44, 7.28)	1	1
	AI	131	16.47	7.95 (6.70, 9.44)	1.25 (0.91, 1.72)	1.16 (0.83, 1.62)
	Past with AI	63	6.76	9.32 (7.28, 11.94)	1.57 (1.09, 2.27)	1.56 (1.00, 2.43)
	Past Tam only	40	4.83	8.28 (6.07, 11.29)	1.31 (0.87, 1.98)	1.40 (0.86, 2.27)
Arrhythmia	Tamoxifen	142	9.74	14.58 (12.37, 17.19)	1	1
	AI	317	14.2	22.32 (19.99, 24.92)	1.46 (1.18, 1.80)	1.45 (1.17, 1.81)
	Past with AI	124	5.64	21.98 (18.43, 26.21)	1.52 (1.18, 1.96)	1.86 (1.38, 2.52)
	Past Tam only	97	4.14	23.40 (19.18, 28.56)	1.56 (1.19, 2.04)	1.90 (1.38, 2.62)
HF	Tamoxifen	65	10.97	5.93 (4.65, 7.56)	1	1
	AI	169	16.37	10.32 (8.88, 12.00)	1.65 (1.21, 2.24)	1.48 (1.07, 2.04)
	Past with AI	72	6.78	10.62 (8.43, 13.38)	1.66 (1.16, 2.37)	1.71 (1.11, 2.63)
	Past Tam only	38	4.84	7.85 (5.71, 10.79)	1.19 (0.78, 1.82)	1.33 (0.81, 2.18)
Pericarditis	Tamoxifen	2	11.35	0.18 (0.04, 0.70)	1	1
	AI	17	17.34	0.98 (0.61, 1.58)	4.85 (1.11, 21.23)	3.61 (0.80, 16.27)
	Past with AI	7	7.09	0.99 (0.47, 2.07)	5.77 (1.19, 28.05)	3.09 (0.49, 19.59)
	Past Tam only	1	5.02	0.20 (0.03, 1.41)	1.19 (0.11, 13.18)	0.65 (0.05, 8.92)
VHD	Tamoxifen	37	11.12	3.33 (2.41, 4.59)	1	1
	AI	117	16.73	6.99 (5.83, 8.38)	1.84 (1.25, 2.70)	1.52 (1.02, 2.25)
	Past with AI	40	6.87	5.82 (4.27, 7.94)	1.46 (0.92, 2.34)	1.18 (0.68, 2.03)
	Past Tam only	38	4.9	7.76 (5.65, 10.67)	1.90 (1.18, 3.03)	1.66 (0.95, 2.91)
VTE	Tamoxifen	117	11.04	10.60 (8.84, 12.70)	1	1
	AI	110	16.07	6.84 (5.68, 8.25)	0.59 (0.45, 0.78)	0.58 (0.43, 0.77)
	Past with AI	69	6.51	10.60 (8.37, 13.42)	0.89 (0.65, 1.21)	1.22 (0.82, 1.80)
	Past Tam only	27	4.81	5.61 (3.85, 8.18)	0.48 (0.31, 0.74)	0.74 (0.45, 1.23)
DVT	Tamoxifen	82	11.12	7.38 (5.94, 9.16)	1	1
	AI	68	16.49	4.12 (3.25, 5.23)	0.50 (0.36, 0.71)	0.50 (0.35, 0.72)
	Past with AI	38	6.64	5.72 (4.16, 7.86)	0.69 (0.46, 1.03)	0.83 (0.49, 1.40)
	Past Tam only	16	4.87	3.28 (2.01, 5.36)	0.39 (0.23, 0.68)	0.50 (0.26, 0.98)
Pulmonary Embolism	Tamoxifen	43	11.29	3.81 (2.83, 5.14)	1	1
	AI	59	16.86	3.50 (2.71, 4.52)	0.84 (0.56, 1.26)	0.77 (0.50, 1.17)
	Past with AI	36	6.89	5.22 (3.77, 7.24)	1.24 (0.78, 1.96)	1.74 (0.99, 3.05)
	Past Tam only	11	4.96	2.22 (1.23, 4.01)	0.51 (0.26, 1.03)	0.84 (0.39, 1.83)

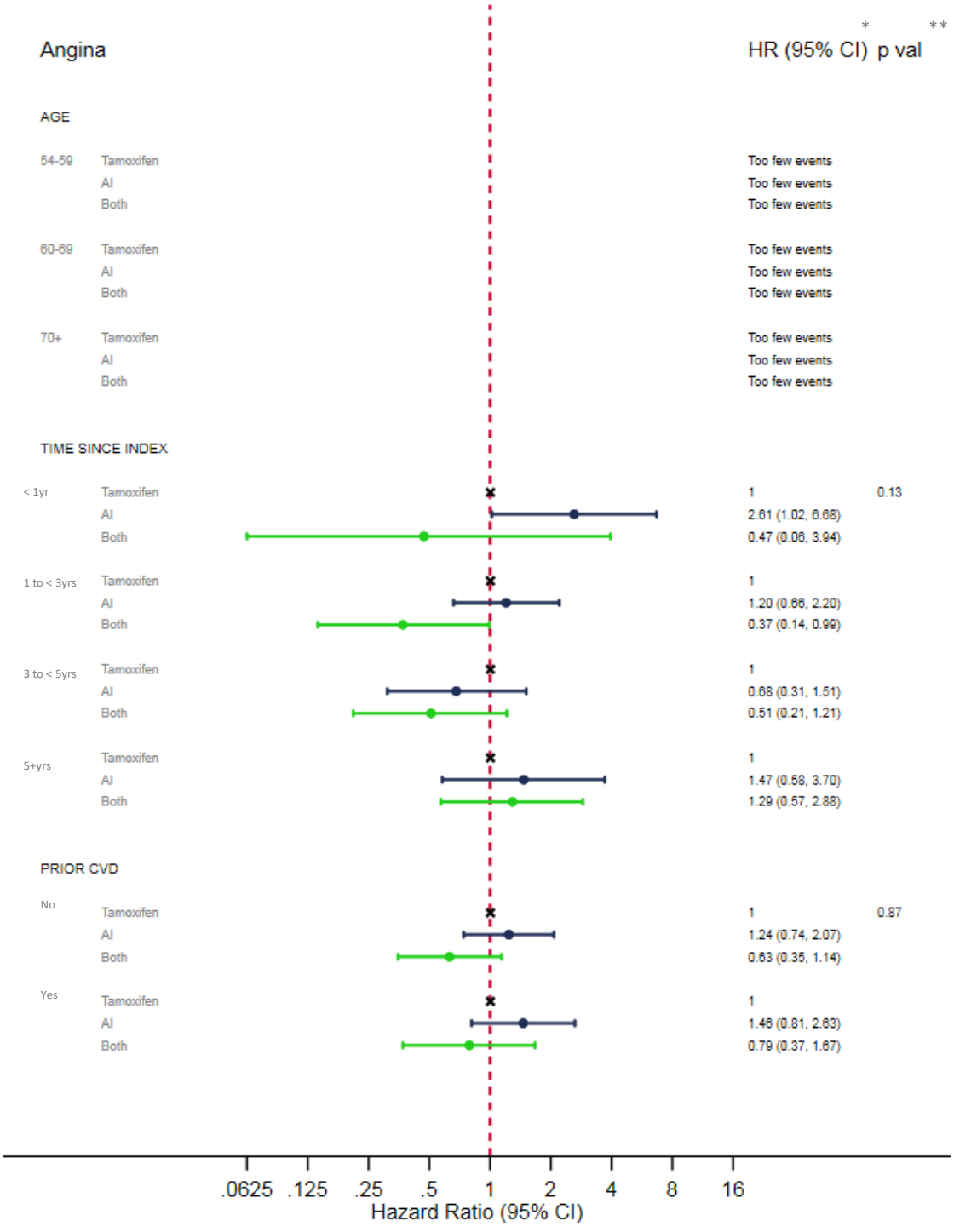
Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+);* smoking status (non-smoker, current smoker, ex-smoker)*;* BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker)*;* IMD score (level 1-5 based on GP level IMD data)*;* use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes*;* chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of VTE; history of non-venous CVD year of breast cancer diagnosis; time since index (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year.

Appendix 5.12 - Adjusted HRs for association between ever exposure to endocrine therapies and risk of coronary artery disease, stratified by age, time since index, and prior CVD



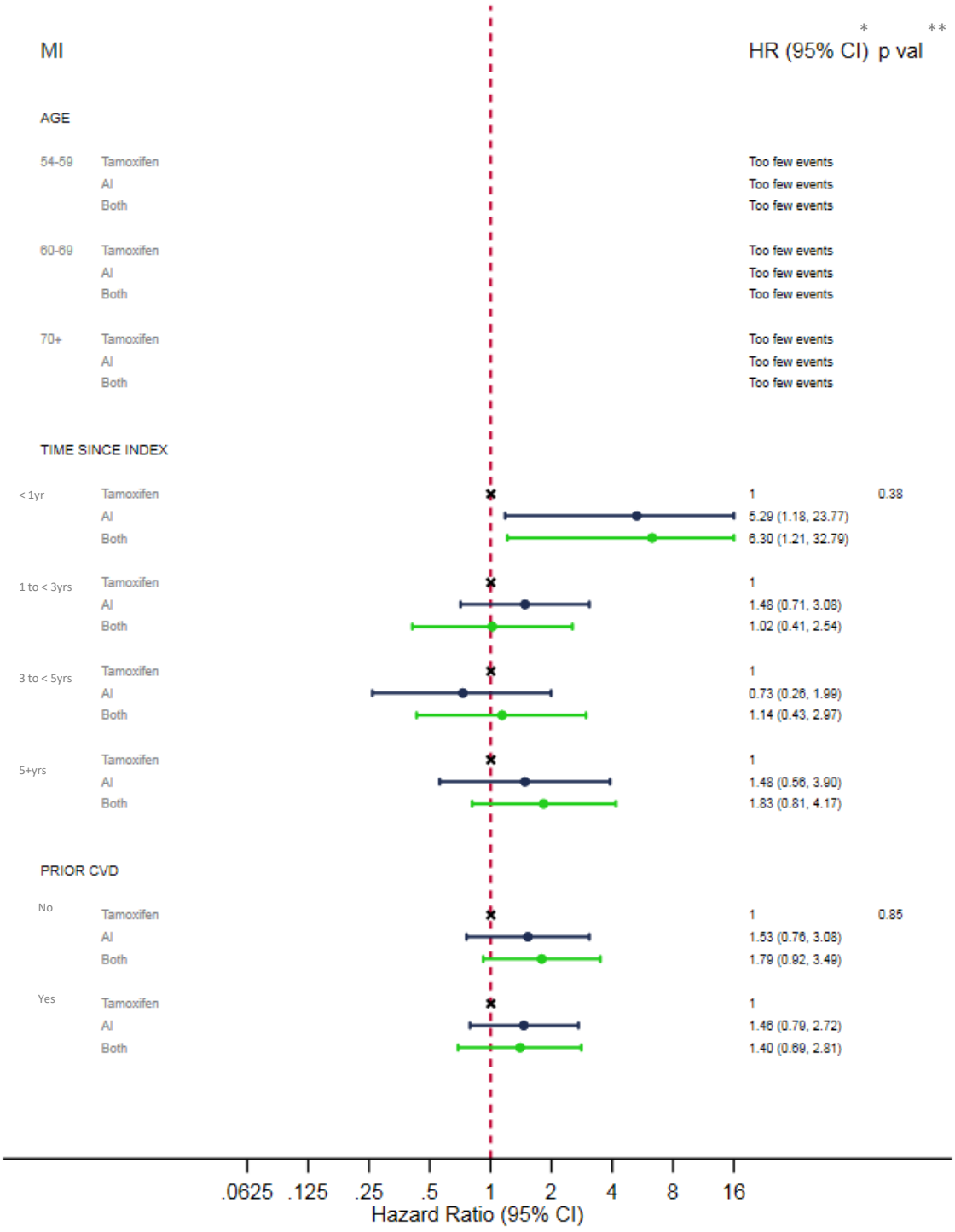
*Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+); smoking status (non-smoker, current smoker, ex-smoker); BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker); IMD score (level 1-5 based on GP level IMD data); use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes; chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of VTE; history of non-venous CVD year of breast cancer diagnosis; time since index (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year.
** p value from Wald test for interaction

Appendix 5.13 - Adjusted HRs for association between ever exposure to endocrine therapies and risk of angina, stratified by age, time since index, and prior CVD



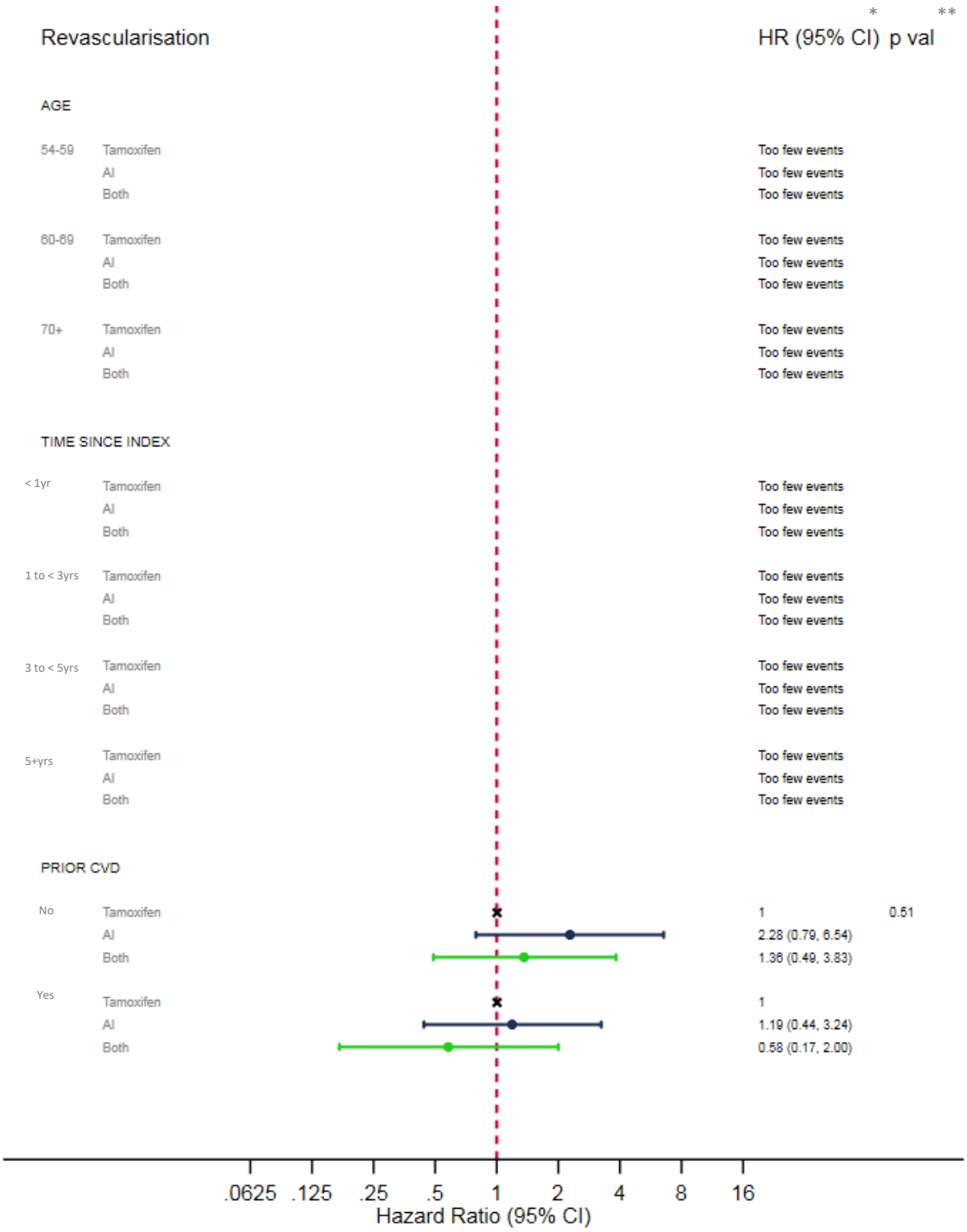
*Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+); smoking status (non-smoker, current smoker, ex-smoker); BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker); IMD score (level 1-5 based on GP level IMD data); use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes; chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of VTE; history of non-venous CVD year of breast cancer diagnosis; time since index (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year.
** p value from Wald test for interaction

Appendix 5.14 - Adjusted HRs for association between ever exposure to endocrine therapies and risk of MI, stratified by age, time since index, and prior CVD



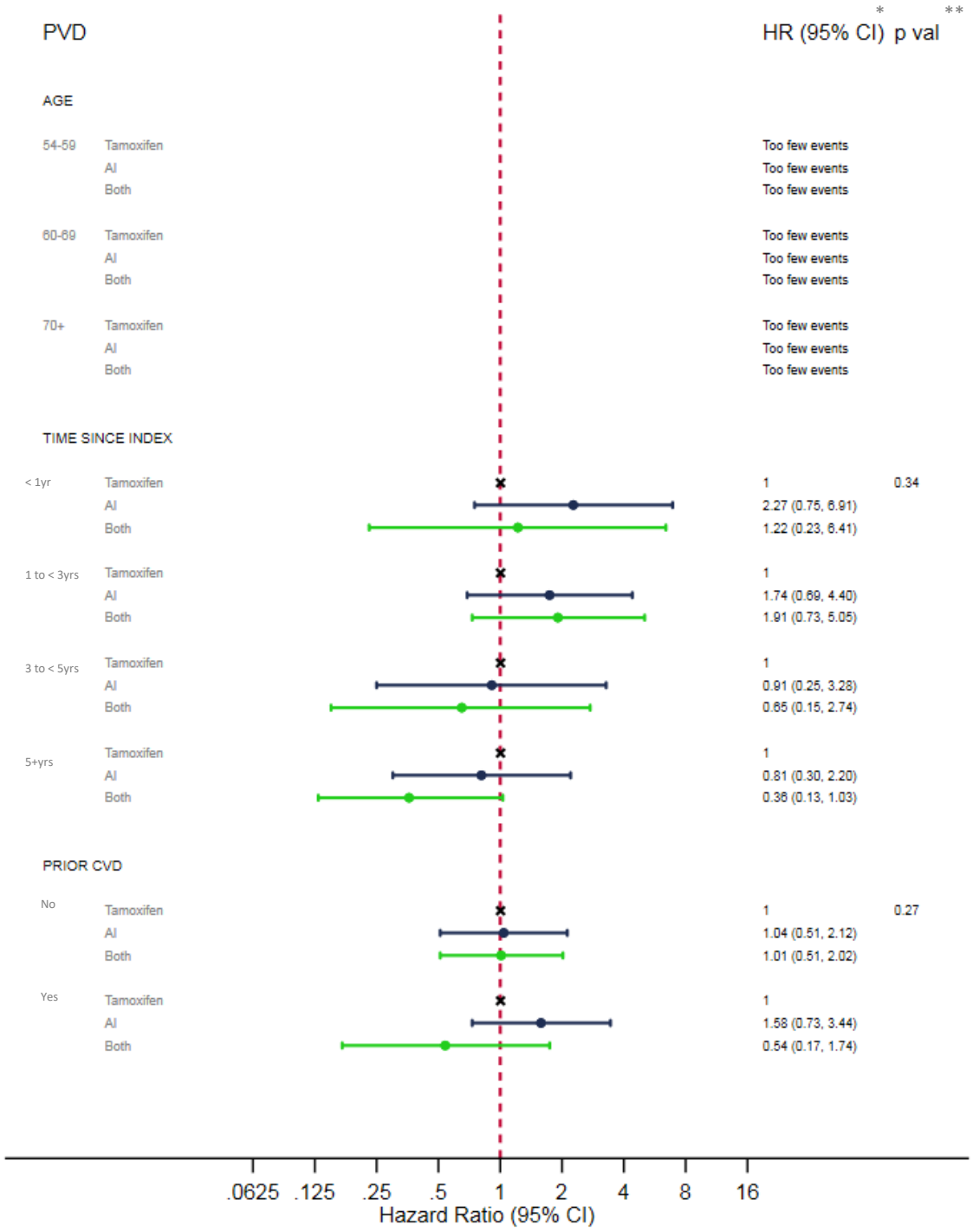
*Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+); smoking status (non-smoker, current smoker, ex-smoker); BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker); IMD score (level 1-5 based on GP level IMD data); use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes; chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of VTE; history of non-venous CVD year of breast cancer diagnosis; time since index (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year.
** p value from Wald test for interaction

Appendix 5.15 - Adjusted HRs for association between ever exposure to endocrine therapies and risk of revascularisation, stratified by age, time since index, and prior CVD



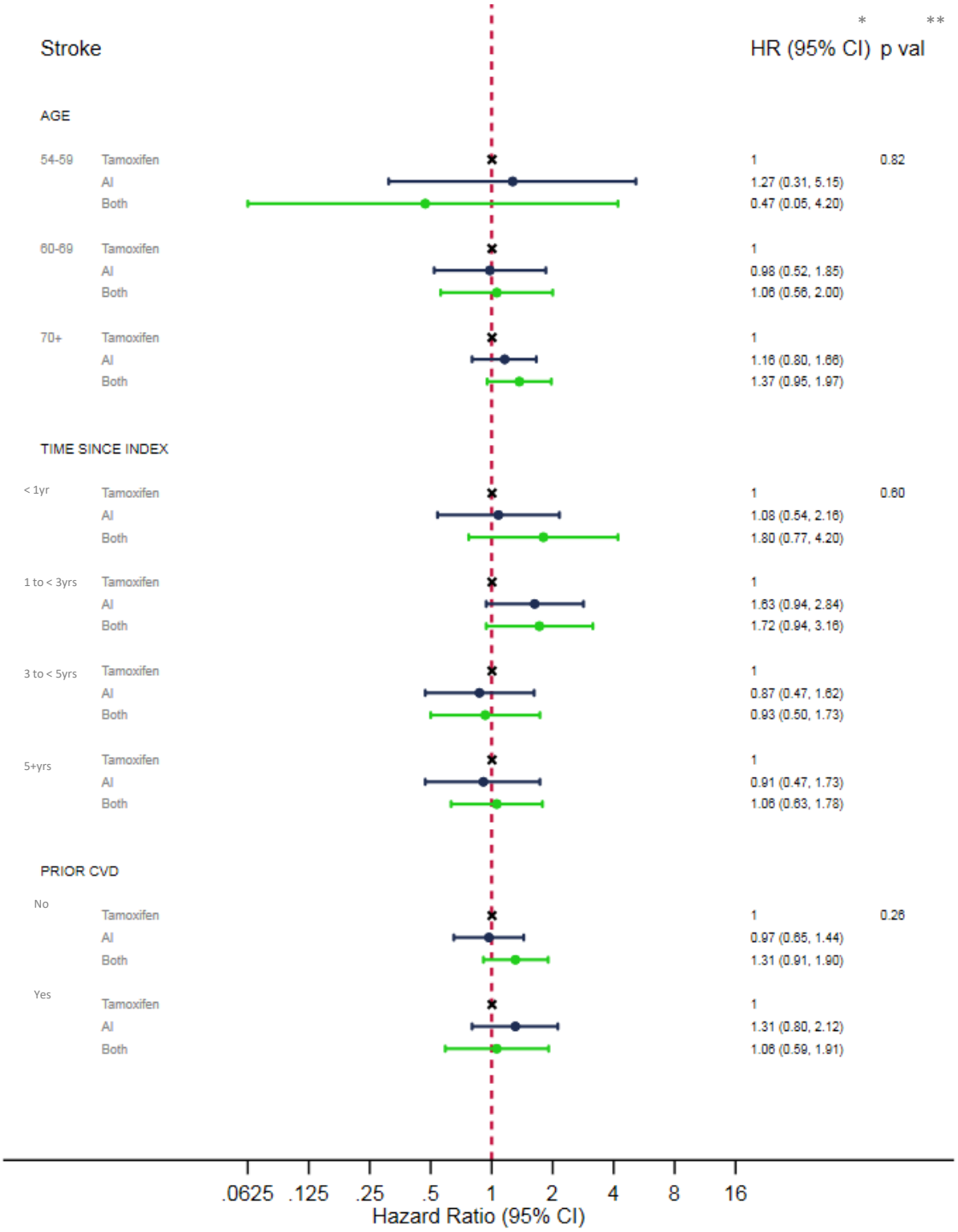
*Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+); smoking status (non-smoker, current smoker, ex-smoker); BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker); IMD score (level 1-5 based on GP level IMD data); use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes; chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of VTE; history of non-venous CVD year of breast cancer diagnosis; time since index (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year.
** p value from Wald test for interaction

