

ABSTRACT

Title of Dissertation: METFORMIN AND STATINS IN RELATION TO OVARIAN CANCER: IMPLICATIONS FOR RISK AND POTENTIAL ETIOLOGIC PATHWAYS

Sarah Renee Irvin, Doctor of Philosophy, 2022

Dissertation directed by: Dr. Cher Dallal,
Associate Professor, Epidemiology and
Biostatistics

Background: Metabolic factors, including obesity, diabetes, and circulating lipids are associated with increased risk of ovarian cancer, the most fatal gynecologic malignancy. Drugs targeting these conditions have drawn interest as potential chemopreventive agents, but limited studies have examined relationships with ovarian cancer overall, by possible benefit groups or with regards to related etiologic pathways. This dissertation examined medications that target metabolic abnormalities with risk of ovarian cancer, and expression of a pro-tumorigenic lipid pathway in ovarian cancer tissue.

Methods: Nested case-control studies of metformin (Aim 1) and statin (Aim 2) use in relation to ovarian cancer (cases n=9,207 and 10:1 frequency-matched controls n=92,070) were conducted in Clinical Practice Research Datalink (CPRD) (1987-2020), a primary care database in the United Kingdom. Aim 3 was a case-only analysis from The Polish Ovarian Cancer Study (POCS) (2001-2003, n=166) determining expression of sphingosine-1-phosphate receptors

S1PR1 and S1PR3 in ovarian cancer tissue. Multivariable logistic regression estimated odds ratios (ORs) and 95% confidence intervals (CI) for associations between medication use and ovarian cancer in CPRD (Aims 1 and 2), and cancer characteristics with marker expression in POCS (Aim 3).

Results: Within the CPRD analyses, metabolic conditions such as obesity (OR 1.19, 95% CI 1.11-1.28) and polycystic ovarian syndrome (OR 1.66, 95% CI 1.17-2.35) were significantly associated with increased ovarian cancer risk. However, neither metformin (OR 1.01, 95% CI 0.91-1.13) nor statin use (OR 1.02, 95% CI 0.96-1.08) was associated with ovarian cancer risk. Associations did not differ by drug indications ($p_{\text{int}} > 0.05$), individual statin or class. In Aim 3, strong expression of lipid markers in ovarian cancer tissue significantly differed by histotype ($p < 0.01$). Lower marker expression was observed in non-serous tumors compared to serous, but expression was not associated with cancer risk factors or survival.

Conclusion: Findings from this dissertation do not support use of metformin or statins as chemopreventive agents. However, as most CPRD users had an indication, it remains unclear whether findings are generalizable to the at-risk population for ovarian cancer. Strong expression of lipid markers in ovarian cancer tissue suggests involvement in carcinogenesis; additional molecular studies are needed to elucidate their role.

METFORMIN AND STATINS IN RELATION TO OVARIAN CANCER:
IMPLICATIONS FOR RISK AND POTENTIAL ETIOLOGIC PATHWAYS

by

Sarah Renee Irvin, MSc

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:

Associate Professor Dr. Cher Dallal, Chair

Dr. Nicolas Wentzensen, Mentor, National Cancer Institute (NCI)

Associate Professor Dr. Typhanye Dyer

Assistant Professor Dr. Tianzhou (Charles) Ma

Associate Professor Dr. Paul Turner

Professor Dr. James Hagberg, Dean's Representative

© Copyright by
Sarah Renee Irvin
2022

Table of Contents

Table of Contents	iii
List of Tables	vii
List of Figures	viii
List of Abbreviations	ix
Chapter 1: Introduction	1
Background and rationale	1
Significance and innovation	7
Chapter 2: Background	8
Classification of ovarian cancer	8
A primer on staging and grading of ovarian cancers	8
Epithelial ovarian cancer histotypes and cells of origin	9
Epidemiology of ovarian cancer	10
Mortality and morbidity	10
Risk factors for epithelial ovarian cancer	11
Ovarian cancer etiologic hypotheses	18
Gonadotropin signaling and incessant ovulation	18
Inflammation and lipid dysregulation	19
Description of the exposures: metformin	20
Metformin drug characteristics	20
Additional indications for metformin	22
Recommendations for metformin prescriptions	22
Existing knowledge	23
Description of the exposures: statins	26
Statin formulations	29
Recommendations for statin prescriptions	30
Existing knowledge	31
Description of the exposures: S1P pathway markers	39
Existing knowledge	40
Chapter 3: Methods	43
Aims 1 and 2: CPRD	43
Study design	43