Appendix 5.16 - Adjusted HRs for association between ever exposure to endocrine therapies and risk of PVD, stratified by age, time since index, and prior CVD



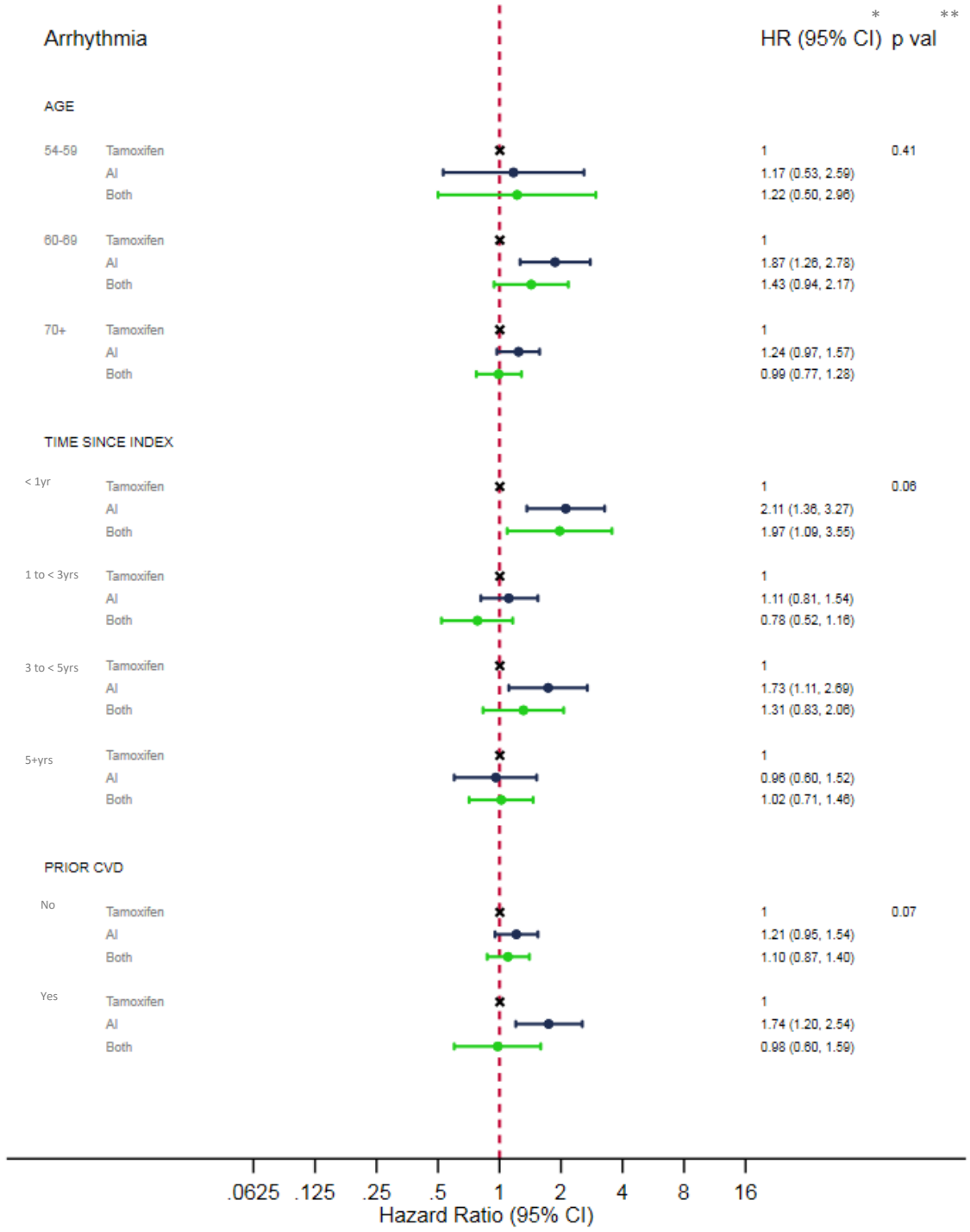
*Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+); smoking status (non-smoker, current smoker, ex-smoker); BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker); IMD score (level 1-5 based on GP level IMD data); use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes; chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of VTE; history of non-venous CVD year of breast cancer diagnosis; time since index <1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year.
** p value from Wald test for interaction

Appendix 5.17 - Adjusted HRs for association between ever exposure to endocrine therapies and risk of stroke, stratified by age, time since index, and prior CVD



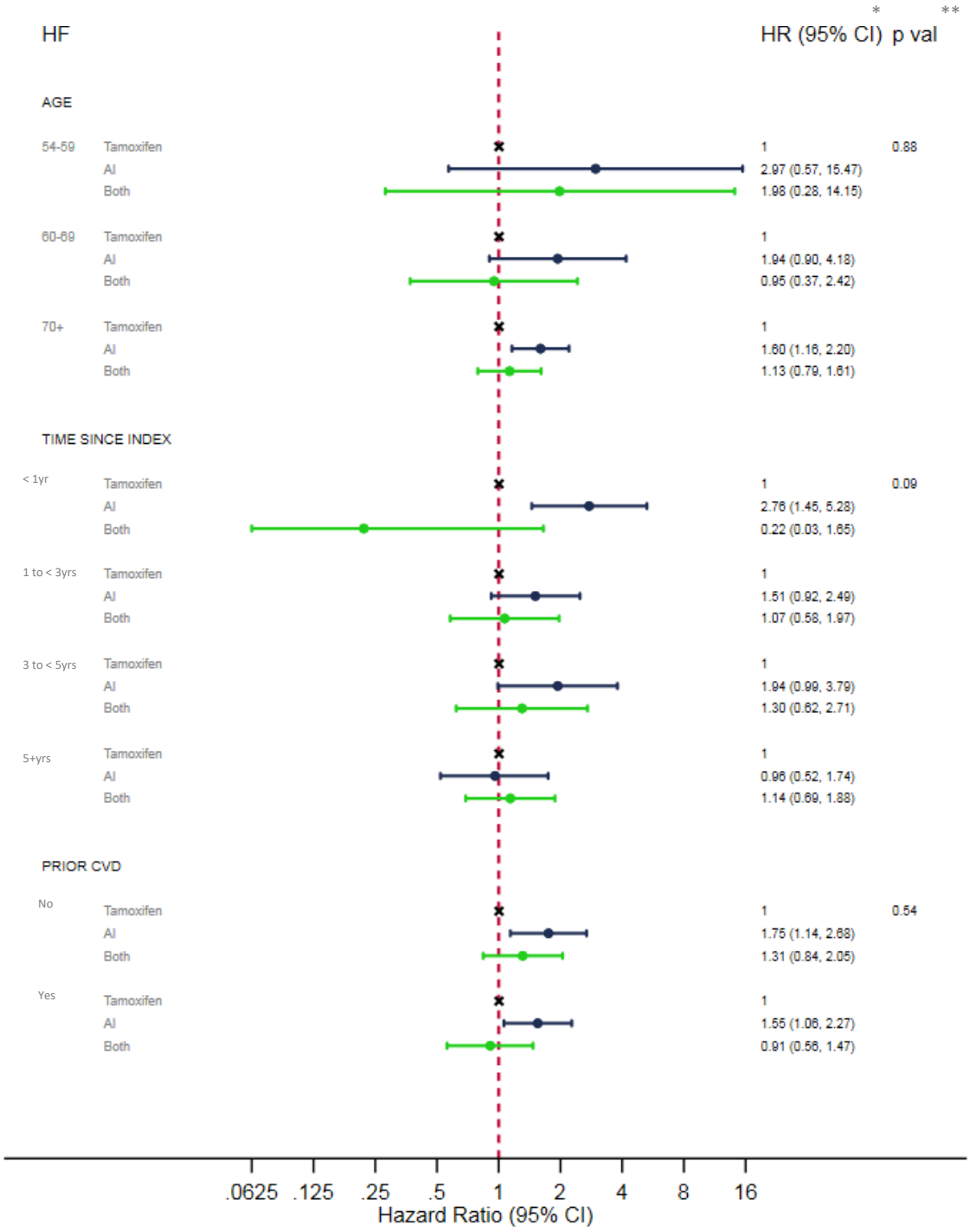
*Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+); smoking status (non-smoker, current smoker, ex-smoker); BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker); IMD score (level 1-5 based on GP level IMD data); use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes; chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of VTE; history of non-venous CVD year of breast cancer diagnosis; time since index (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year.
** p value from Wald test for interaction

Appendix 5.18 - Adjusted HRs for association between ever exposure to endocrine therapies and risk of arrhythmia, stratified by age, time since index, and prior CVD



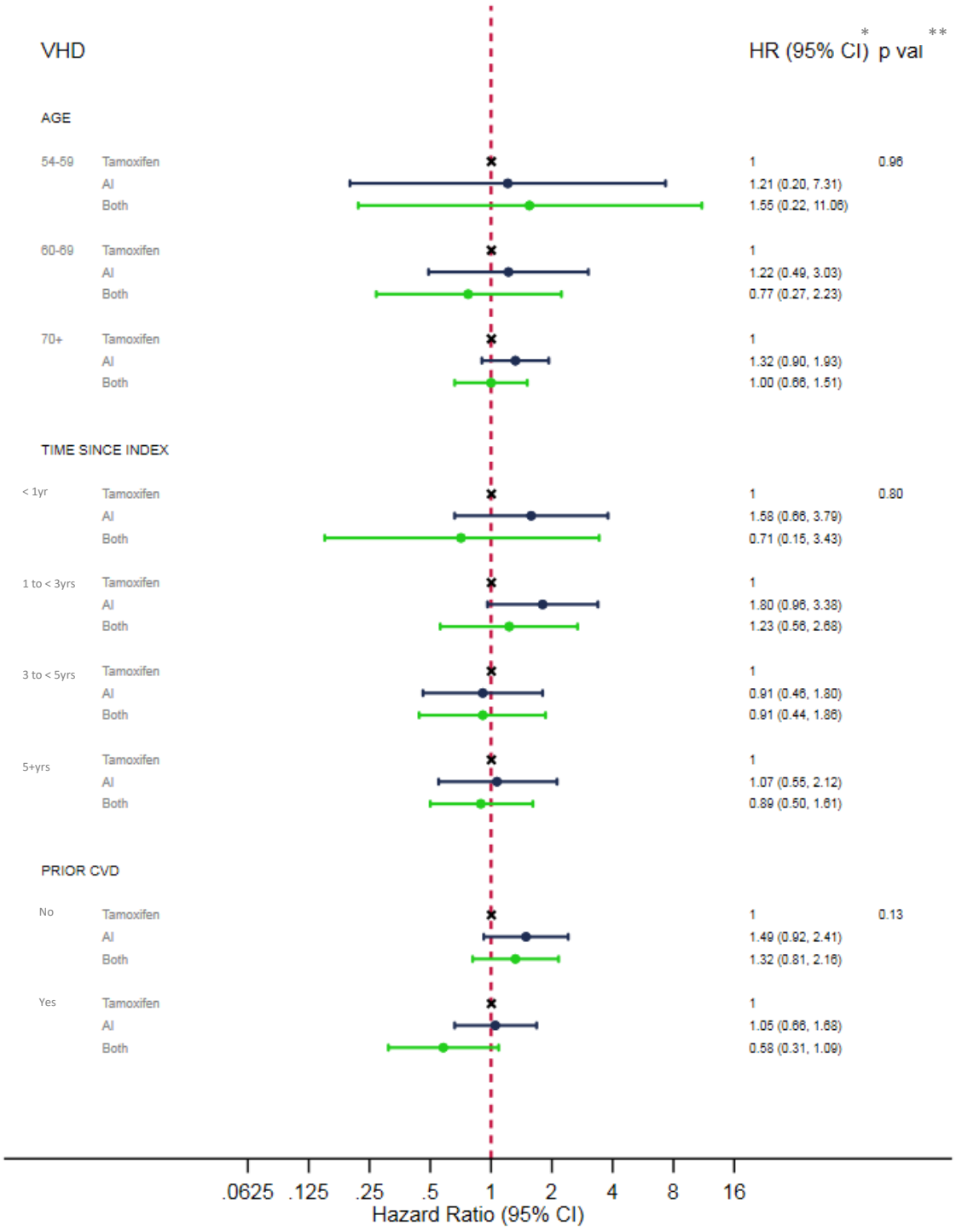
*Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+); smoking status (non-smoker, current smoker, ex-smoker); BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker); IMD score (level 1-5 based on GP level IMD data); use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes; chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of VTE; history of non-venous CVD year of breast cancer diagnosis; time since index (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year.
** p value from Wald test for interaction

Appendix 5.19 - Adjusted HRs for association between ever exposure to endocrine therapies and risk of HF, stratified by age, time since index, and prior CVD



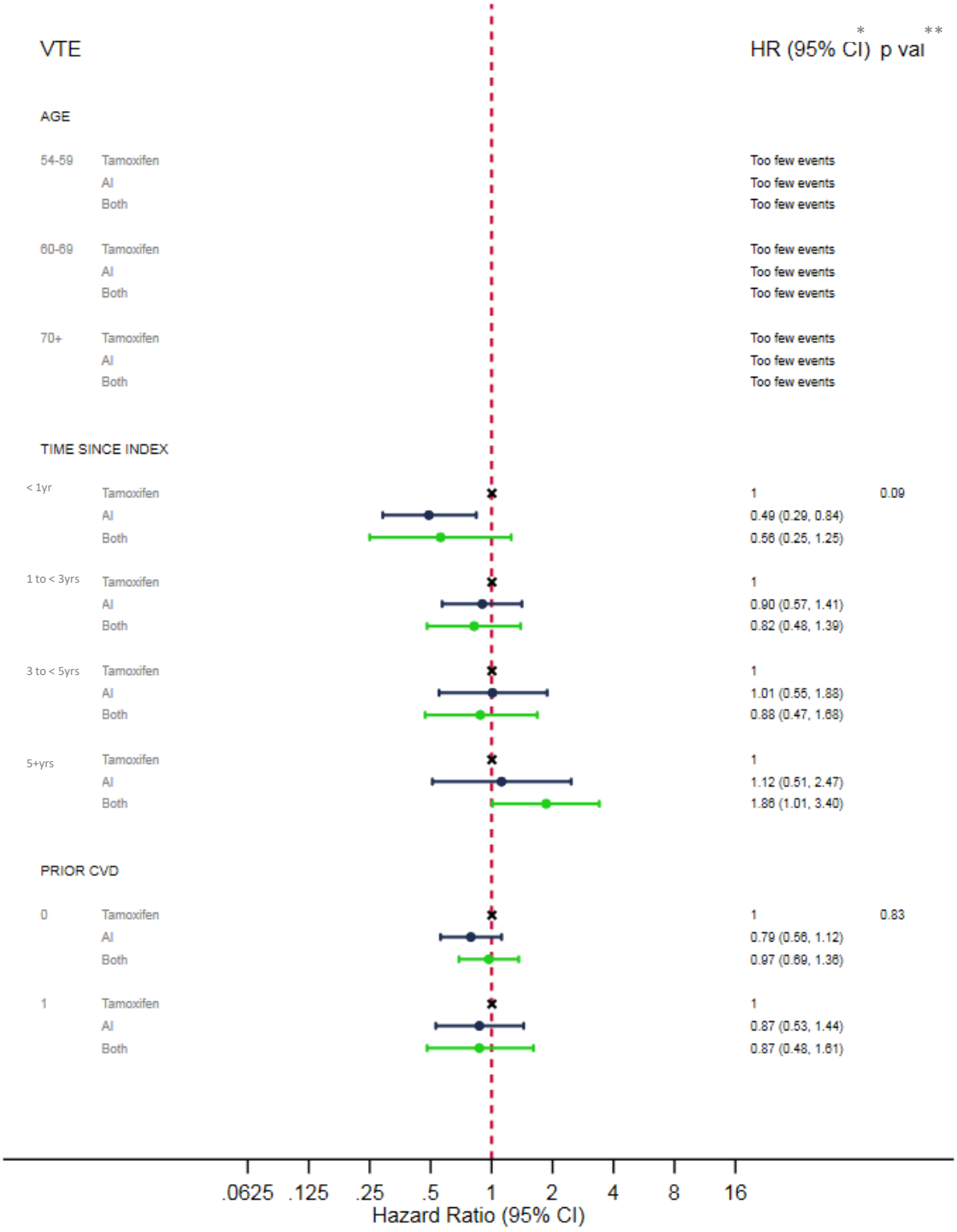
*Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+); smoking status (non-smoker, current smoker, ex-smoker); BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker); IMD score (level 1-5 based on GP level IMD data); use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes; chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of VTE; history of non-venous CVD year of breast cancer diagnosis; time since index (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year.
** p value from Wald test for interaction

Appendix 5.20 - Adjusted HRs for association between ever exposure to endocrine therapies and risk of VHD, stratified by age, time since index, and prior CVD



*Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+); smoking status (non-smoker, current smoker, ex-smoker); BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker); IMD score (level 1-5 based on GP level IMD data); use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes; chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of VTE; history of non-venous CVD year of breast cancer diagnosis; time since index (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year.
** p value from Wald test for interaction

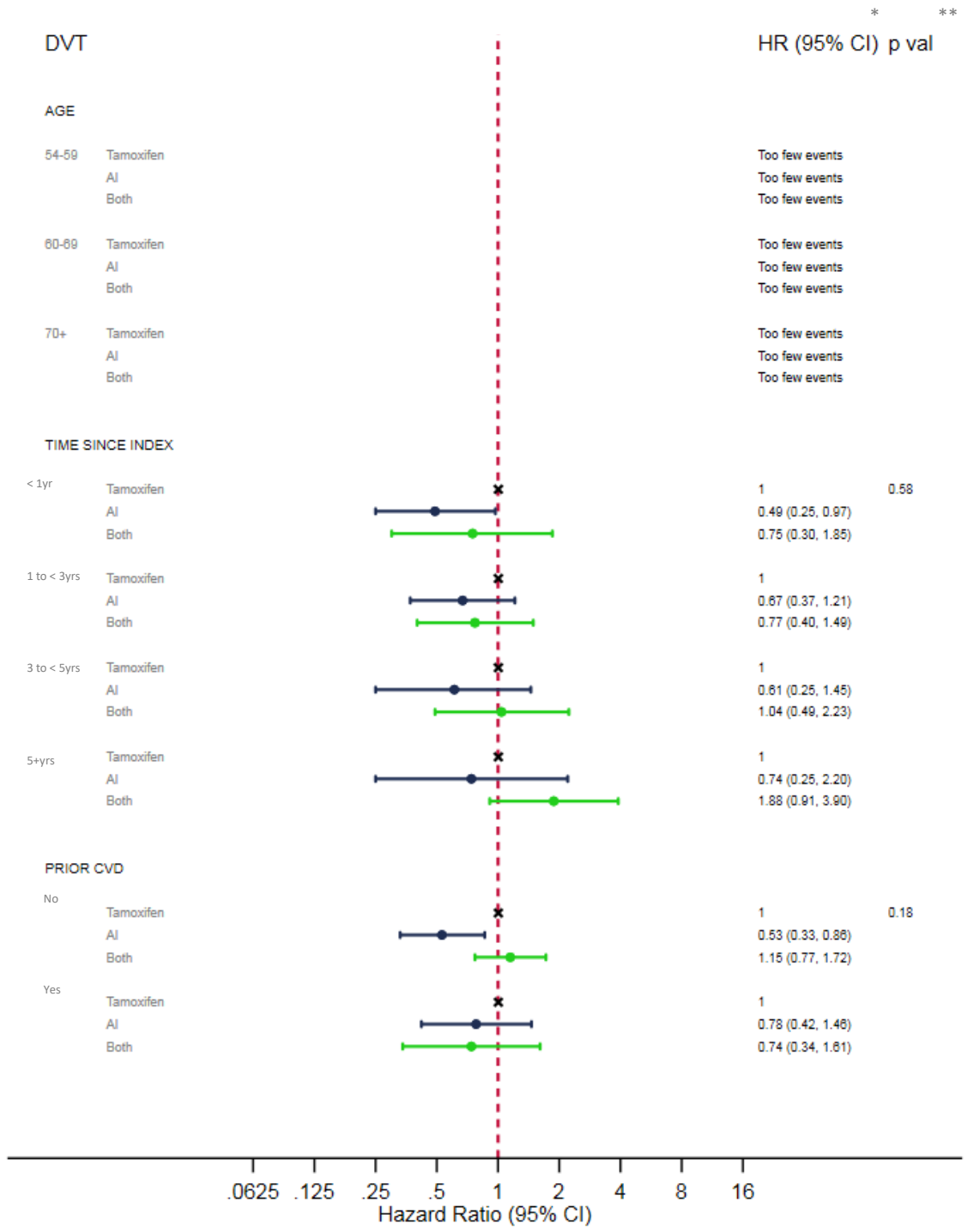
Appendix 5.21 - Adjusted HRs for association between ever exposure to endocrine therapies and risk of VTE, stratified by age, time since index, and prior CVD



*Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+); smoking status (non-smoker, current smoker, ex-smoker); BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker); IMD score (level 1-5 based on GP level IMD data); use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes; chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of VTE; history of non-venous CVD year of breast cancer diagnosis; time since index (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year.
** p value from Wald test for interaction

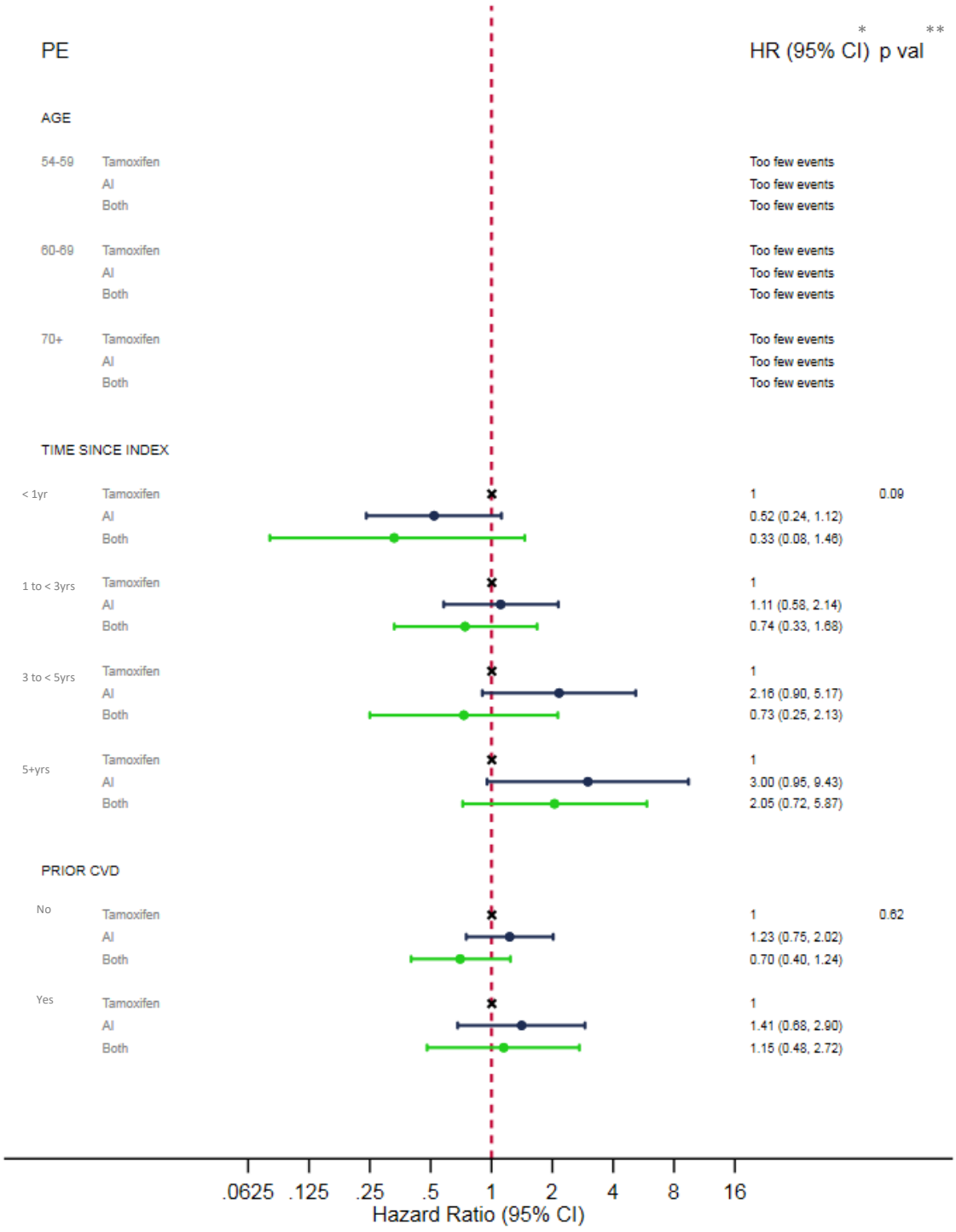
* **

Appendix 5.22 - Adjusted HRs for association between ever exposure to endocrine therapies and risk of DVT, stratified by age, time since index, and prior CVD



*Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+); smoking status (non-smoker, current smoker, ex-smoker); BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker); IMD score (level 1-5 based on GP level IMD data); use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes; chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of VTE; history of non-venous CVD year of breast cancer diagnosis; time since index (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year.
** p value from Wald test for interaction

Appendix 5.23 - Adjusted HRs for association between ever exposure to endocrine therapies and risk of PE, stratified by age, time since index, and prior CVD



*Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+); smoking status (non-smoker, current smoker, ex-smoker); BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker); IMD score (level 1-5 based on GP level IMD data); use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes; chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of VTE; history of non-venous CVD year of breast cancer diagnosis; time since index (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year.
** p value from Wald test for interaction

Appendix 5.24 - Characteristics of study population based on their initial exposure, before and after guideline changes

	Before			After		
	Tamoxifen	**AI**	**Total**	**Tamoxifen**	**AI**	**Total**
N	2267 (100)	670 (100)	2937 (100)	2449 (100)	4619 (100)	7068 (100)
Age (yrs)						
54-59	476 (21)	69 (10.3)	545 (18.6)	435 (17.8)	647 (14)	1082 (15.3)
60-69	815 (36)	208 (31)	1023 (34.8)	1035 (42.3)	1764 (38.2)	2799 (39.6)
70+	976 (43.1)	393 (58.7)	1369 (46.6)	979 (40)	2208 (47.8)	3187 (45.1)
Median (IQR)	68 (61-77)	74 (65-83)	69 (62-78)	68 (62-75)	69 (63-79)	69 (63-78)
Year of breast cancer						
2002	579 (25.5)	72 (10.7)	651 (22.2)	0 (0)	0 (0)	0 (0)
2003	605 (26.7)	123 (18.4)	728 (24.8)	0 (0)	0 (0)	0 (0)
2004	571 (25.2)	186 (27.8)	757 (25.8)	0 (0)	0 (0)	0 (0)
2005	512 (22.6)	289 (43.1)	801 (27.3)	0 (0)	0 (0)	0 (0)
2006	0 (0)	0 (0)	0 (0)	467 (19.1)	370 (8)	837 (11.8)
2007	0 (0)	0 (0)	0 (0)	386 (15.8)	448 (9.7)	834 (11.8)
2008	0 (0)	0 (0)	0 (0)	380 (15.5)	495 (10.7)	875 (12.4)
2009	0 (0)	0 (0)	0 (0)	290 (11.8)	533 (11.5)	823 (11.6)
2010	0 (0)	0 (0)	0 (0)	223 (9.1)	586 (12.7)	809 (11.4)
2011	0 (0)	0 (0)	0 (0)	216 (8.8)	619 (13.4)	835 (11.8)
2012	0 (0)	0 (0)	0 (0)	210 (8.6)	539 (11.7)	749 (10.6)
2013	0 (0)	0 (0)	0 (0)	174 (7.1)	504 (10.9)	678 (9.6)
2014	0 (0)	0 (0)	0 (0)	93 (3.8)	451 (9.8)	544 (7.7)
2015	0 (0)	0 (0)	0 (0)	10 (.4)	74 (1.6)	84 (1.2)
BMI (kg/m2)						
<18	26 (1.1)	11 (1.6)	37 (1.3)	33 (1.3)	52 (1.1)	85 (1.2)
18-24	808 (35.6)	211 (31.5)	1019 (34.7)	885 (36.1)	1408 (30.5)	2293 (32.4)
25-29	757 (33.4)	208 (31)	965 (32.9)	792 (32.3)	1593 (34.5)	2385 (33.7)
30-34	361 (15.9)	108 (16.1)	469 (16)	439 (17.9)	871 (18.9)	1310 (18.5)
≥35	146 (6.4)	55 (8.2)	201 (6.8)	199 (8.1)	493 (10.7)	692 (9.8)
Missing	169 (7.5)	77 (11.5)	246 (8.4)	101 (4.1)	202 (4.4)	303 (4.3)
Median (IQR)	26 (23-30)	27 (23-30)	26 (23-30)	26 (23-30)	27 (24-31)	27 (24-31)
Smoking status						
Never smoker	1235 (54.5)	332 (49.6)	1567 (53.4)	1188 (48.5)	2185 (47.3)	3373 (47.7)
Current smoker	271 (12)	64 (9.6)	335 (11.4)	228 (9.3)	419 (9.1)	647 (9.2)
Ex-smoker	733 (32.3)	261 (39)	994 (33.8)	1032 (42.1)	2006 (43.4)	3038 (43)
Missing	28 (1.2)	13 (1.9)	41 (1.4)	1 (0)	9 (.2)	10 (.1)
Alcohol use						
Non drinker	333 (14.7)	121 (18.1)	454 (15.5)	285 (11.6)	492 (10.7)	777 (11)
Current	1569 (69.2)	409 (61)	1978 (67.3)	1750 (71.5)	3229 (69.9)	4979 (70.4)
Ex-drinker	184 (8.1)	76 (11.3)	260 (8.9)	297 (12.1)	629 (13.6)	926 (13.1)
Missing	181 (8)	64 (9.6)	245 (8.3)	117 (4.8)	269 (5.8)	386 (5.5)
Systolic BP						
Low/ideal	215 (9.5)	62 (9.3)	277 (9.4)	315 (12.9)	538 (11.6)	853 (12.1)
Pre-high	801 (35.3)	225 (33.6)	1026 (34.9)	1061 (43.3)	2102 (45.5)	3163 (44.8)
High	1243 (54.8)	382 (57)	1625 (55.3)	1071 (43.7)	1972 (42.7)	3043 (43.1)
Missing	8 (.4)	1 (.1)	9 (.3)	2 (.1)	7 (.2)	9 (.1)
Diastolic BP						
Low/ideal	922 (40.7)	301 (44.9)	1223 (41.6)	1208 (49.3)	2348 (50.8)	3556 (50.3)
Pre-high	1004 (44.3)	276 (41.2)	1280 (43.6)	984 (40.2)	1782 (38.6)	2766 (39.1)
High	333 (14.7)	92 (13.7)	425 (14.5)	255 (10.4)	482 (10.4)	737 (10.4)
Missing	8 (.4)	1 (.1)	9 (.3)	2 (.1)	7 (.2)	9 (.1)
IMD category						
1	386 (17)	119 (17.8)	505 (17.2)	484 (19.8)	843 (18.3)	1327 (18.8)
2	453 (20)	141 (21)	594 (20.2)	490 (20)	1073 (23.2)	1563 (22.1)
3	431 (19)	124 (18.5)	555 (18.9)	494 (20.2)	930 (20.1)	1424 (20.1)
4	541 (23.9)	142 (21.2)	683 (23.3)	511 (20.9)	794 (17.2)	1305 (18.5)
5	456 (20.1)	144 (21.5)	600 (20.4)	470 (19.2)	978 (21.2)	1448 (20.5)
Missing	0 (0)	0 (0)	0 (0)	0 (0)	1 (0)	1 (0)
Statins before index	385 (17)	190 (28.4)	575 (19.6)	715 (29.2)	1713 (37.1)	2428 (34.4)
ACEI before index	475 (21)	204 (30.4)	679 (23.1)	720 (29.4)	1619 (35.1)	2339 (33.1)
CCB before index	510 (22.5)	204 (30.4)	714 (24.3)	685 (28)	1560 (33.8)	2245 (31.8)
ARB before index	180 (7.9)	77 (11.5)	257 (8.8)	313 (12.8)	718 (15.5)	1031 (14.6)
Anti-platelets before	516 (22.8)	245 (36.6)	761 (25.9)	616 (25.2)	1394 (30.2)	2010 (28.4)
RA before index	52 (2.3)	26 (3.9)	78 (2.7)	86 (3.5)	111 (2.4)	197 (2.8)
Diabetes before index	204 (9)	76 (11.3)	280 (9.5)	258 (10.5)	652 (14.1)	910 (12.9)
CKD before index	464 (20.5)	218 (32.5)	682 (23.2)	403 (16.5)	872 (18.9)	1275 (18)
Non-venous CVD before	455 (20.1)	235 (35.1)	690 (23.5)	576 (23.5)	1401 (30.3)	1977 (28)
VTE before index	64 (2.8)	75 (11.2)	139 (4.7)	80 (3.3)	249 (5.4)	329 (4.7)

Appendix 5.25 - Sensitivity analysis – Adjusted HRs for current exposure to endocrine therapy for Arrhythmia, DVT, and HF, with different grace periods, where the grace period is the maximum gap in prescription coverage for exposure to be considered continuous

Outcome	Current exposure	Adjusted HR – 1m grace (original)	Adjusted HR – 3m grace	Adjusted HR – 6m grace	Adjusted HR – 1y grace
Arrhythmia	Tamoxifen	1	1	1	1
	AI	1.46 (1.17, 1.82)	1.49 (1.20, 1.84)	1.45 (1.18, 1.79)	1.37 (1.12, 1.67)
	Past with AI	1.88 (1.39, 2.54)	1.74 (1.27, 2.39)	1.89 (1.36, 2.63)	1.72 (1.21, 2.44)
	Past Tam only	1.88 (1.37, 2.60)	1.90 (1.36, 2.66)	2.05 (1.46, 2.89)	1.91 (1.32, 2.74)
DVT	Tamoxifen	1	1	1	1
	AI	0.51 (0.35, 0.73)	0.53 (0.37, 0.75)	0.59 (0.42, 0.83)	0.62 (0.45, 0.87)
	Past with AI	0.83 (0.49, 1.41)	0.81 (0.46, 1.42)	0.42 (0.22, 0.81)	0.41 (0.20, 0.84)
	Past Tam only	0.50 (0.26, 0.98)	0.46 (0.23, 0.95)	0.33 (0.15, 0.70)	0.41 (0.19, 0.89)
HF	Tamoxifen	1	1	1	1
	AI	1.49 (1.08, 2.06)	1.49 (1.09, 2.03)	1.55 (1.14, 2.10)	1.54 (1.14, 2.09)
	Past with AI	1.72 (1.12, 2.65)	1.50 (0.96, 2.36)	1.41 (0.87, 2.27)	1.47 (0.89, 2.43)
	Past Tam only	1.31 (0.80, 2.15)	1.15 (0.69, 1.94)	1.18 (0.69, 2.01)	1.37 (0.79, 2.36)

Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+);* smoking status (non-smoker, current smoker, ex-smoker)*;* BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker)*;* IMD score (level 1-5 based on GP level IMD data)*;* use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes*;* chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high history of VTE; history of non-venous CVD ; year of breast cancer diagnosis; time since index (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); and current year

Appendix 5.26 - Crude rates, unadjusted HRs, and adjusted HRs by ever exposure to endocrine therapies for a range of clinical CVDs, additionally adjusted for HRT use prior to index

Outcome	Ever exposure	Number of Events	Person years (1000s)	Crude Rate (per 1000 Person-Years)	Unadjusted HR (95% CI)	Adjusted HR (95% CI) *
Coronary Artery Disease	Tamoxifen	93	13.54	6.87 (5.60, 8.41)	1	1
	AI	131	12.87	10.18 (8.58, 12.08)	1.40 (1.06, 1.86)	1.29 (0.94, 1.76)
	Both	67	10.6	6.32 (4.98, 8.03)	0.95 (0.69, 1.31)	0.95 (0.68, 1.32)
Angina	Tamoxifen	56	13.73	4.08 (3.14, 5.30)	1	1
	AI	80	13.15	6.09 (4.89, 7.58)	1.43 (1.00, 2.05)	1.31 (0.88, 1.97)
	Both	31	10.73	2.89 (2.03, 4.11)	0.70 (0.44, 1.10)	0.69 (0.43, 1.11)
MI	Tamoxifen	32	14.6	2.19 (1.55, 3.10)	1	1
	AI	61	14.11	4.32 (3.36, 5.56)	1.73 (1.12, 2.68)	1.56 (0.96, 2.52)
	Both	39	11.24	3.47 (2.54, 4.75)	1.62 (1.01, 2.60)	1.62 (0.99, 2.63)
Revascularisation	Tamoxifen	15	14.72	1.02 (0.61, 1.69)	1	1
	AI	20	14.3	1.40 (0.90, 2.17)	1.36 (0.69, 2.68)	1.84 (0.85, 4.01)
	Both	12	11.33	1.06 (0.60, 1.86)	0.99 (0.46, 2.11)	1.00 (0.46, 2.21)
SCA	Tamoxifen	13	14.88	0.87 (0.51, 1.50)	1	1
	AI	21	14.53	1.45 (0.94, 2.22)	1.87 (0.83, 4.22)	1.65 (0.65, 4.19)
	Both	5	11.43	0.44 (0.18, 1.05)	0.73 (0.24, 2.19)	0.69 (0.22, 2.10)
PVD	Tamoxifen	35	14.64	2.39 (1.72, 3.33)	1	1
	AI	41	14.08	2.91 (2.14, 3.96)	1.25 (0.77, 2.01)	1.31 (0.76, 2.25)
	Both	22	11.23	1.96 (1.29, 2.97)	0.81 (0.45, 1.43)	0.87 (0.48, 1.57)
Stroke	Tamoxifen	91	14.45	6.30 (5.13, 7.73)	1	1
	AI	118	13.71	8.61 (7.19, 10.31)	1.22 (0.91, 1.63)	1.11 (0.81, 1.53)
	Both	88	10.98	8.02 (6.51, 9.88)	1.33 (0.98, 1.80)	1.25 (0.91, 1.70)
Arrhythmia	Tamoxifen	219	12.7	17.25 (15.11, 19.69)	1	1
	AI	287	11.52	24.90 (22.18, 27.96)	1.37 (1.14, 1.64)	1.37 (1.12, 1.69)
	Both	174	9.51	18.30 (15.77, 21.23)	1.07 (0.87, 1.32)	1.08 (0.87, 1.34)
HF	Tamoxifen	90	14.33	6.28 (5.11, 7.72)	1	1
	AI	178	13.59	13.10 (11.31, 15.17)	1.87 (1.43, 2.45)	1.67 (1.24, 2.25)
	Both	76	11.04	6.89 (5.50, 8.62)	1.12 (0.81, 1.55)	1.12 (0.81, 1.56)
Pericarditis	Tamoxifen	3	14.85	0.20 (0.07, 0.63)	1	1
	AI	14	14.53	0.96 (0.57, 1.63)	3.96 (1.12, 14.03)	3.26 (0.86, 12.29)
	Both	10	11.4	0.88 (0.47, 1.63)	4.29 (1.18, 15.61)	3.46 (0.92, 13.11)
VHD	Tamoxifen	66	14.54	4.54 (3.57, 5.78)	1	1
	AI	114	13.93	8.18 (6.81, 9.83)	1.58 (1.15, 2.17)	1.31 (0.92, 1.86)
	Both	52	11.14	4.67 (3.56, 6.13)	1.08 (0.74, 1.57)	0.97 (0.66, 1.42)
VTE	Tamoxifen	122	14.41	8.47 (7.09, 10.11)	1	1
	AI	116	13.4	8.66 (7.22, 10.38)	0.93 (0.71, 1.21)	0.82 (0.61, 1.10)
	Both	85	10.63	8.00 (6.47, 9.89)	0.91 (0.69, 1.22)	0.95 (0.71, 1.28)
DVT	Tamoxifen	83	14.52	5.72 (4.61, 7.09)	1	1
	AI	62	13.78	4.50 (3.51, 5.77)	0.69 (0.48, 0.98)	0.62 (0.42, 0.92)
	Both	59	10.83	5.45 (4.22, 7.03)	0.98 (0.69, 1.38)	1.06 (0.74, 1.52)
PE	Tamoxifen	48	14.75	3.25 (2.45, 4.32)	1	1
	AI	68	14.06	4.84 (3.81, 6.13)	1.39 (0.95, 2.04)	1.21 (0.80, 1.85)
	Both	33	11.19	2.95 (2.10, 4.15)	0.80 (0.51, 1.28)	0.76 (0.47, 1.23)

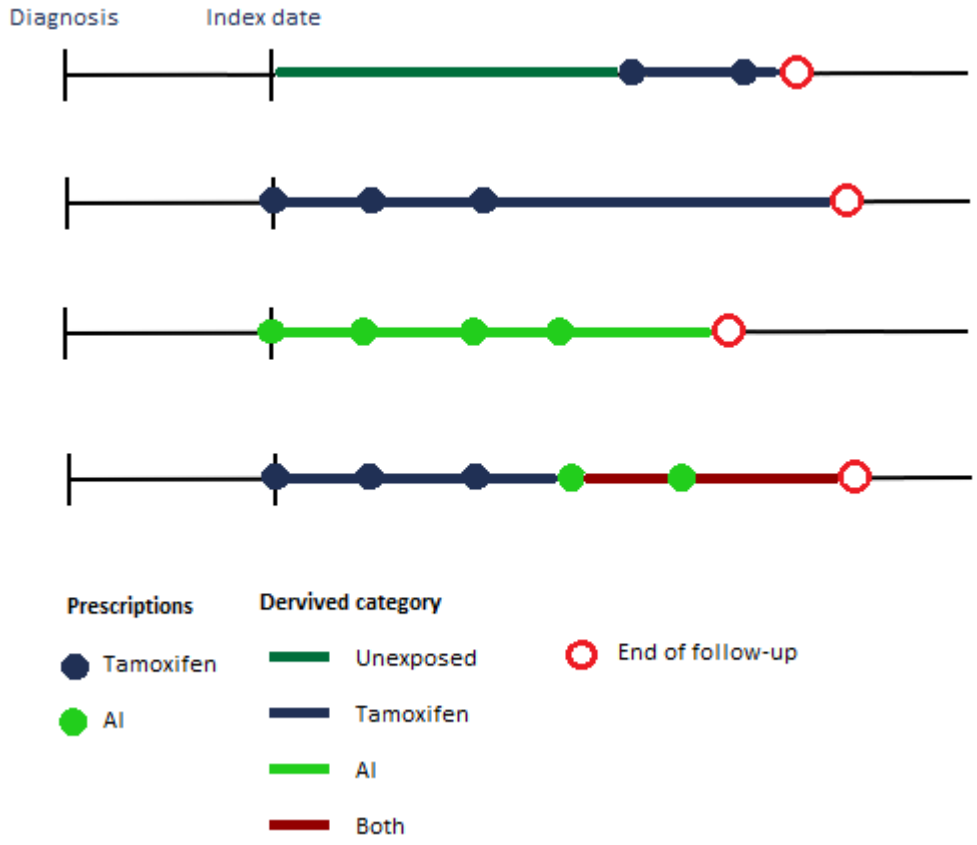
Adjusted the following covariates at baseline: for age (54-59, 60-69, 70+);* smoking status (non-smoker, current smoker, ex-smoker)*;* BMI (underweight/healthy weight, overweight, obese); alcohol status (non-drinker, current drinker, ex-drinker)*;* IMD score (level 1-5 based on GP level IMD data)*;* use of statins; use of ACE inhibitors; use of calcium channel blockers (CCB); use of angiotensin II receptor blockers (ARB); diabetes*;* chronic kidney disease; rheumatoid arthritis; systolic blood pressure (low/normal, pre-high, high); diastolic blood pressure (low/normal, pre-high, high); history of VTE; history of non-venous CVD year of breast cancer diagnosis; time since index (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current year; and use of HRT.