Study population	44
Data management	46
Selection, information and confounding bias	47
Statistical analysis	50
Power analysis	50
Theoretical directed acyclic graphs (DAGs) Aims 1 and 2	52
Outcome, exposure and covariate window considerations in electronic medical record (EMR)-derived data	53
Aim 3: The Polish Ovarian Cancer Study (POCS)	58
Data assembly	59
Selection, information and confounding bias	59
Statistical analysis	61
Chapter 4: Manuscript 1	64
Abstract	64
Introduction	66
Methods	67
Study population	67
Outcome definition	68
Exposure definition	68
Study covariates	69
Statistical analysis	70
Results	72
Study characteristics by 10-year birth cohort	72
Primary association between metformin and ovarian cancer risk	73
Association between metformin use and ovarian cancer risk by BMI	74
Evaluation of metformin and ovarian cancer risk by indication group	75
Joint effects of metformin and concomitant medications	76
Discussion	76
Tables for manuscript 1	82
Chapter 5: Manuscript 2	89
Abstract	89
Introduction	91
Methods	92
Study population	Error! Bookmark not defined.

Outcome definition	93
Exposure definition	94
Study covariates	94
Statistical methods	96
Results	97
Study characteristics by 10-year birth cohort	98
Primary association between statin use and ovarian cancer risk	99
Evaluation of statin use and ovarian cancer risk by indication groups	100
Evaluation of statin use and ovarian cancer risk by other factors	101
Joint effects of statins and concomitant medications	102
Discussion	103
Tables for manuscript 2	108
Chapter 6: Manuscript 3	116
Abstract	116
Introduction	118
Methods	120
Study population	120
Tissue specimen collection and creation of tissue micro arrays (TMAs)	121
Pilot evaluation for S1P pathway marker antibodies	121
Processing of TMAs and scoring	122
Covariates	123
Statistical analysis	123
Results	125
Characteristics of the study population	125
Descriptive characteristics of S1PR1 and S1PR3 marker expression overall and by histotype	125
Associations between ovarian cancer risk factors and expression of S1PR1 and S1PR3	128
Evaluation of survival by S1PR1 and S1PR3 expression status	128
Discussion	130
Tables and figures for Manuscript 3	134
Chapter 7: Public Health Significance	141
Conclusions	141
Public Health Implications	141
Future directions	143

Appendices	144
Bibliography	210

List of Tables

Table 1 Summary of studies related to reproductive and demographic risk factors of ovarian cancer	15
Table 2 Summary of estimates for associations between metabolic conditions and risk of ovarian cancer	17
Table 3 Studies reporting effect measures for the association between metformin use and risk of ovarian cancer	24
Table 4 Studies assessing the association between statin use and ovarian cancer risk.....	36
Table 5. Ovarian cancer histotype-specific estimates by study	38
Table 6. Key S1P pathway markers.....	42
Table 7. Characteristics of ovarian cancer cases and controls in CPRD	82
Table 8. Associations between metformin use and ovarian cancer risk in CPRD.....	83
Table 9. Association between metformin use and ovarian cancer risk in CPRD by decade of birth.....	84
Table 10. Association between metformin and ovarian cancer risk by BMI categories in CPRD	86
Table 11 Associations between metformin and ovarian cancer risk by indication group in CPRD	87
Table 12 Joint effects of metformin use with other diabetes and cardiovascular disease medications in CPRD.....	88
Table 13. Characteristics of ovarian cancer cases and controls in CPRD	108
Table 14. Overall association between statin use and ovarian cancer risk in CPRD.....	109
Table 15. Associations between statin use and ovarian cancer risk in CPRD by decade of birth	110
Table 16. Association between statin dose and duration with ovarian cancer risk in CPRD	112
Table 17. Associations between statin use and ovarian cancer risk by indication group in CPRD	113
Table 18. Joint effects of statin use with other CVD and diabetes medications in CPRD	115
Table 19. Characteristics of ovarian cancer cases in POCS	134
Table 20. Associations between tumor characteristics and S1PR1 and S1PR3 expression	136
Table 21. Associations between S1P receptor staining and ovarian cancer risk factors	137