CHAPTER 6

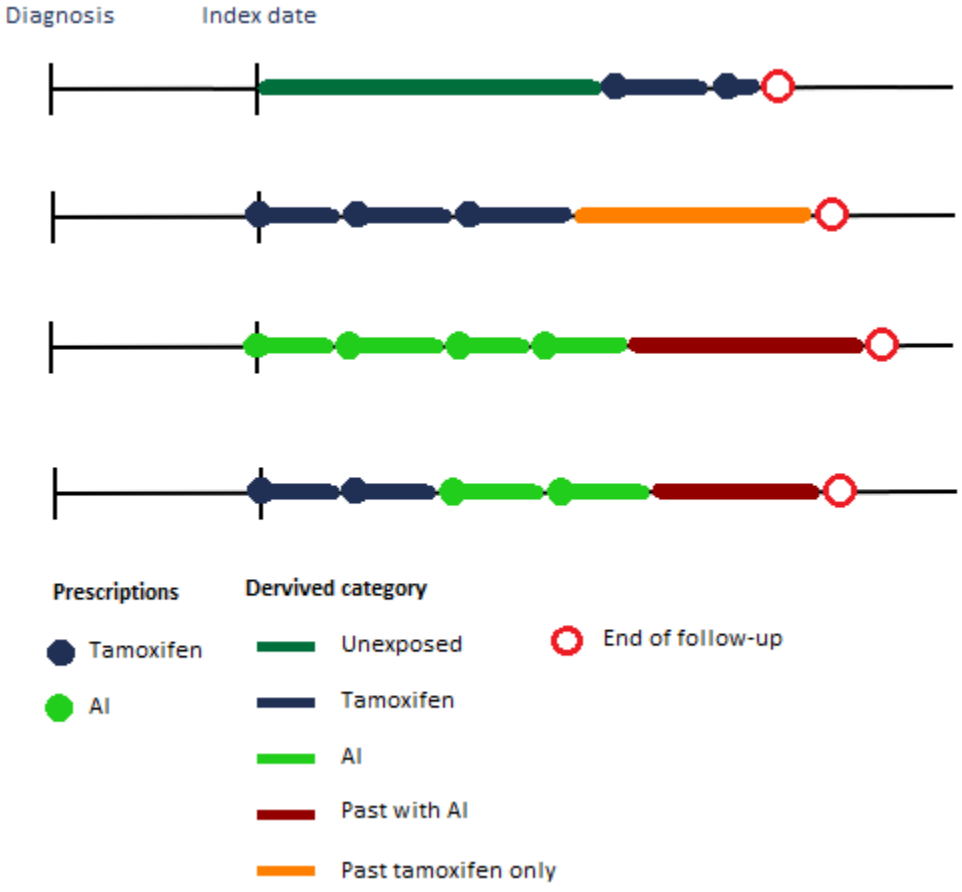
Appendix 6.1 - Endocrine therapy billing codes

	HCPCS	NDC
Tamoxifen	S0187	00038060025, 00038060060, 00054483121, 00054483126, 00054483413, 00054483422, 00054883125, 00054883425, 00093078210, 00093078256, 00093078405, 00093078406, 00093078410, 00172565658, 00172565670, 00172565680, 00172565746, 00172565755, 00172565760, 00172565770, 00310060018, 00310060075, 00310060412, 00310060430, 00310060490, 00310073060, 00378014405, 00378014491, 00378027401, 00378027493, 00440845092, 00440845130, 00440845160, 00440845192, 00555044603, 00555044605, 00555044609, 00555044663, 00555090401, 00591223218, 00591223319, 00591223330, 00591247319, 00591247330, 38779034101, 38779034103, 38779034108, 49452775301, 51552083802, 51927297600, 53002103203, 53002103230, 54569038201, 54569038202, 54569376501, 54569571600, 54569853100, 54569860200, 54868300401, 54868300402, 54868300403, 54868300404, 54868428700, 54868428703, 54868428704, 57866661501, 58016065760, 60346004832, 62991115101, 62991115103, 63304060060, 63304060130, 63304060190, 63739026910, 63739026915, 00093078201, 00093078205, 00093078486, 00172565649, 00172565780, 00310060025, 00310060060, 00310073130, 00440845030, 00440845060, 00555090405, 00555090414, 00591223260, 00591247218, 00591247260, 13632012301, 38779034104, 38779034105, 42254034390, 52372075601, 52372075602, 54569038200, 54569376500, 54868300405, 54868428701, 54868428702, 55175550006, 55289058530, 57866661801, 62991115104, 63304060028
Aromatase inhibitors	S0170 S0156	Anastrozole: 00054016413, 00093753656, 00310020130, 00310020137, 00310020150, 00378603405, 00378603477, 00781535631, 00904619546, 00904622961, 16571042103, 16729003510, 16729003515, 16729003516, 21695099030, 35356027030, 38779227406, 38779255503, 38779255504, 38779255506, 42043018003, 42291010530, 51079032301, 51079032306, 51927443500, 51991062010, 51991062033, 54569573100, 54569619800, 54868500000, 54868613000, 54868613001, 55111064730, 55175550503, 60258086603, 62756025013, 62756025083, 63275993001, 63275993002, 63323012930, 66435041530, 67877017110, 67877017130, 68084044811, 68084044821, 68258903501, 68382020906, 42254016130, 60429028630, 60429028690, 60505298503, 63275993003, 63275993004, 63275993005, 68382020910 Exemestane: 00054008013, 49999098630, 54569573200, 54868526100, 59762285801, 00009766304 Letrozole: 00054026913, 00078024915, 00093762056, 00378207105, 00378207193, 00603418016, 16729003410, 35356040930, 51991075910, 51991075933, 54569571400, 54868415100, 54868625200, 55111064630, 62756051183, 63323077230 42254024330, 60505325503, 60505325508

Appendix 6.2 - Visualisation of potential categorisations in the ever exposure



Appendix 6.3 - Visualisation of potential categorisations in the current exposure



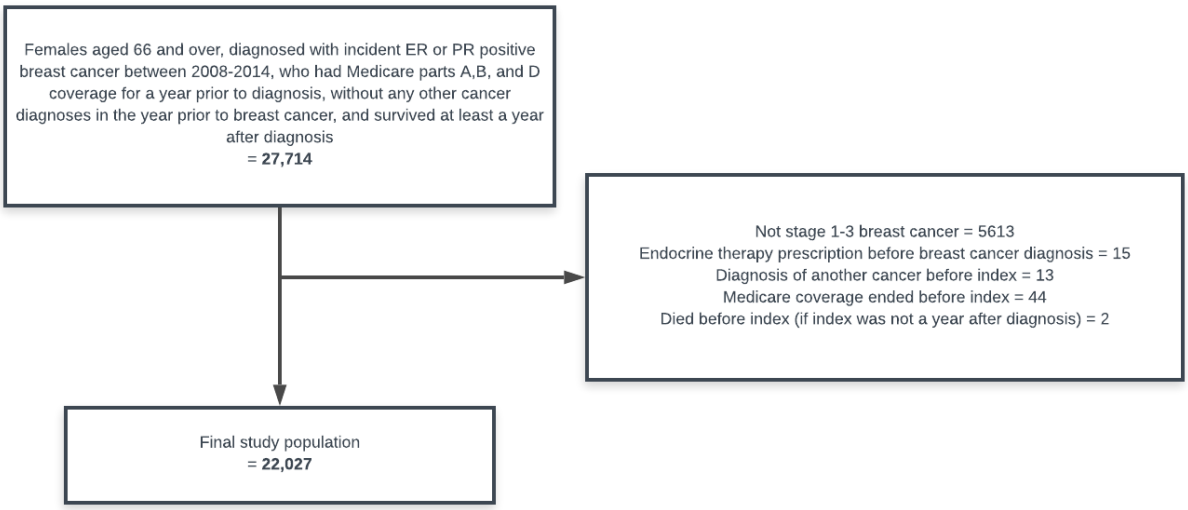
Appendix 6.4 - CVD outcome billing codes

Composite outcomes	Individual outcomes	ICD 9 diagnosis and procedure codes	CPT/HCPCS codes
Coronary Artery Disease	Angina	411.1, 413.1, 413.9	
	Myocardial infarction	410.11, 410.01, 410.31, 410.21, 410.41, 410.81, 410.51, 410.61, 410.91, 410.71	
	Revascularization procedures	36.0x, 36.1x, 36.2, 36.3	33140, 33510, 33511, 33512, 33513, 33514, 33516, 33517, 33518, 33519, 33521, 33522, 33523, 33533, 33534, 33535, 33536, 92920, 92924, 92928, 92933, 92937, 92941, 92943, 92980, 92981, 92982, 92984, 92995, 92996, 92997, 92998
	Sudden cardiac arrest	427.5	
	Peripheral vascular disease	443.89, 443.9	
	Stroke (haemorrhagic and ischaemic)	430.x, 431.x, 432.1, 432.0, 432.9, 433.91, 433.21, 433.01, 433.11, 433.31, 433.81, 434.01, 434.x, 435.0, 435.1, 435.3, 435.8, 435.9, 436.x	
	Arrhythmia	427.31, 427.32, 427.41, 427.42, 427.61, 427.0, 427.69, 427.60, 427.81, 427.89, 427.9	
	Heart failure	HF - 428.1, 428.0, 428.20, 428.21, 428.22, 428.23, 428.30, 428.31, 428.32, 428.33, 428.40, 428.41, 428.42, 428.43, 428.9 Cardiomyopathy - 414.8, 425.4, 425.11, 425.18, 425.0, 425.3, 425.5, 425.9, 425.2, 425.8	
	Pericarditis	420.91, 420.90, 420.99, 423.1, 423.2, 423.0, 423.9, 423.3, 423.8 420.0	
	Valvular heart disease	394.0, 394.1, 394.2, 394.9, 395.0, 395.1, 395.2, 395.9, 397.0, 396.0, 396.1, 396.2, 396.3, 397.9, 396.8, 396.9, 424.0, 424.1, 424.2, 424.3	
VTE	Deep venous thromboembolism	415.1x	
	Pulmonary Embolism	415.0, 415.12, 415.13, 415.19	

Appendix 6.5 - Covariate drugs billing codes

Chemotherapy billing codes	
Drug Class	**HCPCS/CPT *(variable)***
Taxanes	C9127, C9431, J9170, J9171, J9264, J9265, Q0125
Anthracyclines	C9415, J9000, J9001, J9178, J9180
Trastuzumab	J9355
Others	96408, 96409, 96410, 96411, 96412, 96413, 96414, 96415, 96416, 96417, 96545, 96549, C1167, C8953, C8954, C8955, C9214, C9240, C9257, C9280, C9414, C9418, C9420, C9421, C9425, C9438, C9440, G0359, G0360, G9021, G9022, G9023, G9024, G9025, G9026, G9027, G9028, G9029, G9030, G9031, G9032, J8520, J8521, J8530, J8560, J8999, J9035, J9045, J9060, J9062, J9070, J9080, J9090, J9091, J9092, J9093, J9094, J9095, J9096, J9097, J9179, J9181, J9182, J9190, J9201, J9207, J9250, J9260, J9293, J9360, J9390, J9999, Q0083, Q0084, Q0085
Concomitant medication billing codes	
Drug class	**ATC codes**
Hypertensives	C02
Statins	C10AA (C10AA01-C10AA08)
Ace Inhibitors	C09AA (C09AA01-C09AA16)
Calcium Channel Blockers	C08CA (C08CA01-C08CA16, C08CA55)
Angiotensin 2 receptor blocker	C09CA (C09CA01-C09CA09)

Appendix 6.6 - STROBE flow diagram for inclusion in US study



Appendix 6.7 - Adjusted HRs, events, and crude rate for the association between ever exposure to endocrine therapy and a range of clinical CVD outcomes

Outcome	Ever exposure	Events, Follow-up, Rate (95% CI) (per 1000 pyears) *	Unadjusted HR (95% CI)	Adjusted HR (95% CI)	Adjusted HR (95% CI)**
Coronary Artery Disease	Unexposed	318, 8.22, 38.70 (34.67, 43.19)	1.00 (., .)	1.00 (., .)	1.34 (1.09, 1.66)
	Tamoxifen	129, 4.79, 26.95 (22.68, 32.03)	0.74 (0.60, 0.91)	0.74 (0.60, 0.92)	1.00 (., .)
	AI	1059, 30.03, 35.26 (33.20, 37.45)	1.01 (0.89, 1.15)	0.96 (0.83, 1.10)	1.29 (1.06, 1.55)
	Both	118, 3.87, 30.51 (25.47, 36.54)	0.86 (0.69, 1.07)	0.89 (0.71, 1.10)	1.19 (0.92, 1.54)
Angina	Unexposed	189, 8.40, 22.51 (19.52, 25.96)	1.00 (., .)	1.00 (., .)	1.13 (0.87, 1.46)
	Tamoxifen	87, 4.83, 18.00 (14.59, 22.21)	0.85 (0.66, 1.10)	0.88 (0.68, 1.14)	1.00 (., .)
	AI	710, 30.61, 23.20 (21.55, 24.97)	1.06 (0.90, 1.26)	1.05 (0.89, 1.25)	1.19 (0.95, 1.50)
	Both	76, 3.97, 19.13 (15.28, 23.95)	0.89 (0.68, 1.17)	0.95 (0.72, 1.25)	1.07 (0.78, 1.46)
MI	Unexposed	153, 9.39, 16.30 (13.91, 19.10)	1.00 (., .)	1.00 (., .)	2.29 (1.58, 3.33)
	Tamoxifen	38, 5.41, 7.02 (5.11, 9.65)	0.45 (0.31, 0.65)	0.44 (0.30, 0.63)	1.00 (., .)
	AI	407, 34.29, 11.87 (10.77, 13.08)	0.90 (0.74, 1.10)	0.79 (0.64, 0.97)	1.81 (1.28, 2.58)
	Both	42, 4.35, 9.66 (7.14, 13.07)	0.71 (0.50, 1.01)	0.67 (0.47, 0.96)	1.53 (0.97, 2.42)
Revascularisation	Unexposed	65, 9.40, 6.91 (5.42, 8.81)	1.00 (., .)	1.00 (., .)	1.60 (0.99, 2.57)
	Tamoxifen	25, 5.37, 4.66 (3.15, 6.89)	0.65 (0.41, 1.05)	0.63 (0.39, 1.01)	1.00 (., .)
	AI	234, 34.29, 6.82 (6.00, 7.76)	0.92 (0.69, 1.23)	0.91 (0.68, 1.23)	1.46 (0.95, 2.24)
	Both	30, 4.33, 6.94 (4.85, 9.92)	0.93 (0.60, 1.46)	0.99 (0.63, 1.55)	1.57 (0.91, 2.72)
SCA	Unexposed	80, 9.70, 8.25 (6.63, 10.27)	1.00 (., .)	1.00 (., .)	1.49 (0.96, 2.33)
	Tamoxifen	29, 5.50, 5.28 (3.67, 7.59)	0.69 (0.44, 1.07)	0.67 (0.43, 1.04)	1.00 (., .)
	AI	222, 35.21, 6.31 (5.53, 7.19)	0.91 (0.69, 1.20)	0.78 (0.59, 1.04)	1.17 (0.78, 1.76)
	Both	28, 4.48, 6.25 (4.32, 9.05)	0.82 (0.52, 1.31)	0.80 (0.50, 1.28)	1.20 (0.69, 2.08)
PVD	Unexposed	331, 7.81, 42.40 (38.07, 47.23)	1.00 (., .)	1.00 (., .)	1.10 (0.91, 1.34)
	Tamoxifen	158, 4.56, 34.64 (29.64, 40.48)	0.89 (0.74, 1.08)	0.91 (0.75, 1.10)	1.00 (., .)
	AI	1075, 28.85, 37.26 (35.10, 39.55)	1.05 (0.92, 1.19)	1.00 (0.87, 1.14)	1.10 (0.92, 1.31)
	Both	129, 3.81, 33.85 (28.48, 40.22)	0.91 (0.73, 1.12)	0.93 (0.75, 1.15)	1.02 (0.80, 1.30)
Stroke	Unexposed	404, 7.66, 52.76 (47.86, 58.16)	1.00 (., .)	1.00 (., .)	1.22 (1.02, 1.45)
	Tamoxifen	190, 4.65, 40.85 (35.43, 47.09)	0.85 (0.71, 1.02)	0.82 (0.69, 0.98)	1.00 (., .)
	AI	1126, 28.82, 39.07 (36.85, 41.42)	0.93 (0.82, 1.05)	0.87 (0.76, 0.98)	1.05 (0.90, 1.24)
	Both	134, 3.72, 36.06 (30.44, 42.71)	0.84 (0.69, 1.03)	0.81 (0.66, 1.00)	0.99 (0.79, 1.24)
Arrhythmia	Unexposed	510, 5.97, 85.40 (78.30, 93.14)	1.00 (., .)	1.00 (., .)	1.34 (1.14, 1.58)
	Tamoxifen	222, 3.67, 60.49 (53.04, 69.00)	0.73 (0.62, 0.86)	0.75 (0.63, 0.88)	1.00 (., .)
	AI	1640, 22.80, 71.92 (68.52, 75.48)	0.94 (0.85, 1.05)	0.91 (0.81, 1.01)	1.22 (1.05, 1.41)
	Both	189, 2.87, 65.90 (57.14, 76.00)	0.85 (0.71, 1.01)	0.88 (0.74, 1.04)	1.17 (0.96, 1.43)
HF	Unexposed	488, 6.77, 72.07 (65.95, 78.76)	1.00 (., .)	1.00 (., .)	1.15 (0.98, 1.36)
	Tamoxifen	233, 4.17, 55.90 (49.16, 63.56)	0.85 (0.72, 1.00)	0.87 (0.74, 1.02)	1.00 (., .)
	AI	1368, 26.03, 52.56 (49.85, 55.42)	0.93 (0.83, 1.04)	0.84 (0.75, 0.94)	0.96 (0.83, 1.12)
	Both	167, 3.44, 48.56 (41.72, 56.51)	0.85 (0.71, 1.02)	0.83 (0.69, 1.00)	0.96 (0.78, 1.18)
Pericarditis	Unexposed	74, 9.48, 7.80 (6.21, 9.80)	1.00 (., .)	1.00 (., .)	2.69 (1.54, 4.71)
	Tamoxifen	16, 5.44, 2.94 (1.80, 4.80)	0.37 (0.21, 0.64)	0.37 (0.21, 0.65)	1.00 (., .)
	AI	197, 34.48, 5.71 (4.97, 6.57)	0.74 (0.56, 0.98)	0.67 (0.50, 0.90)	1.81 (1.06, 3.08)
	Both	20, 4.43, 4.52 (2.91, 7.00)	0.55 (0.33, 0.93)	0.57 (0.34, 0.96)	1.53 (0.77, 3.05)
VHD	Unexposed	447, 6.08, 73.51 (67.00, 80.65)	1.00 (., .)	1.00 (., .)	1.24 (1.05, 1.47)
	Tamoxifen	205, 3.69, 55.62 (48.50, 63.78)	0.78 (0.66, 0.93)	0.81 (0.68, 0.96)	1.00 (., .)
	AI	1513, 22.37, 67.63 (64.31, 71.13)	1.01 (0.90, 1.13)	0.98 (0.87, 1.09)	1.21 (1.04, 1.41)
	Both	175, 2.86, 61.13 (52.72, 70.90)	0.90 (0.75, 1.07)	0.94 (0.78, 1.13)	1.16 (0.94, 1.43)
VTE	Unexposed	78, 9.34, 8.35 (6.69, 10.43)	1.00 (., .)	1.00 (., .)	0.72 (0.51, 1.02)
	Tamoxifen	58, 5.41, 10.73 (8.30, 13.88)	1.39 (0.98, 1.97)	1.39 (0.98, 1.98)	1.00 (., .)
	AI	284, 33.99, 8.36 (7.44, 9.39)	1.17 (0.90, 1.52)	1.11 (0.84, 1.46)	0.80 (0.59, 1.07)
	Both	54, 4.31, 12.54 (9.60, 16.37)	1.67 (1.16, 2.41)	1.71 (1.18, 2.47)	1.23 (0.83, 1.81)
DVT	Unexposed	72, 9.36, 7.69 (6.11, 9.69)	1.00 (., .)	1.00 (., .)	0.71 (0.49, 1.02)
	Tamoxifen	54, 5.42, 9.97 (7.63, 13.01)	1.41 (0.98, 2.03)	1.42 (0.98, 2.04)	1.00 (., .)
	AI	263, 34.05, 7.72 (6.85, 8.72)	1.18 (0.89, 1.56)	1.14 (0.86, 1.52)	0.81 (0.59, 1.09)
	Both	53, 4.31, 12.29 (9.39, 16.09)	1.79 (1.23, 2.60)	1.85 (1.26, 2.71)	1.31 (0.88, 1.94)
PE**	Unexposed	-, -, 0.93 (0.48, 1.78)	1.00 (., .)	1.00 (., .)	0.95 (0.32, 2.87)
	Tamoxifen	-, -, 0.91 (0.38, 2.18)	1.02 (0.34, 3.06)	1.05 (0.35, 3.16)	1.00 (., .)
	AI	29, 35.26, 0.82 (0.57, 1.18)	1.01 (0.47, 2.18)	0.80 (0.36, 1.79)	0.77 (0.29, 2.04)
	Both	-, -, 0.22 (0.03, 1.58)	0.27 (0.03, 2.19)	0.26 (0.03, 2.08)	0.24 (0.03, 2.11)

*Events and follow-up suppressed if number of events ≤ 11

**With reference category changed to ever tamoxifen

† HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

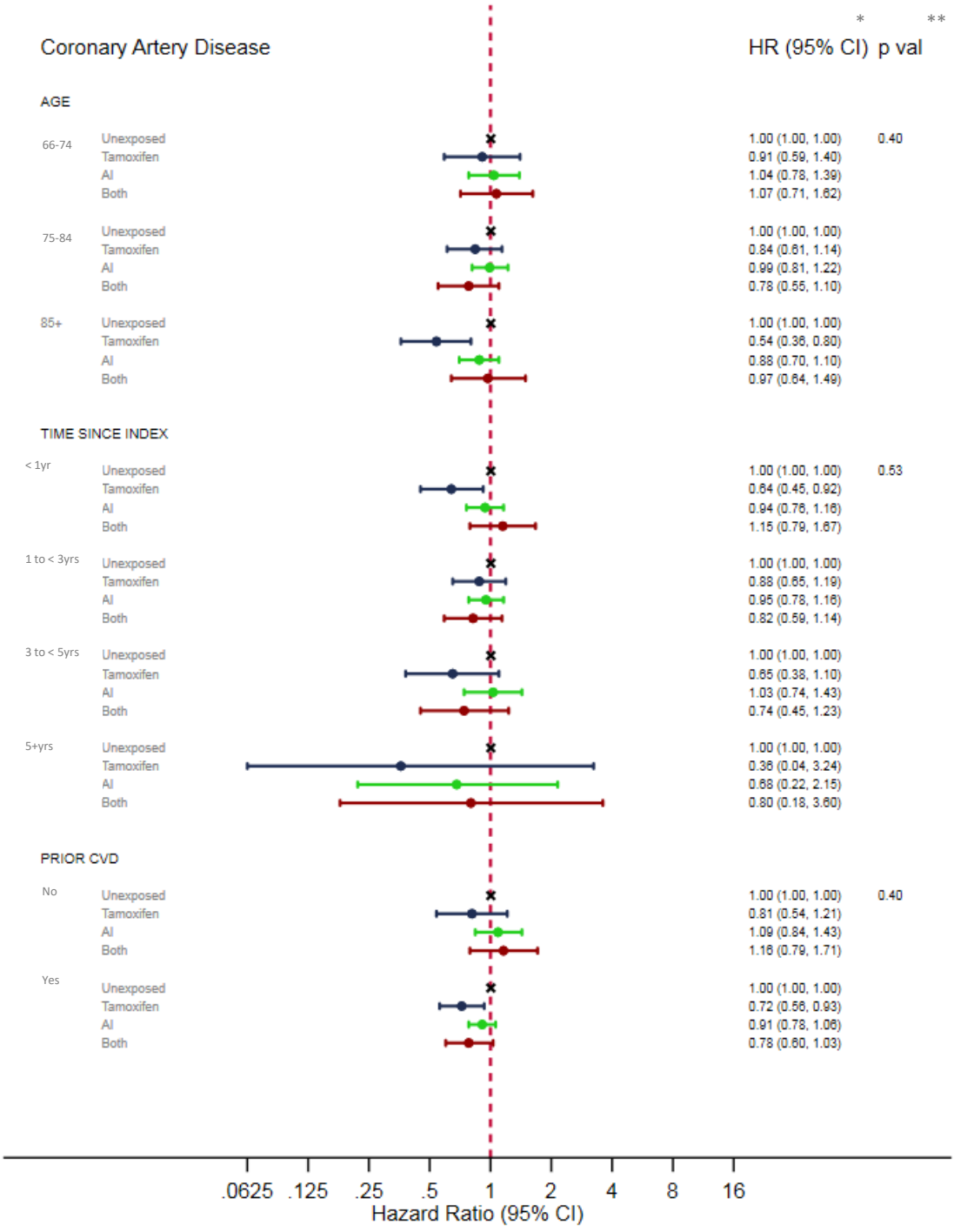
Appendix 6.8 - Adjusted HRs, events, and crude rate for the association between current exposure to endocrine therapy and a range of clinical CVD outcomes

Outcome	Current exposure	Events, Follow-up, Rate (95% CI)* (per 1000 pyears)	Unadjusted HR (95% CI)	Adjusted HR (95% CI)
Coronary Artery Disease	Unexposed	318, 8.22, 38.70 (34.67,	1.00 (., .)	1.00 (., .)
	Tamoxifen	141, 5.25, 26.85 (22.76,	0.77 (0.63, 0.94)	0.77 (0.63, 0.94)
	AI	807, 24.32, 33.19 (30.97,	0.96 (0.84, 1.10)	0.89 (0.77, 1.02)
	Past with AI	316, 7.76, 40.71 (36.46,	1.11 (0.95, 1.31)	1.11 (0.94, 1.32)
	Past Tam only	42, 1.35, 31.00 (22.91, 41.95)	0.78 (0.56, 1.08)	0.85 (0.61, 1.19)
Angina	Unexposed	189, 8.40, 22.51 (19.52,	1.00 (., .)	1.00 (., .)
	Tamoxifen	99, 5.33, 18.58 (15.26, 22.63)	0.88 (0.69, 1.13)	0.90 (0.70, 1.15)
	AI	561, 24.73, 22.69 (20.89,	1.05 (0.88, 1.25)	1.03 (0.86, 1.23)
	Past with AI	188, 7.99, 23.54 (20.40,	1.06 (0.86, 1.30)	1.10 (0.89, 1.36)
	Past Tam only	25, 1.37, 18.22 (12.31, 26.96)	0.83 (0.55, 1.27)	0.93 (0.61, 1.41)
MI	Unexposed	153, 9.39, 16.30 (13.91,	1.00 (., .)	1.00 (., .)
	Tamoxifen	40, 5.87, 6.82 (5.00, 9.29)	0.49 (0.34, 0.70)	0.47 (0.33, 0.67)
	AI	299, 27.59, 10.84 (9.68,	0.83 (0.67, 1.02)	0.71 (0.57, 0.88)
	Past with AI	135, 9.03, 14.96 (12.63,	1.06 (0.84, 1.35)	0.98 (0.77, 1.26)
	Past Tam only	13, 1.57, 8.29 (4.82, 14.28)	0.45 (0.25, 0.84)	0.46 (0.25, 0.85)
Revascularisation	Unexposed	65, 9.40, 6.91 (5.42, 8.81)	1.00 (., .)	1.00 (., .)
	Tamoxifen	37, 5.82, 6.35 (4.60, 8.77)	0.92 (0.61, 1.38)	0.89 (0.59, 1.35)
	AI	185, 27.54, 6.72 (5.82, 7.76)	0.89 (0.66, 1.20)	0.89 (0.65, 1.21)
	Past with AI	60, 9.06, 6.62 (5.14, 8.53)	0.90 (0.63, 1.30)	0.89 (0.62, 1.30)
	Past Tam only	7, 1.56, 4.49 (2.14, 9.41)	0.56 (0.24, 1.29)	0.54 (0.23, 1.27)
SCA	Unexposed	80, 9.70, 8.25 (6.63, 10.27)	1.00 (., .)	1.00 (., .)
	Tamoxifen	25, 5.96, 4.20 (2.83, 6.21)	0.59 (0.37, 0.94)	0.58 (0.36, 0.92)
	AI	150, 28.25, 5.31 (4.53, 6.23)	0.77 (0.57, 1.03)	0.64 (0.48, 0.87)
	Past with AI	87, 9.38, 9.28 (7.52, 11.44)	1.21 (0.88, 1.67)	1.11 (0.80, 1.54)
	Past Tam only	17, 1.60, 10.62 (6.60, 17.09)	1.30 (0.76, 2.22)	1.31 (0.76, 2.27)
PVD	Unexposed	331, 7.81, 42.40 (38.07,	1.00 (., .)	1.00 (., .)
	Tamoxifen	156, 5.01, 31.16 (26.64,	0.84 (0.69, 1.03)	0.85 (0.70, 1.03)
	AI	830, 23.42, 35.44 (33.11,	1.01 (0.88, 1.15)	0.95 (0.82, 1.09)
	Past with AI	312, 7.50, 41.62 (37.25,	1.09 (0.93, 1.28)	1.10 (0.93, 1.29)
	Past Tam only	64, 1.30, 49.07 (38.41, 62.69)	1.14 (0.86, 1.50)	1.21 (0.91, 1.60)
Stroke	Unexposed	404, 7.66, 52.76 (47.86,	1.00 (., .)	1.00 (., .)
	Tamoxifen	197, 5.11, 38.52 (33.50,	0.87 (0.73, 1.03)	0.83 (0.69, 0.99)
	AI	856, 23.48, 36.46 (34.10,	0.88 (0.78, 1.00)	0.81 (0.71, 0.93)
	Past with AI	346, 7.29, 47.43 (42.69,	1.06 (0.91, 1.23)	1.02 (0.87, 1.18)
	Past Tam only	51, 1.30, 39.13 (29.74, 51.48)	0.73 (0.53, 0.99)	0.73 (0.54, 1.00)
Arrhythmia	Unexposed	510, 5.97, 85.40 (78.30,	1.00 (., .)	1.00 (., .)
	Tamoxifen	231, 4.09, 56.52 (49.68,	0.71 (0.60, 0.83)	0.72 (0.61, 0.84)
	AI	1285, 18.65, 68.91 (65.24,	0.91 (0.81, 1.01)	0.86 (0.77, 0.97)
	Past with AI	456, 5.62, 81.09 (73.98,	1.02 (0.90, 1.17)	1.02 (0.89, 1.17)
	Past Tam only	79, 0.98, 80.35 (64.45,	0.92 (0.73, 1.18)	0.99 (0.77, 1.26)
HF	Unexposed	488, 6.77, 72.07 (65.95,	1.00 (., .)	1.00 (., .)
	Tamoxifen	217, 4.62, 47.00 (41.14,	0.78 (0.66, 0.92)	0.77 (0.65, 0.91)
	AI	1058, 21.14, 50.04 (47.12,	0.90 (0.80, 1.00)	0.79 (0.70, 0.89)
	Past with AI	412, 6.69, 61.59 (55.92,	1.02 (0.89, 1.17)	0.99 (0.86, 1.14)
	Past Tam only	81, 1.19, 68.32 (54.95, 84.94)	0.96 (0.76, 1.23)	1.05 (0.82, 1.34)
Pericarditis	Unexposed	74, 9.48, 7.80 (6.21, 9.80)	1.00 (., .)	1.00 (., .)
	Tamoxifen	18, 5.89, 3.06 (1.93, 4.85)	0.41 (0.25, 0.70)	0.42 (0.25, 0.70)
	AI	159, 27.66, 5.75 (4.92, 6.72)	0.75 (0.56, 1.00)	0.67 (0.50, 0.90)
	Past with AI	50, 9.21, 5.43 (4.11, 7.16)	0.66 (0.46, 0.97)	0.65 (0.45, 0.96)
	Past Tam only	-, -, 3.77 (1.69, 8.39)	0.41 (0.16, 1.01)	0.43 (0.17, 1.06)
VHD	Unexposed	447, 6.08, 73.51 (67.00,	1.00 (., .)	1.00 (., .)
	Tamoxifen	224, 4.01, 55.84 (48.99,	0.82 (0.69, 0.96)	0.82 (0.69, 0.97)
	AI	1181, 18.20, 64.91 (61.31,	0.97 (0.87, 1.09)	0.92 (0.82, 1.04)
	Past with AI	423, 5.68, 74.43 (67.66,	1.08 (0.94, 1.24)	1.11 (0.97, 1.28)
	Past Tam only	65, 1.03, 63.11 (49.49, 80.47)	0.83 (0.64, 1.09)	0.92 (0.70, 1.20)
VTE	Unexposed	78, 9.34, 8.35 (6.69, 10.43)	1.00 (., .)	1.00 (., .)
	Tamoxifen	75, 5.88, 12.75 (10.17, 15.99)	1.71 (1.23, 2.38)	1.68 (1.21, 2.35)
	AI	222, 27.31, 8.13 (7.13, 9.27)	1.15 (0.87, 1.51)	1.07 (0.81, 1.42)
	Past with AI	87, 8.95, 9.72 (7.88, 11.99)	1.28 (0.93, 1.76)	1.27 (0.91, 1.76)
	Past Tam only	12, 1.56, 7.68 (4.36, 13.52)	0.98 (0.53, 1.81)	1.04 (0.56, 1.92)
DVT	Unexposed	72, 9.36, 7.69 (6.11, 9.69)	1.00 (., .)	1.00 (., .)
	Tamoxifen	71, 5.90, 12.04 (9.54, 15.20)	1.77 (1.26, 2.48)	1.74 (1.23, 2.46)
	AI	207, 27.35, 7.57 (6.60, 8.67)	1.17 (0.88, 1.55)	1.11 (0.83, 1.49)
	Past with AI	81, 8.96, 9.04 (7.27, 11.24)	1.29 (0.92, 1.80)	1.31 (0.93, 1.84)
	Past Tam only	-, -, 7.03 (3.89, 12.69)	0.98 (0.52, 1.86)	1.04 (0.55, 1.98)
PE	Unexposed	-, -, 0.93 (0.48, 1.78)	1.00 (., .)	1.00 (., .)
	Tamoxifen	-, -, 0.67 (0.25, 1.78)	0.79 (0.24, 2.58)	0.77 (0.22, 2.53)
	AI	20, 28.27, 0.71 (0.46, 1.10)	0.86 (0.38, 1.95)	0.68 (0.29, 1.57)
	Past with AI	-, -, 0.96 (0.50, 1.84)	1.14 (0.45, 2.91)	0.97 (0.37, 2.53)
	Past Tam only	-, -, 1.25 (0.31, 4.98)	1.38 (0.30, 6.43)	1.51 (0.32, 7.14)

*Events and follow-up suppressed if number of events ≤ 11

† HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

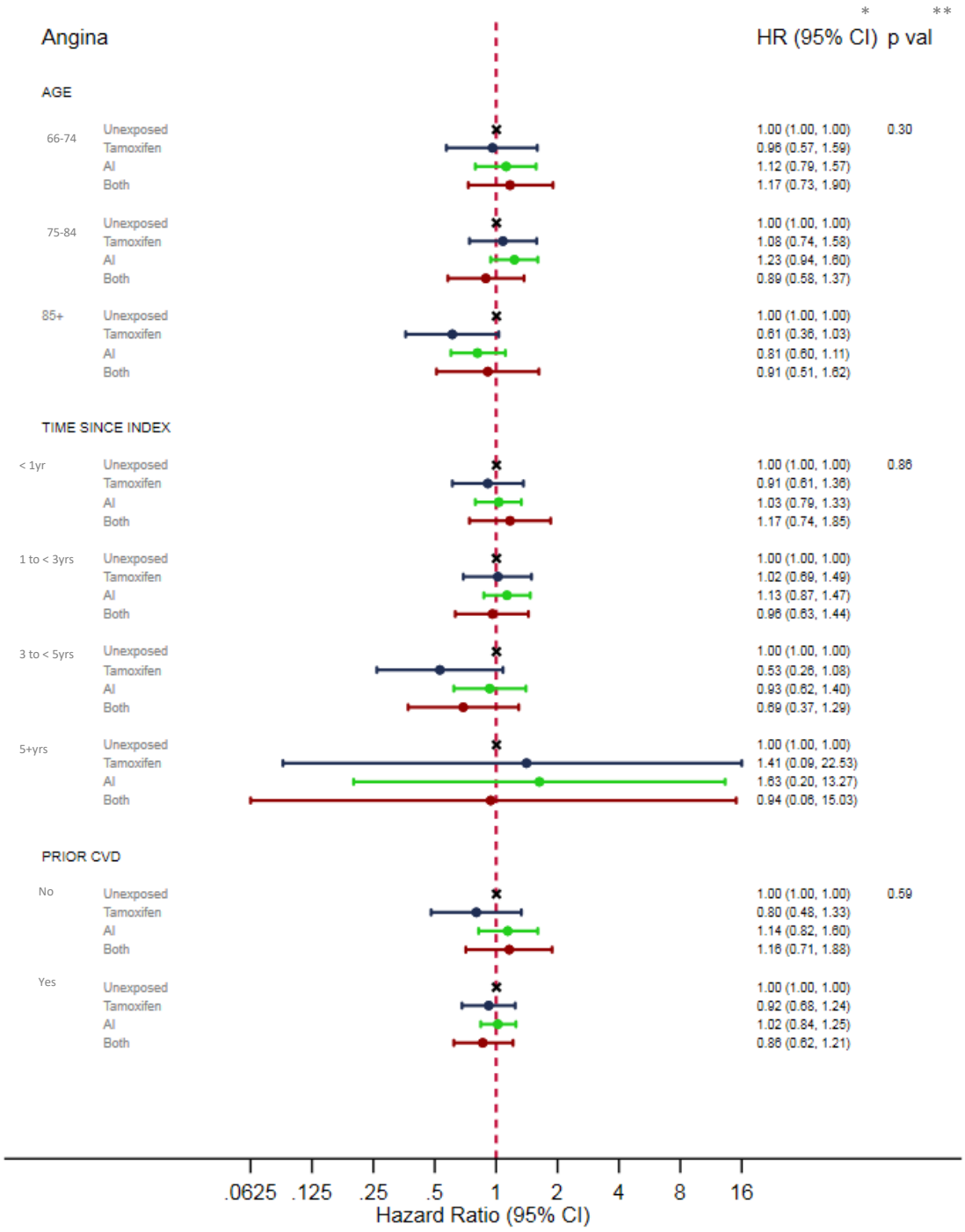
Appendix 6.9 – Adjusted HRs for association between ever exposure to endocrine therapies and risk of coronary artery disease, stratified by age, time since index, and prior CVD



* HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

** p value from Wald test for interaction

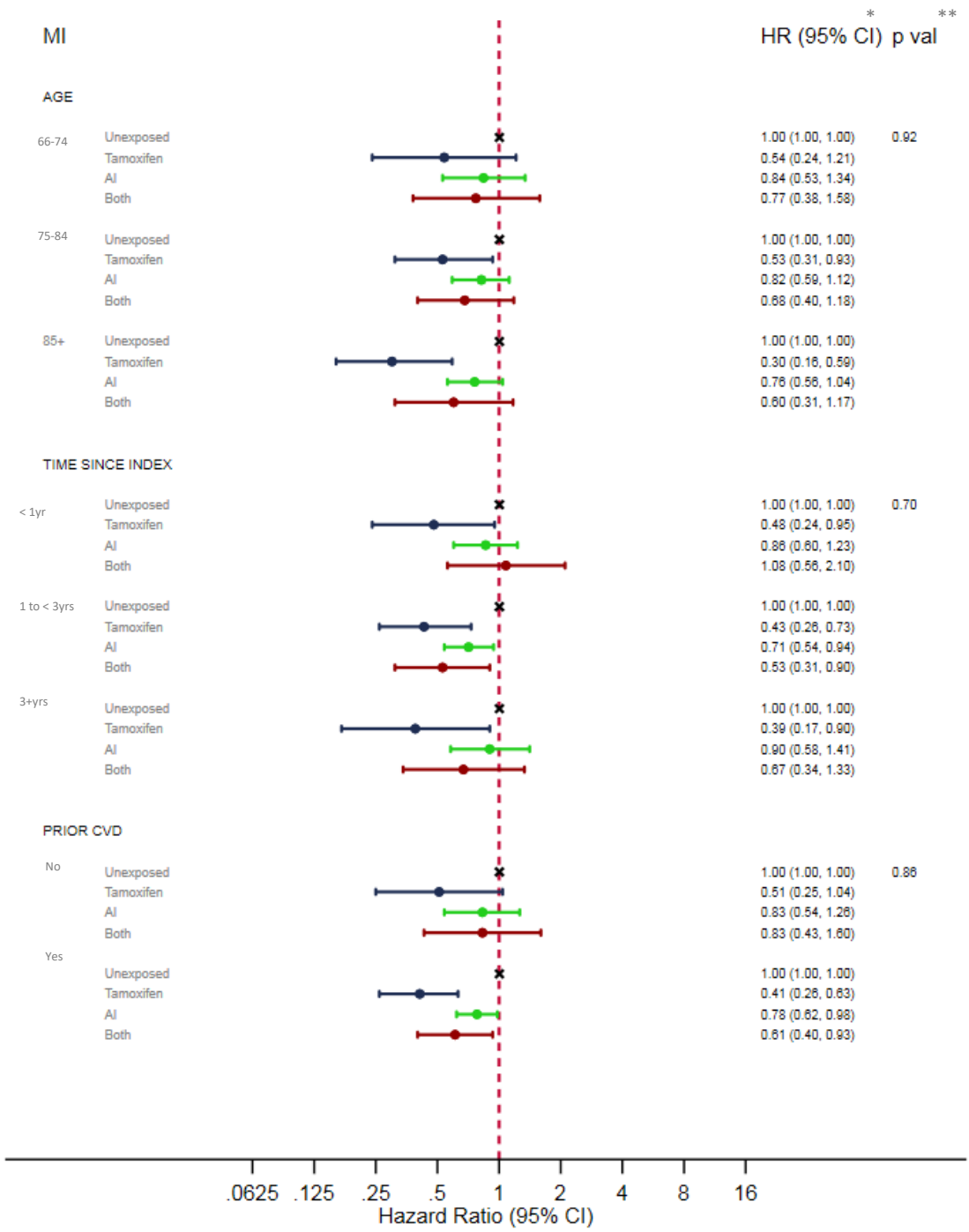
Appendix 6.10 – Adjusted HRs for association between ever exposure to endocrine therapies and risk of angina, stratified by age, time since index, and prior CVD



* HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date <1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

** p value from Wald test for interaction

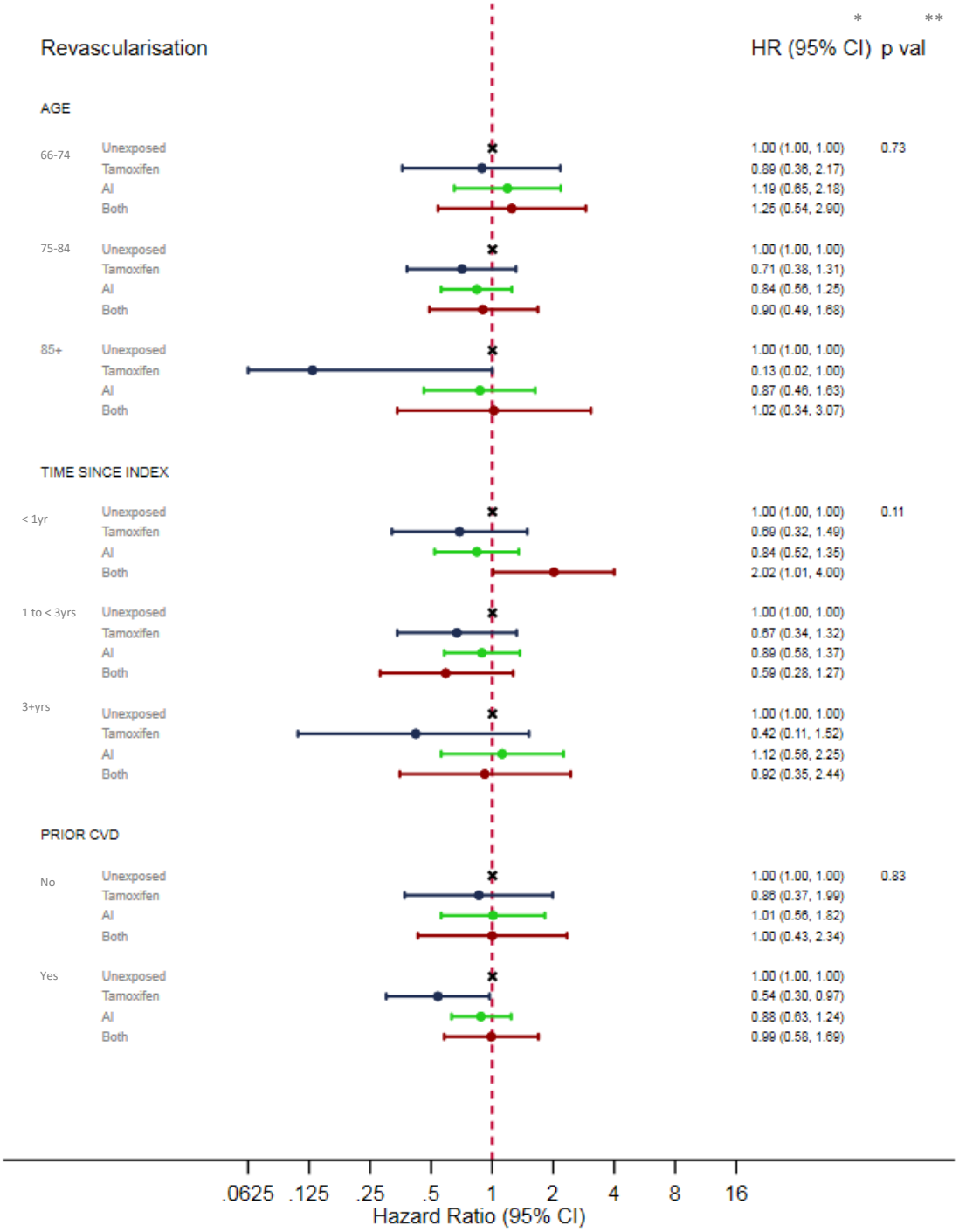
Appendix 6.11 – Adjusted HRs for association between ever exposure to endocrine therapies and risk of MI, stratified by age, time since index, and prior CVD



* HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

** p value from Wald test for interaction

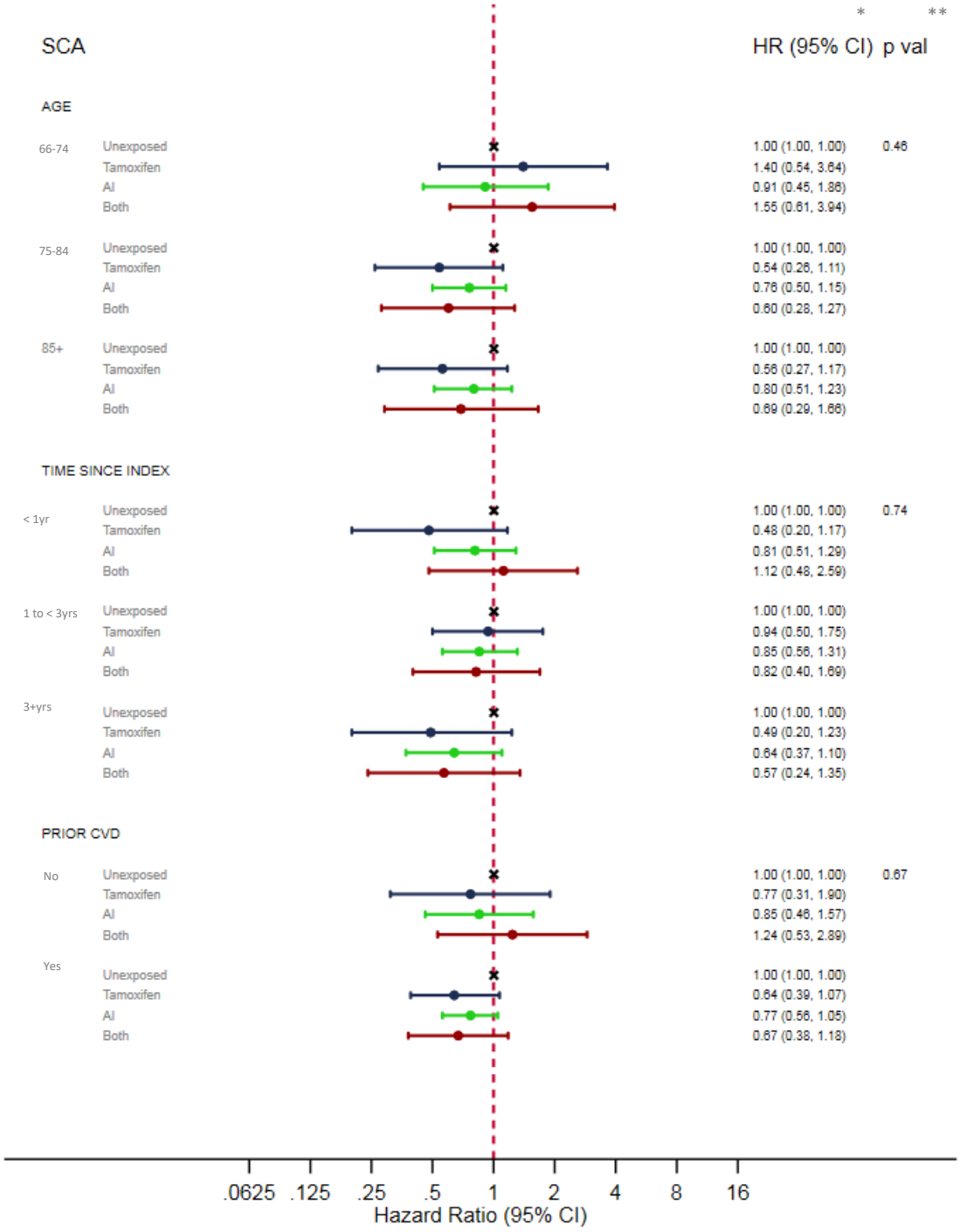
Appendix 6.12 – Adjusted HRs for association between ever exposure to endocrine therapies and risk of revascularisation, stratified by age, time since index, and prior CVD



* HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

** p value from Wald test for interaction

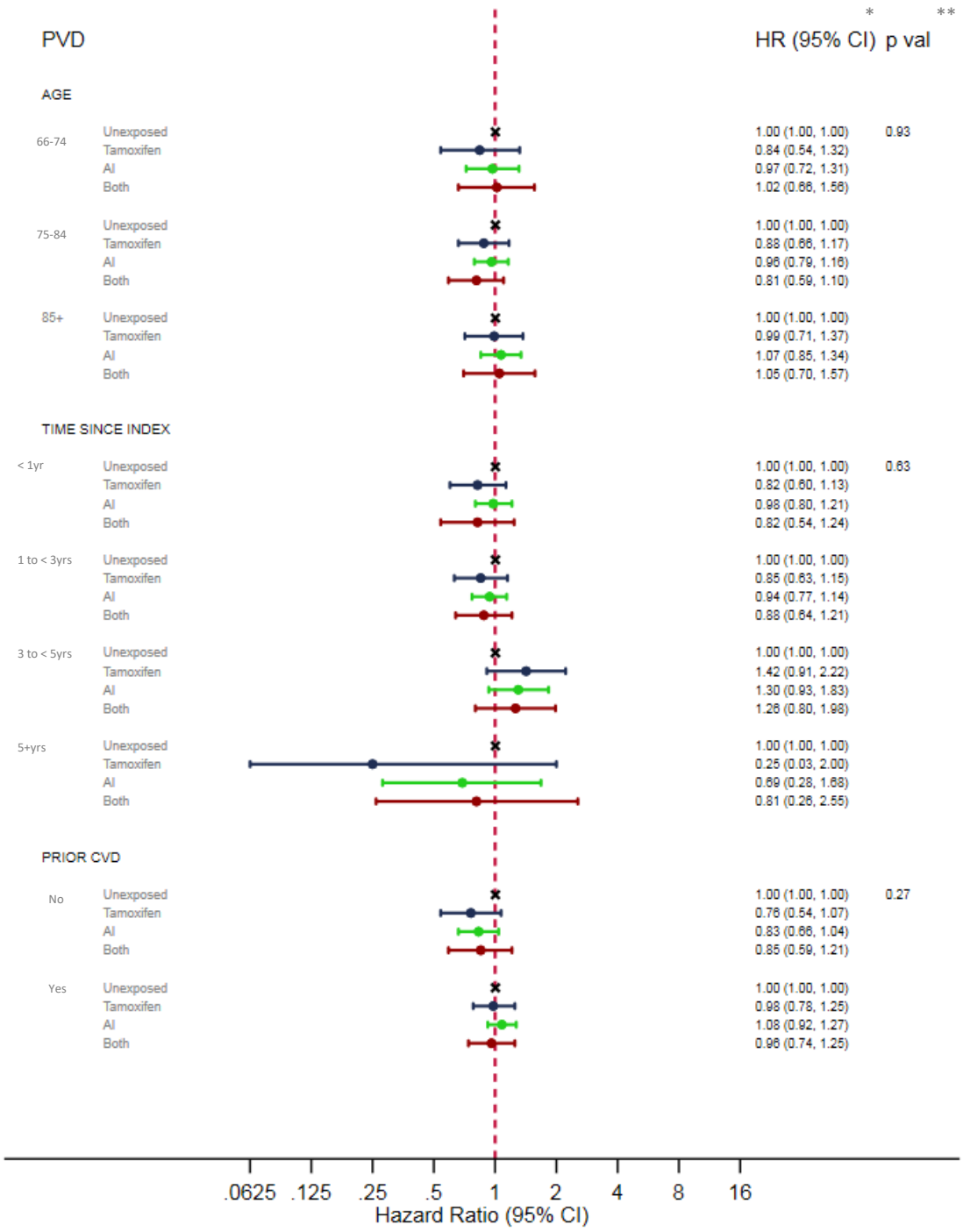
Appendix 6.13 – Adjusted HRs for association between ever exposure to endocrine therapies and risk of SCA, stratified by age, time since index, and prior CVD



* HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

** p value from Wald test for interaction

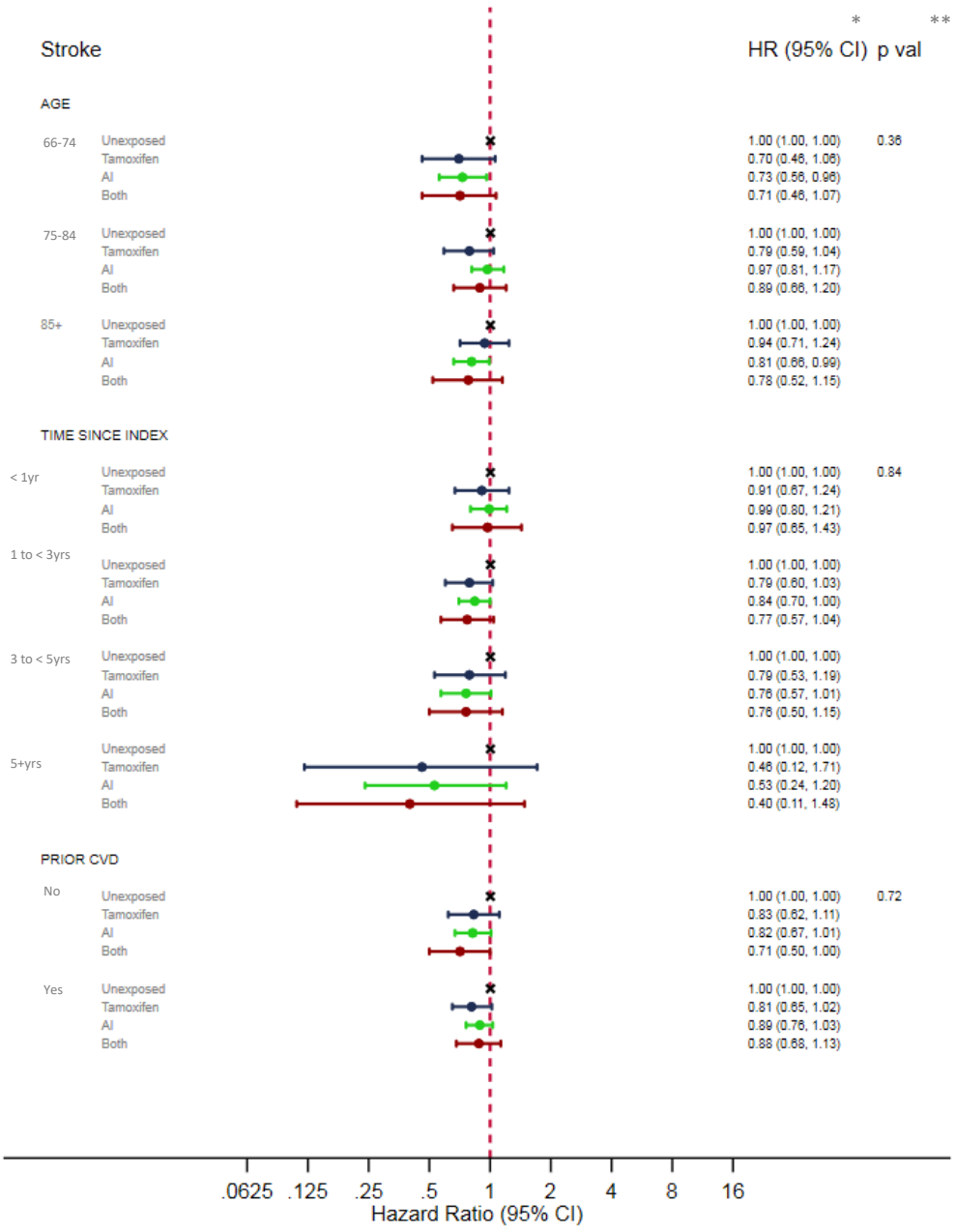
Appendix 6.14 – Adjusted HRs for association between ever exposure to endocrine therapies and risk of PVD, stratified by age, time since index, and prior CVD



* HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

** p value from Wald test for interaction

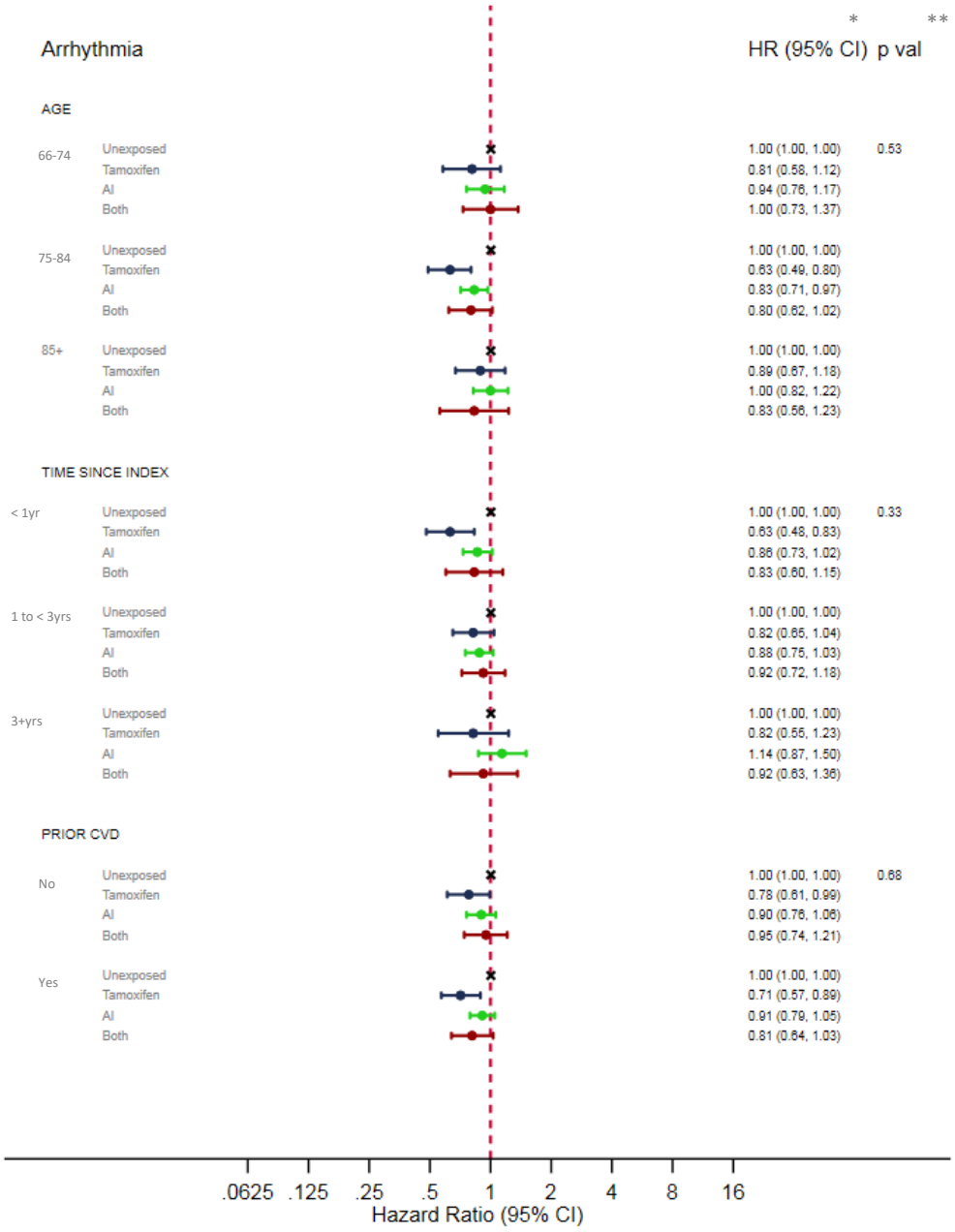
Appendix 6.15 – Adjusted HRs for association between ever exposure to endocrine therapies and risk of stroke, stratified by age, time since index, and prior CVD



* HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

** p value from Wald test for interaction

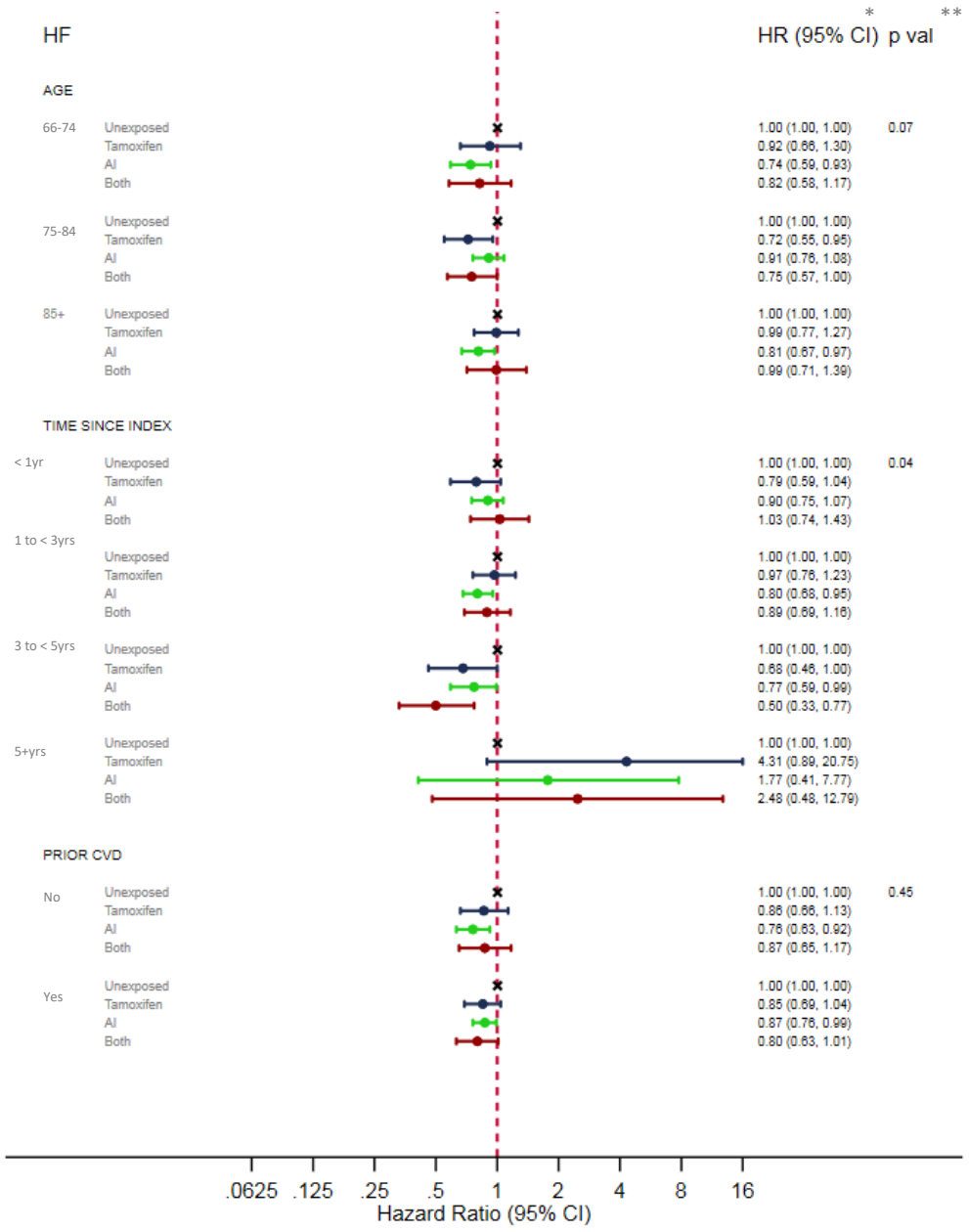
Appendix 6.16 – Adjusted HRs for association between ever exposure to endocrine therapies and risk of arrhythmia, stratified by age, time since index, and prior CVD



* HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date <1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

** p value from Wald test for interaction

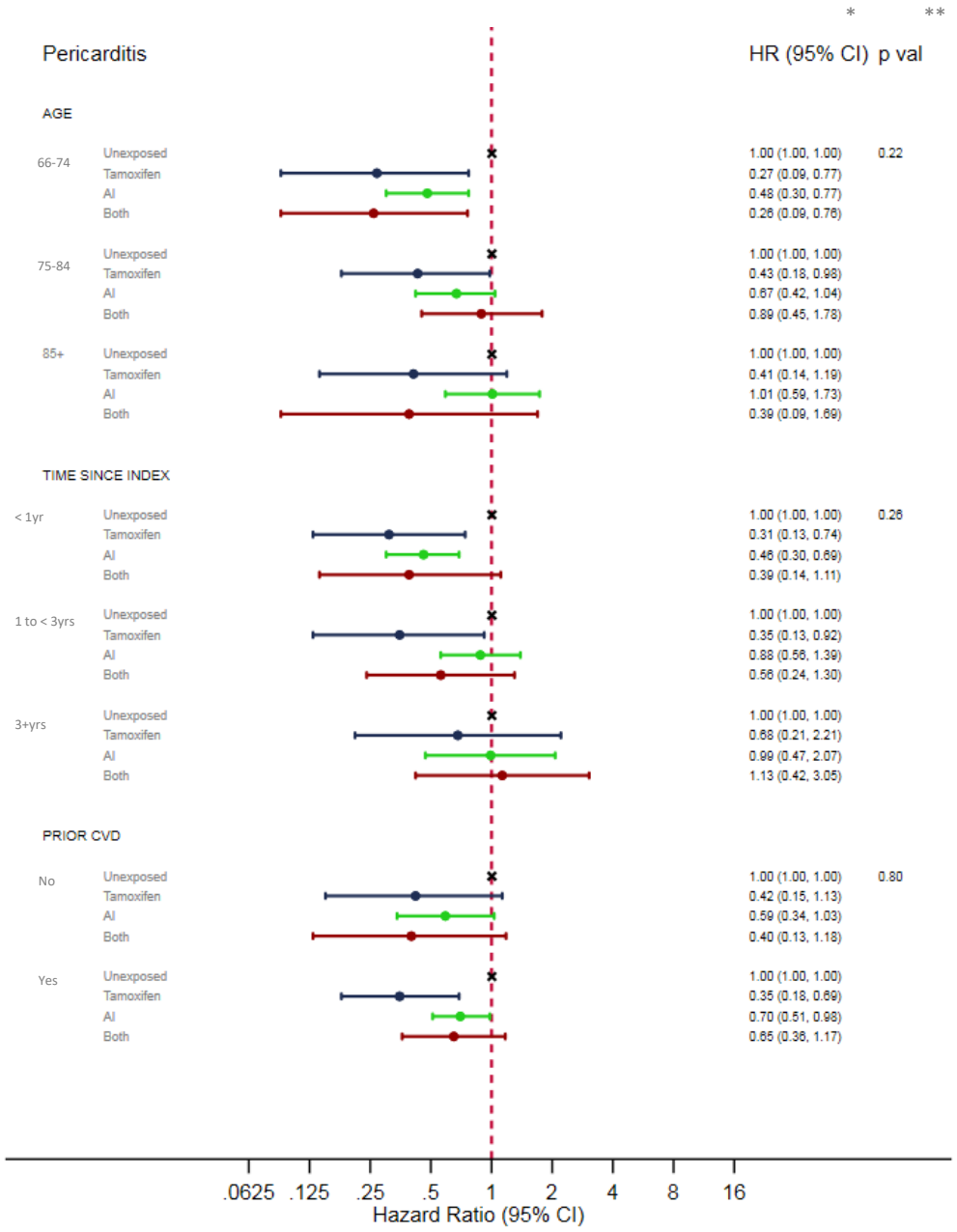
Appendix 6.17 – Adjusted HRs for association between ever exposure to endocrine therapies and risk of HF, stratified by age, time since index, and prior CVD



* HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

** p value from Wald test for interaction

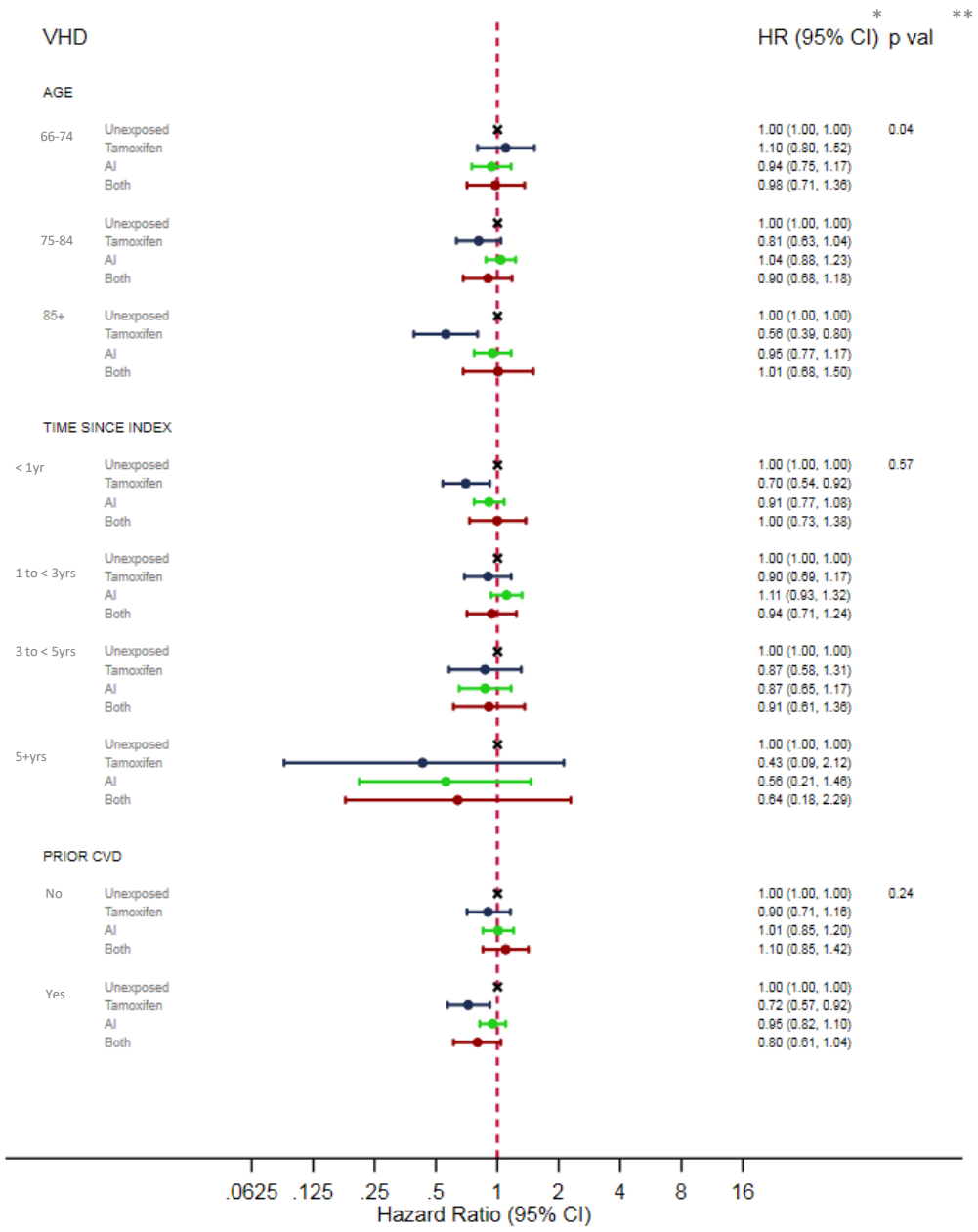
Appendix 6.18 – Adjusted HRs for association between ever exposure to endocrine therapies and risk of pericarditis, stratified by age, time since index, and prior CVD



* HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

** p value from Wald test for interaction

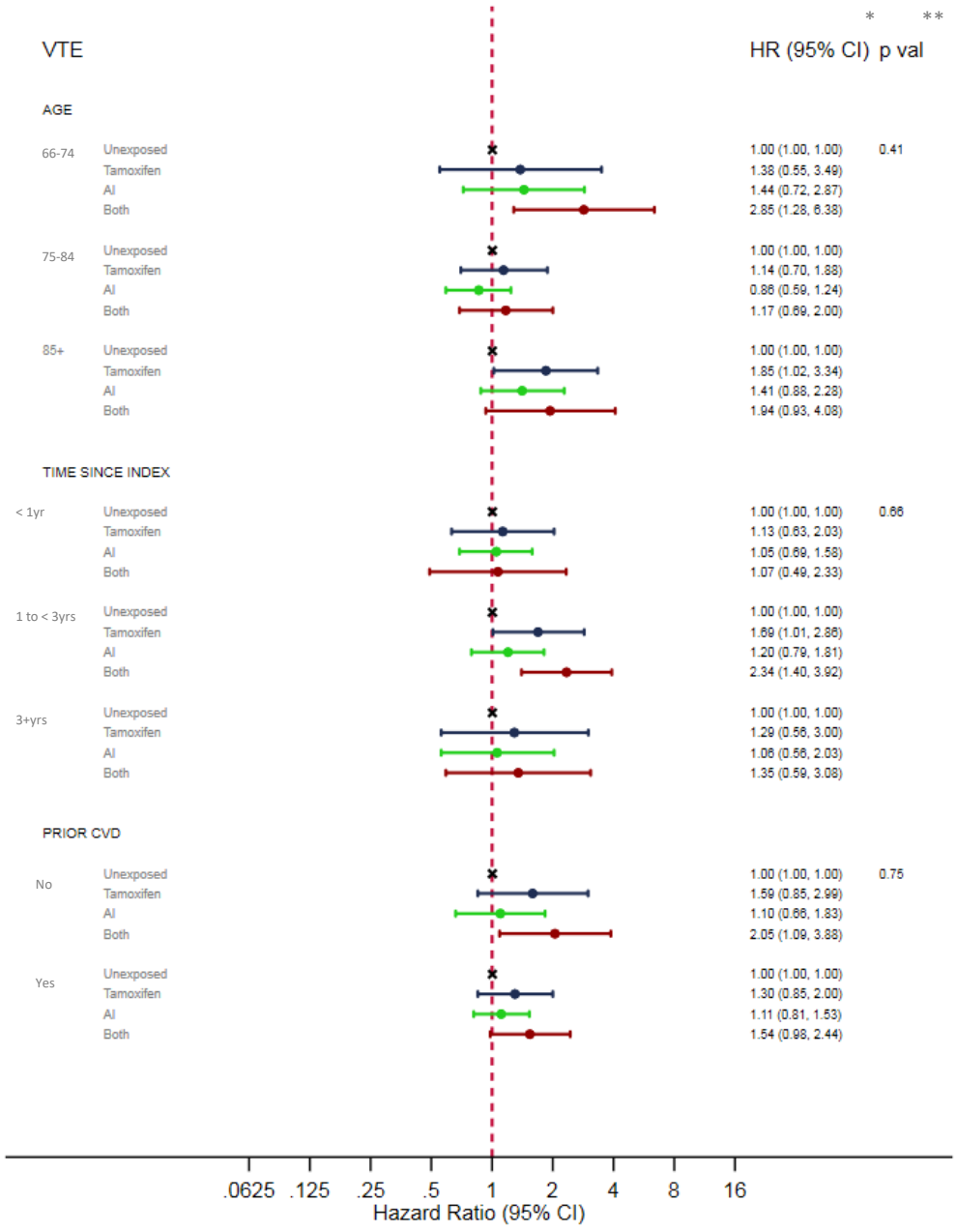
Appendix 6.19 – Adjusted HRs for association between ever exposure to endocrine therapies and risk of VHD, stratified by age, time since index, and prior CVD



* HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

** p value from Wald test for interaction

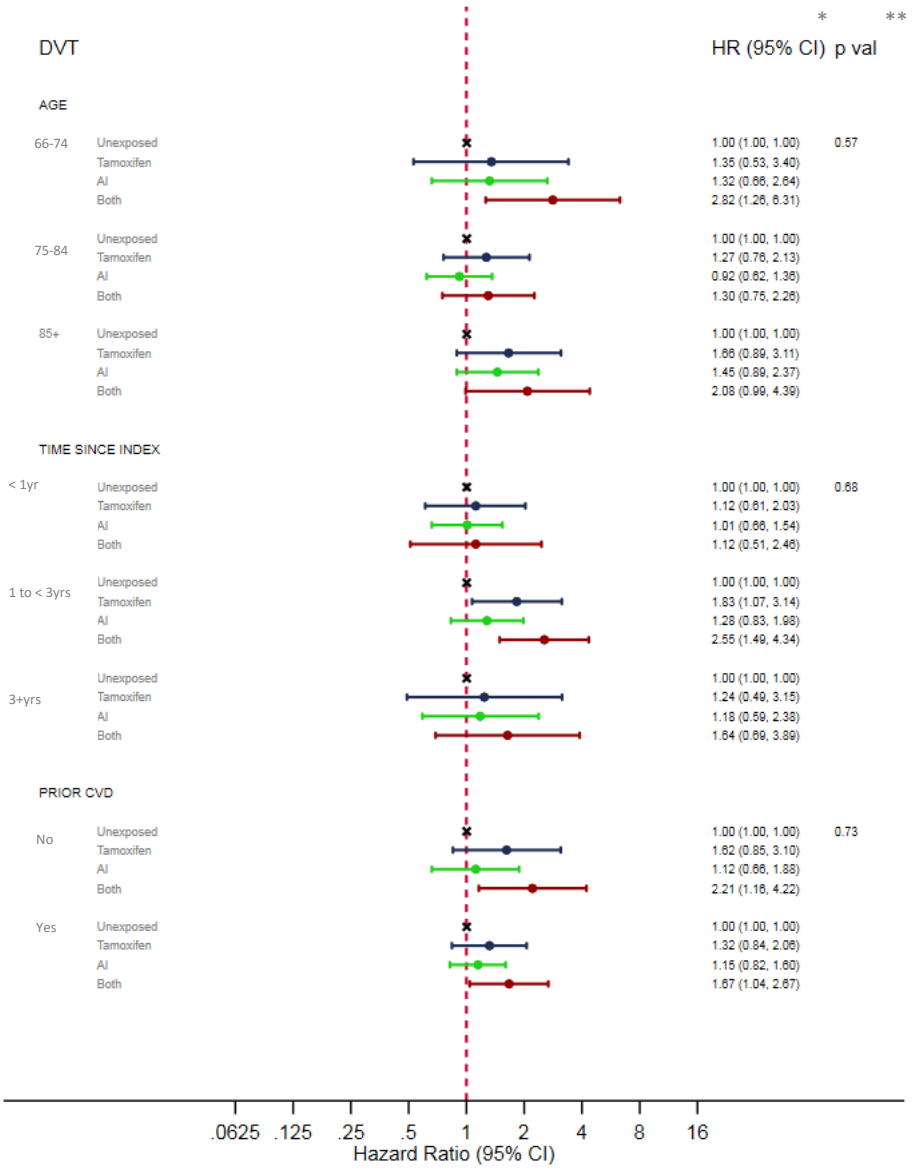
Appendix 6.20 – Adjusted HRs for association between ever exposure to endocrine therapies and risk of VTE, stratified by age, time since index, and prior CVD



* HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

** p value from Wald test for interaction

Appendix 6.21 – Adjusted HRs for association between ever exposure to endocrine therapies and risk of DVT, stratified by age, time since index, and prior CVD



* HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

** p value from Wald test for interaction

Appendix 6.22 - Adjusted HRs for the association between ever exposure to endocrine therapy and a range of clinical CVD outcomes, directly comparing the risk in AI and tamoxifen users

Outcome	Ever tamoxifen vs unexposed to any endocrine therapy (original study population)	Ever AI vs unexposed to any endocrine therapy (original study population)	Ever AI vs ever tamoxifen (original study population)	Ever AI vs ever tamoxifen (study population restricted to only those with tamoxifen/AI prescription)
Coronary Artery Disease	0.74 (0.60, 0.92)	0.96 (0.83, 1.10)	1.29 (1.06, 1.55)	1.35 (1.13, 1.63)
Angina	0.88 (0.68, 1.14)	1.05 (0.89, 1.25)	1.19 (0.95, 1.50)	1.27 (1.01, 1.58)
MI	0.44 (0.30, 0.63)	0.79 (0.64, 0.97)	1.81 (1.28, 2.58)	1.88 (1.34, 2.64)
Revascularisation	0.63 (0.39, 1.01)	0.91 (0.68, 1.23)	1.46 (0.95, 2.24)	1.38 (0.91, 2.09)
SCA	0.67 (0.43, 1.04)	0.78 (0.59, 1.04)	1.17 (0.78, 1.76)	1.29 (0.87, 1.91)
PVD	0.91 (0.75, 1.10)	1.00 (0.87, 1.14)	1.10 (0.92, 1.31)	1.13 (0.96, 1.34)
Stroke	0.82 (0.69, 0.98)	0.87 (0.76, 0.98)	1.05 (0.90, 1.24)	1.07 (0.92, 1.25)
Arrhythmia	0.75 (0.63, 0.88)	0.91 (0.81, 1.01)	1.22 (1.05, 1.41)	1.21 (1.05, 1.40)
HF	0.87 (0.74, 1.02)	0.84 (0.75, 0.94)	0.96 (0.83, 1.12)	1.01 (0.87, 1.16)
Pericarditis	0.37 (0.21, 0.65)	0.67 (0.50, 0.90)	1.81 (1.06, 3.08)	1.97 (1.18, 3.29)
VHD	0.81 (0.68, 0.96)	0.98 (0.87, 1.09)	1.21 (1.04, 1.41)	1.25 (1.08, 1.45)
VTE	1.39 (0.98, 1.98)	1.11 (0.84, 1.46)	0.80 (0.59, 1.07)	0.85 (0.64, 1.13)
DVT	1.42 (0.98, 2.04)	1.14 (0.86, 1.52)	0.81 (0.59, 1.09)	0.85 (0.63, 1.15)
PE	1.05 (0.35, 3.16)	0.80 (0.36, 1.79)	0.77 (0.29, 2.04)	0.95 (0.36, 2.48)

† HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

Appendix 6.23 - Sensitivity analysis – Adjusted HRs for the association between current exposure to endocrine therapy and the risk of a rand of CVDs, with different grace periods, where the grace period is the maximum gap in prescription coverage for exposure to be considered continuous

Outcome	Current exposure	Adjusted HR – 1m grace	Adjusted HR – 3m grace	Adjusted HR – 6m grace	Adjusted HR – 1y grace
Coronary Artery Disease	Unexposed	1.00 (., .)	1.00 (., .)	1.00 (., .)	1.00 (., .)
	Tamoxifen	0.77 (0.63, 0.94)	0.75 (0.62, 0.92)	0.78 (0.64, 0.95)	0.79 (0.66, 0.96)
	AI	0.89 (0.77, 1.02)	0.91 (0.79, 1.05)	0.92 (0.80, 1.06)	0.93 (0.81, 1.07)
	Past with AI	1.11 (0.94, 1.32)	1.08 (0.90, 1.28)	1.06 (0.87, 1.28)	1.02 (0.82, 1.27)
	Past Tam only	0.85 (0.61, 1.19)	0.85 (0.60, 1.22)	0.81 (0.55, 1.20)	0.78 (0.49, 1.23)
Angina	Unexposed	1.00 (., .)	1.00 (., .)	1.00 (., .)	1.00 (., .)
	Tamoxifen	0.90 (0.70, 1.15)	0.86 (0.67, 1.10)	0.90 (0.71, 1.15)	0.93 (0.74, 1.17)
	AI	1.03 (0.86, 1.23)	1.02 (0.86, 1.22)	1.03 (0.86, 1.22)	1.03 (0.87, 1.23)
	Past with AI	1.10 (0.89, 1.36)	1.11 (0.89, 1.39)	1.10 (0.86, 1.40)	1.10 (0.83, 1.45)
	Past Tam only	0.93 (0.61, 1.41)	0.98 (0.63, 1.52)	0.94 (0.58, 1.51)	0.74 (0.40, 1.36)
MI	Unexposed	1.00 (., .)	1.00 (., .)	1.00 (., .)	1.00 (., .)
	Tamoxifen	0.47 (0.33, 0.67)	0.49 (0.35, 0.70)	0.50 (0.36, 0.70)	0.49 (0.36, 0.68)
	AI	0.71 (0.57, 0.88)	0.73 (0.59, 0.91)	0.75 (0.61, 0.93)	0.77 (0.63, 0.95)
	Past with AI	0.98 (0.77, 1.26)	0.94 (0.72, 1.22)	0.91 (0.69, 1.21)	0.85 (0.61, 1.17)
	Past Tam only	0.46 (0.25, 0.85)	0.43 (0.22, 0.85)	0.39 (0.18, 0.84)	0.44 (0.19, 1.00)
Revascularisation	Unexposed	1.00 (., .)	1.00 (., .)	1.00 (., .)	1.00 (., .)
	Tamoxifen	0.89 (0.59, 1.35)	0.89 (0.59, 1.33)	0.84 (0.56, 1.27)	0.84 (0.56, 1.25)
	AI	0.89 (0.65, 1.21)	0.88 (0.65, 1.20)	0.89 (0.66, 1.20)	0.89 (0.66, 1.21)
	Past with AI	0.89 (0.62, 1.30)	0.90 (0.61, 1.34)	0.89 (0.58, 1.35)	0.84 (0.52, 1.36)
	Past Tam only	0.54 (0.23, 1.27)	0.53 (0.21, 1.31)	0.61 (0.24, 1.52)	0.66 (0.24, 1.83)
SCA	Unexposed	1.00 (., .)	1.00 (., .)	1.00 (., .)	1.00 (., .)
	Tamoxifen	0.58 (0.36, 0.92)	0.63 (0.41, 0.98)	0.66 (0.43, 1.02)	0.67 (0.44, 1.01)
	AI	0.64 (0.48, 0.87)	0.72 (0.54, 0.96)	0.73 (0.55, 0.98)	0.75 (0.57, 1.00)
	Past with AI	1.11 (0.80, 1.54)	0.93 (0.65, 1.34)	0.94 (0.64, 1.39)	0.87 (0.55, 1.36)
	Past Tam only	1.31 (0.76, 2.27)	1.24 (0.69, 2.25)	1.11 (0.57, 2.16)	1.34 (0.67, 2.70)
PVD	Unexposed	1.00 (., .)	1.00 (., .)	1.00 (., .)	1.00 (., .)
	Tamoxifen	0.85 (0.70, 1.03)	0.84 (0.69, 1.01)	0.87 (0.72, 1.05)	0.93 (0.77, 1.11)
	AI	0.95 (0.82, 1.09)	0.97 (0.85, 1.12)	0.98 (0.86, 1.13)	0.99 (0.86, 1.13)
	Past with AI	1.10 (0.93, 1.29)	1.06 (0.89, 1.26)	1.01 (0.83, 1.22)	0.99 (0.80, 1.23)
	Past Tam only	1.21 (0.91, 1.60)	1.28 (0.96, 1.70)	1.29 (0.95, 1.75)	1.06 (0.72, 1.55)
Stroke	Unexposed	1.00 (., .)	1.00 (., .)	1.00 (., .)	1.00 (., .)
	Tamoxifen	0.83 (0.69, 0.99)	0.82 (0.69, 0.98)	0.82 (0.69, 0.97)	0.83 (0.70, 0.98)
	AI	0.81 (0.71, 0.93)	0.82 (0.72, 0.93)	0.84 (0.74, 0.95)	0.85 (0.75, 0.96)
	Past with AI	1.02 (0.87, 1.18)	1.02 (0.87, 1.20)	0.99 (0.83, 1.17)	0.92 (0.75, 1.13)
	Past Tam only	0.73 (0.54, 1.00)	0.74 (0.53, 1.02)	0.77 (0.55, 1.09)	0.80 (0.55, 1.18)
Arrhythmia	Unexposed	1.00 (., .)	1.00 (., .)	1.00 (., .)	1.00 (., .)
	Tamoxifen	0.72 (0.61, 0.84)	0.73 (0.63, 0.86)	0.75 (0.64, 0.87)	0.76 (0.65, 0.88)
	AI	0.86 (0.77, 0.97)	0.89 (0.79, 0.99)	0.89 (0.80, 1.00)	0.89 (0.80, 0.99)
	Past with AI	1.02 (0.89, 1.17)	0.98 (0.85, 1.13)	0.95 (0.81, 1.10)	0.96 (0.81, 1.15)
	Past Tam only	0.99 (0.77, 1.26)	0.94 (0.72, 1.22)	0.96 (0.73, 1.28)	1.11 (0.81, 1.50)
HF	Unexposed	1.00 (., .)	1.00 (., .)	1.00 (., .)	1.00 (., .)
	Tamoxifen	0.77 (0.65, 0.91)	0.82 (0.70, 0.96)	0.82 (0.70, 0.96)	0.81 (0.69, 0.94)
	AI	0.79 (0.70, 0.89)	0.80 (0.71, 0.90)	0.82 (0.73, 0.92)	0.83 (0.74, 0.93)
	Past with AI	0.99 (0.86, 1.14)	0.96 (0.83, 1.12)	0.94 (0.80, 1.10)	0.95 (0.79, 1.14)
	Past Tam only	1.05 (0.82, 1.34)	1.01 (0.77, 1.31)	1.00 (0.76, 1.33)	1.06 (0.78, 1.46)
Pericarditis	Unexposed	1.00 (., .)	1.00 (., .)	1.00 (., .)	1.00 (., .)
	Tamoxifen	0.42 (0.25, 0.70)	0.42 (0.25, 0.70)	0.42 (0.25, 0.69)	0.38 (0.23, 0.63)
	AI	0.60 (0.50, 0.90)	0.70 (0.52, 0.94)	0.69 (0.51, 0.93)	0.70 (0.52, 0.94)
	Past with AI	0.65 (0.45, 0.96)	0.57 (0.37, 0.88)	0.58 (0.37, 0.92)	0.50 (0.28, 0.89)
	Past Tam only	0.43 (0.17, 1.06)	0.48 (0.19, 1.20)	0.57 (0.23, 1.41)	0.76 (0.30, 1.89)
VHD	Unexposed	1.00 (., .)	1.00 (., .)	1.00 (., .)	1.00 (., .)
	Tamoxifen	0.82 (0.69, 0.97)	0.82 (0.70, 0.97)	0.83 (0.71, 0.98)	0.82 (0.70, 0.96)
	AI	0.92 (0.82, 1.04)	0.93 (0.83, 1.05)	0.94 (0.84, 1.06)	0.95 (0.85, 1.07)
	Past with AI	1.11 (0.97, 1.28)	1.10 (0.95, 1.28)	1.13 (0.96, 1.32)	1.17 (0.97, 1.40)
	Past Tam only	0.92 (0.70, 1.20)	0.95 (0.72, 1.26)	0.85 (0.62, 1.16)	1.00 (0.71, 1.40)
VTE	Unexposed	1.00 (., .)	1.00 (., .)	1.00 (., .)	1.00 (., .)
	Tamoxifen	1.68 (1.21, 2.35)	1.70 (1.22, 2.35)	1.68 (1.21, 2.32)	1.59 (1.15, 2.19)
	AI	1.07 (0.81, 1.42)	1.09 (0.83, 1.45)	1.09 (0.83, 1.44)	1.12 (0.85, 1.48)
	Past with AI	1.27 (0.91, 1.76)	1.19 (0.84, 1.68)	1.22 (0.84, 1.77)	1.10 (0.71, 1.69)
	Past Tam only	1.04 (0.56, 1.92)	0.90 (0.45, 1.80)	0.81 (0.37, 1.76)	0.94 (0.40, 2.16)
DVT	Unexposed	1.00 (., .)	1.00 (., .)	1.00 (., .)	1.00 (., .)
	Tamoxifen	1.74 (1.23, 2.46)	1.77 (1.26, 2.48)	1.75 (1.25, 2.45)	1.66 (1.19, 2.31)
	AI	1.11 (0.83, 1.49)	1.14 (0.85, 1.52)	1.14 (0.85, 1.52)	1.15 (0.87, 1.54)
	Past with AI	1.31 (0.93, 1.84)	1.22 (0.85, 1.75)	1.23 (0.83, 1.81)	1.18 (0.76, 1.84)
	Past Tam only	1.04 (0.55, 1.98)	0.88 (0.42, 1.83)	0.76 (0.33, 1.75)	0.85 (0.34, 2.13)
PE	Unexposed	1.00 (., .)	1.00 (., .)	1.00 (., .)	1.00 (., .)
	Tamoxifen	0.77 (0.22, 2.53)	0.74 (0.22, 2.44)	0.70 (0.21, 2.30)	0.65 (0.20, 2.14)
	AI	0.68 (0.29, 1.57)	0.71 (0.31, 1.64)	0.67 (0.29, 1.54)	0.77 (0.34, 1.74)
	Past with AI	0.97 (0.37, 2.53)	0.97 (0.34, 2.72)	1.23 (0.44, 3.47)	0.79 (0.21, 3.07)
	Past Tam only	1.51 (0.32, 7.14)	1.80 (0.38, 8.55)	2.19 (0.46, 10.41)	2.81 (0.59, 13.49)

ǂ HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84, 85+); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

Appendix 6.24 - Sensitivity analysis – Crude rates, unadjusted HRs, and adjusted HRs for the association between ever exposure to endocrine therapies and the risk of a range of clinical CVDs, excluding women over 85

Outcome	Ever exposure	Events, Follow-up, Rate (95% CI) (per 1000 pyears) *	Unadjusted HR (95% CI)	Adjusted HR (95% CI)
Coronary Artery	Unexposed	177, 5.55, 31.86 (27.50, 36.92)	1.00 (., .)	1.00 (., .)
Disease	Tamoxifen	96, 3.68, 26.08 (21.36, 31.86)	0.83 (0.65, 1.07)	0.85 (0.66, 1.10)
	AI	849, 25.84, 32.85 (30.71, 35.14)	1.06 (0.90, 1.25)	1.00 (0.84, 1.18)
	Both	91, 3.33, 27.30 (22.23, 33.53)	0.85 (0.66, 1.10)	0.88 (0.68, 1.15)
Angina	Unexposed	111, 5.65, 19.64 (16.31, 23.65)	1.00 (., .)	1.00 (., .)
	Tamoxifen	69, 3.72, 18.55 (14.65, 23.49)	0.99 (0.73, 1.34)	1.03 (0.76, 1.39)
	AI	604, 26.30, 22.96 (21.20, 24.87)	1.19 (0.97, 1.47)	1.18 (0.96, 1.46)
	Both	62, 3.41, 18.20 (14.19, 23.35)	0.95 (0.69, 1.30)	1.01 (0.73, 1.39)
MI	Unexposed	74, 6.32, 11.71 (9.33, 14.71)	1.00 (., .)	1.00 (., .)
	Tamoxifen	27, 4.13, 6.53 (4.48, 9.52)	0.55 (0.35, 0.86)	0.52 (0.33, 0.83)
	AI	298, 29.35, 10.15 (9.06, 11.38)	0.93 (0.71, 1.20)	0.80 (0.61, 1.04)
	Both	32, 3.71, 8.63 (6.11, 12.21)	0.74 (0.49, 1.14)	0.70 (0.46, 1.08)
Revascularisation	Unexposed	46, 6.26, 7.34 (5.50, 9.80)	1.00 (., .)	1.00 (., .)
	Tamoxifen	24, 4.09, 5.86 (3.93, 8.74)	0.79 (0.48, 1.31)	0.75 (0.45, 1.25)
	AI	207, 29.27, 7.07 (6.17, 8.10)	0.94 (0.68, 1.31)	0.95 (0.68, 1.33)
	Both	26, 3.67, 7.08 (4.82, 10.39)	0.94 (0.57, 1.54)	1.01 (0.61, 1.67)
SCA	Unexposed	40, 6.48, 6.17 (4.53, 8.42)	1.00 (., .)	1.00 (., .)
	Tamoxifen	18, 4.19, 4.29 (2.70, 6.81)	0.73 (0.42, 1.27)	0.74 (0.42, 1.31)
	AI	161, 30.07, 5.35 (4.59, 6.25)	0.88 (0.62, 1.26)	0.77 (0.54, 1.11)
	Both	20, 3.79, 5.28 (3.40, 8.18)	0.83 (0.48, 1.46)	0.89 (0.50, 1.57)
PVD	Unexposed	200, 5.43, 36.81 (32.05, 42.28)	1.00 (., .)	1.00 (., .)
	Tamoxifen	107, 3.57, 29.98 (24.81, 36.24)	0.84 (0.66, 1.07)	0.86 (0.68, 1.09)
	AI	847, 25.08, 33.77 (31.58, 36.13)	0.99 (0.85, 1.16)	0.96 (0.82, 1.13)
	Both	97, 3.28, 29.57 (24.23, 36.07)	0.83 (0.65, 1.07)	0.87 (0.67, 1.12)
Stroke	Unexposed	224, 5.33, 42.06 (36.90, 47.94)	1.00 (., .)	1.00 (., .)
	Tamoxifen	116, 3.68, 31.49 (26.25, 37.78)	0.77 (0.61, 0.97)	0.76 (0.60, 0.96)
	AI	879, 25.20, 34.88 (32.65, 37.27)	0.94 (0.81, 1.10)	0.88 (0.75, 1.03)
	Both	104, 3.24, 32.11 (26.50, 38.92)	0.84 (0.66, 1.07)	0.83 (0.65, 1.06)
Arrhythmia	Unexposed	328, 4.23, 77.46 (69.51, 86.31)	1.00 (., .)	1.00 (., .)
	Tamoxifen	155, 2.92, 53.08 (45.34, 62.13)	0.67 (0.55, 0.81)	0.69 (0.57, 0.84)
	AI	1346, 20.14, 66.85 (63.37, 70.51)	0.90 (0.79, 1.02)	0.86 (0.76, 0.98)
	Both	159, 2.53, 62.82 (53.78, 73.39)	0.82 (0.68, 1.00)	0.86 (0.71, 1.05)
HF	Unexposed	260, 4.88, 53.29 (47.19, 60.18)	1.00 (., .)	1.00 (., .)
	Tamoxifen	139, 3.34, 41.60 (35.23, 49.13)	0.78 (0.63, 0.96)	0.80 (0.65, 0.99)
	AI	1063, 23.01, 46.19 (43.50, 49.05)	0.92 (0.80, 1.06)	0.83 (0.72, 0.95)
	Both	126, 3.06, 41.12 (34.53, 48.97)	0.79 (0.64, 0.99)	0.78 (0.63, 0.98)
Pericarditis	Unexposed	51, 6.37, 8.01 (6.09, 10.54)	1.00 (., .)	1.00 (., .)
	Tamoxifen	12, 4.16, 2.89 (1.64, 5.08)	0.33 (0.17, 0.64)	0.35 (0.18, 0.67)
	AI	155, 29.46, 5.26 (4.50, 6.16)	0.63 (0.45, 0.87)	0.55 (0.39, 0.76)
	Both	17, 3.75, 4.53 (2.82, 7.29)	0.54 (0.31, 0.94)	0.54 (0.31, 0.96)
VHD	Unexposed	283, 4.33, 65.37 (58.19, 73.45)	1.00 (., .)	1.00 (., .)
	Tamoxifen	163, 2.93, 55.63 (47.71, 64.86)	0.88 (0.72, 1.07)	0.90 (0.74, 1.10)
	AI	1245, 19.65, 63.35 (59.93, 66.97)	1.02 (0.89, 1.16)	0.98 (0.86, 1.12)
	Both	143, 2.53, 56.44 (47.91, 66.49)	0.88 (0.72, 1.08)	0.92 (0.75, 1.14)
VTE	Unexposed	49, 6.26, 7.83 (5.92, 10.36)	1.00 (., .)	1.00 (., .)
	Tamoxifen	37, 4.13, 8.96 (6.49, 12.37)	1.18 (0.76, 1.82)	1.17 (0.76, 1.81)
	AI	225, 29.10, 7.73 (6.78, 8.81)	1.03 (0.75, 1.42)	1.00 (0.72, 1.38)
	Both	42, 3.66, 11.46 (8.47, 15.51)	1.53 (1.00, 2.34)	1.56 (1.01, 2.40)
DVT	Unexposed	44, 6.27, 7.02 (5.23, 9.44)	1.00 (., .)	1.00 (., .)
	Tamoxifen	36, 4.14, 8.71 (6.28, 12.07)	1.28 (0.82, 2.01)	1.27 (0.81, 2.00)
	AI	206, 29.16, 7.07 (6.16, 8.10)	1.05 (0.75, 1.46)	1.03 (0.73, 1.45)
	Both	41, 3.67, 11.18 (8.23, 15.18)	1.66 (1.07, 2.57)	1.70 (1.09, 2.66)
PE	Unexposed	-, -, 0.92 (0.41, 2.06)	1.00 (., .)	1.00 (., .)
	Tamoxifen	-, -, 0.24 (0.03, 1.69)	0.26 (0.03, 2.17)	0.27 (0.03, 2.27)
	AI	25, 30.11, 0.83 (0.56, 1.23)	0.96 (0.39, 2.36)	0.84 (0.33, 2.14)
	Both	-, -, 0.26 (0.04, 1.86)	0.31 (0.04, 2.57)	0.31 (0.04, 2.67)

*Events and follow-up supressed if number of events ≤ 11

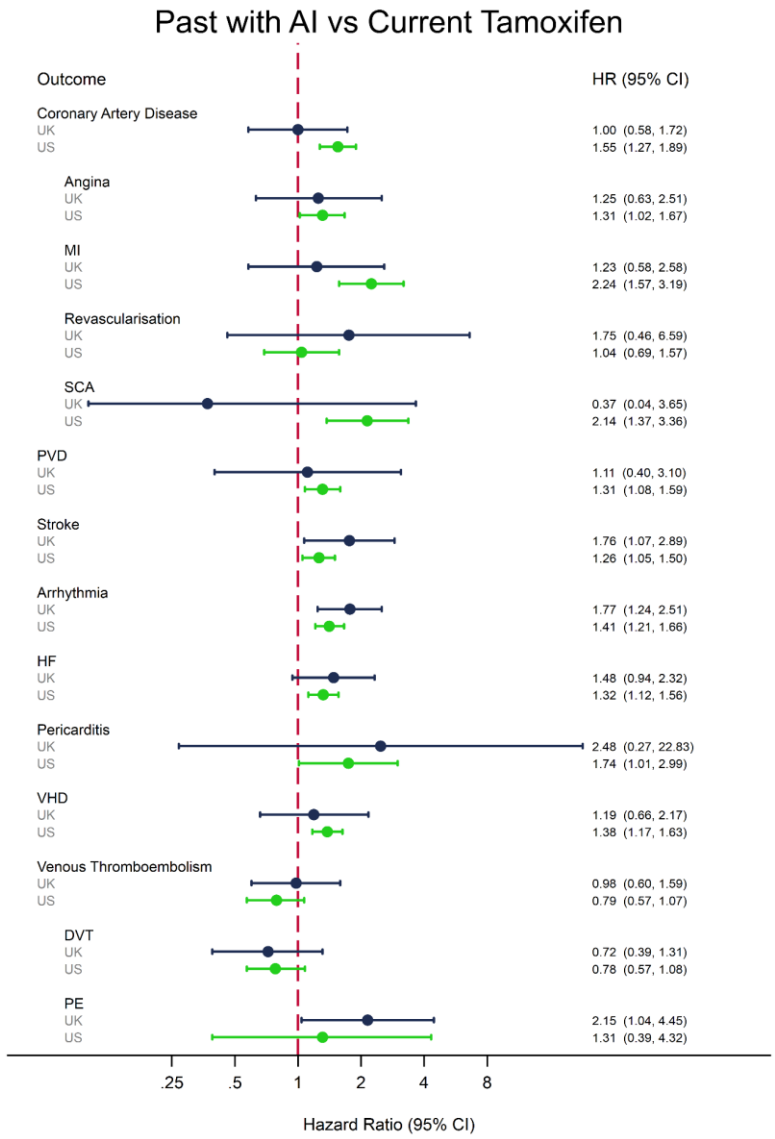
† HRs adjusted for: year of breast cancer diagnosis; age at index date (66-74, 75-84); race (White, Black Asian, Hispanic, Native American, other); SEER region (North East, South, North Central, West); breast cancer stage (1-3); breast cancer grade (1-3); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of taxanes, anthracyclines, trastuzumab, other systemic cancer treatments, statins, anti-hypertensive drugs, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers; diagnosis of rheumatoid arthritis, chronic kidney disease, hypertension, diabetes, VTE, and non-venous CVD

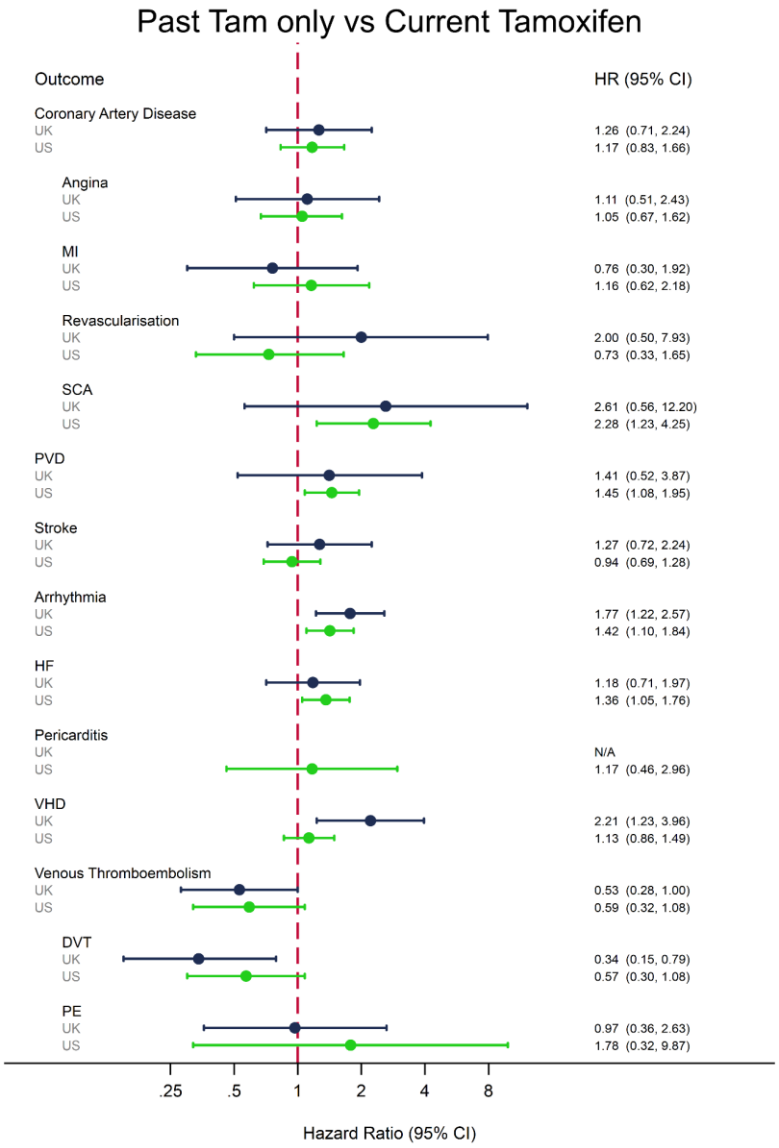
CHAPTER 7

Appendix 7.1 – HRs for the association between ever AI use compared with ever tamoxifen use and a range of clinical CVD outcomes, first unadjusted, then individually adjusted for antiplatelet use

Outcome	Adjustments	HR (95% CI)	Percentage difference between crude and adjusted (%)
Coronary artery disease	Crude	1.40 (1.06, 1.86)	4.3
	Adjusted for antiplatelet use	1.34 (1.01, 1.77)	
Angina	Crude	1.43 (1.00, 2.05)	7
	Adjusted for antiplatelet use	1.33 (0.93, 1.90)	
MI	Crude	1.73 (1.12, 2.68)	4
	Adjusted for antiplatelet use	1.66 (1.07, 2.57)	
Revascularisation	Crude	1.36 (0.69, 2.68)	4.4
	Adjusted for antiplatelet use	1.30 (0.66, 2.57)	
SCA	Crude	1.87 (0.83, 4.22)	1.6
	Adjusted for antiplatelet use	1.84 (0.82, 4.16)	
PVD	Crude	1.25 (0.77, 2.01)	1.6
	Adjusted for antiplatelet use	1.23 (0.76, 1.99)	
Stroke	Crude	1.22 (0.91, 1.63)	2.5
	Adjusted for antiplatelet use	1.19 (0.89, 1.59)	
Arrhythmia	Crude	1.37 (1.14, 1.64)	2.2
	Adjusted for antiplatelet use	1.34 (1.11, 1.61)	
HF	Crude	1.87 (1.43, 2.45)	6.4
	Adjusted for antiplatelet use	1.75 (1.34, 2.29)	
Pericarditis	Crude	3.96 (1.12, 14.03)	1.3
	Adjusted for antiplatelet use	3.91 (1.10, 13.89)	
VHD	Crude	1.58 (1.15, 2.17)	3.8
	Adjusted for antiplatelet use	1.52 (1.11, 2.10)	
VTE	Crude	0.92 (0.71, 1.20)	1.1
	Adjusted for antiplatelet use	0.91 (0.70, 1.19)	
DVT	Crude	0.69 (0.48, 0.98)	1.4
	Adjusted for antiplatelet use	0.68 (0.48, 0.97)	
PE	Crude	1.39 (0.95, 2.04)	0
	Adjusted for antiplatelet use	1.39 (0.95, 2.03)	

Appendix 7.2 – Adjusted HRs for the association between current exposure to endocrine therapy and a range of clinical CVD outcomes in restricted analyses in both the UK and US





†Adjusted for: year of breast cancer; age at index date (66-74, 75-84, 85+); time since index date (<1yr, 1 to <3yrs, 3 to <5yrs, 5+yrs); current calendar year; use of statins; use of ACE inhibitors; use of calcium channel blockers; use of angiotensin receptor blockers; rheumatoid arthritis; chronic kidney disease; diabetes; VTE; and non-venous CVD

Appendix 7.3 – HF QOF Read code list

Read code	Read term
585f.00	echocardiogram shows left ventricular systolic dysfunction
662f.00	new york heart association classification - class i
662g.00	new york heart association classification - class ii
662h.00	new york heart association classification - class iii
662i.00	new york heart association classification - class iv
G58..00	heart failure
G58..11	cardiac failure
G5yy900	left ventricular systolic dysfunction
G5yyD00	left ventricular cardiac dysfunction
G1yz100	rheumatic left ventricular failure

REFERENCES - APPENDICES

1. StataCorp, *Stata Statistical Software: Release 15*. 2017, StataCorp LLC: College Station, TX.
2. Levey, A.S., et al., *A new equation to estimate glomerular filtration rate.* Ann Intern Med, 2009. **150**(9): p. 604-12.
3. Fox, C.S., et al., *Associations of kidney disease measures with mortality and end-stage renal disease in individuals with and without diabetes: a meta-analysis.* Lancet, 2012. **380**(9854): p. 1662-73.
4. Bhaskaran, K., et al., *Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5.24 million UK adults.* Lancet, 2014. **384**(9945): p. 755-65.
5. WHO, *Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee. WHO Technical Report Series 854. Geneva: World Health Organization, 1995. .* WHO Techical Report, 1995. **Series 854. Geneva**.
6. Noble, M.M., et al., *The English Indices of Deprivation 2010.* D.f.C.a.L. Government, 2011.