List of Figures

Figure 1 Ovarian cancer risk factor associations accounting for cell of origin	10
Figure 2. Mechanisms by which metformin regulates glucose.....	21
Figure 3 Statin mechanism of action	26
Figure 4 Prevalence of statin use in the general population among women in the United States in NHANES 1999-2018.....	28
Figure 5. Prevalence of statin use by type in women from the United States in NHANES 1999-2018.....	29
Figure 6. Association between ever-use of statins and ovarian cancer risk.....	33
Figure 7. Association between statin use and risk of ovarian cancer by statin class.....	34
Figure 8. S1P pathway components and possible downstream genetic targets	41
Figure 9. Case selection from CPRD, Manuscripts 1 and 2	45
Figure 10 Power analysis for Manuscripts 1 and 2.....	51
Figure 11. Directed Acyclic Graphs for Manuscripts 1 and 2	52
Figure 12. Ovarian cancer capture by birth cohort	54
Figure 13. Metformin use windows by birth cohort	55
Figure 14. Statin use windows by birth cohort	56
Figure 15. Oral contraceptive use in CPRD according to birth cohort.....	57
Figure 16. Reproductive window overlap with CPRD study window.....	57
Figure 17. Power analysis for Manuscript 3	62
Figure 18. Distributions of area, intensity and staining score measures for S1PR1 and S1PR3	135
Figure 19. Kaplan-Meier overall survival curves for S1PR1 and S1PR3 expression in ovarian cancer tissue	139
Figure 20. S1P depiction of S1P pathway components	140

List of Abbreviations

ADOPT- A Diabetes Outcome Progression Trial

AFCAPS/TexCAPS - Air Force/Texas Coronary Atherosclerosis Prevention Study

AMP - Adenosine monophosphate

ANOVA - Analysis of Variance

ASMR - Age standardized mortality rate

ATP - Adenosine triphosphate

BMI - Body mass index

CA-125 – Cancer Antigen 125

CDC - Centers for Disease Control

CI - Confidence interval

CPRD - Clinical Practice Research Database

CVD - Cardiovascular disease

DAG - Directed acyclic graph

DDD - Defined Daily Doses

DETECT - Targets and Essential Data for Commitment of Treatment

EML - Essential medicines list

EOC - Epithelial ovarian cancer

FDA - Food and Drug Administration

FFPE - Formalin fixed paraffin-embedded

FinDM - Finland Diabetes Database

FPG- Fasting plasma glucose

FSH - Follicle-stimulating hormone

GP - General practice

GPRD - General Practice Research Database

H&E - hematoxylin & eosin

HDL - High-density lipoprotein

HMG-CoA - 3-hydroxy-3-methyl-glutaryl-coenzyme A

HR - Hazard ratio

HRT - Hormone replacement therapy

IHC - Immunohistochemistry

IGF-1 - Insulin-like growth factor

IRB - International Review Board

CPMCP - Kaiser Permanente Medical Care Program

LDL - Low-density lipoprotein

LH - Luteinizing hormone

LOC - Lifetime ovulatory cycles

MetS - Metabolic syndrome

MHT - Menopausal hormone therapy

MMS - Multimodal screening

NCI - National Cancer Institute

NEC - New England Case Control Study

NICE - National Institute for Health Care Excellence

NHANES - National Health and Nutrition Examination Survey

NHS - National Health Service

NHS - Nurse's Health Study

NICE - National Institutes for Healthcare Excellence

OCAC - Ovarian Cancer Awareness Consortium

OC3 - Ovarian Cancer Cohort Consortium

OC - Oral contraceptives

ONS - Office of National Statistics

OR - Odds ratio

OSE - Ovarian surface epithelium

PCOS - Polycystic ovarian syndrome

PDD - Prescribed daily dose

PHI - Personal health information

PID - Pelvic inflammatory disease

PLCO - Prostate, Lung, Colorectal, and Ovarian

POCS - Polish Ovarian Cancer Study

QoF - Quality and Outcomes Framework

Ref - Reference

RECORD - Rosiglitazone Evaluated for Cardiovascular Outcomes and Regulation of Glycaemia
in Diabetes

RR - Risk ratio

SEER - Surveillance, Epidemiology and End Results Program

SIR - Standardized incidence ratio

S1P - Sphingosine-1-phosphate

S1PR - S1P receptor

SPHK - Sphingosine kinase

TAM - Tumor-associated macrophage

THIN - The Health Improvement Network

TMA - Tissue microarray

TNM - T = size of the main tumor, N = number of lymph nodes with cancer, M = metastatic status

Treg - T-regulatory cell

UK - United Kingdom

UMD - University of Maryland

US – United States

USPSTF - United States Preventative Services Task Force

VAMP - Value-Added medical products

WHI - Women's Health Initiative

WHO - World Health Organization

XR - Extended release

YOB - Year of birth

Chapter 1: Introduction

Background and rationale

Ovarian cancer is the fifth leading cause of cancer deaths among women in the United States (U.S) [1]. Although it comprises only 2.5% of incident cancers in women, it accounts for 5% of cancer deaths and is the deadliest gynecologic malignancy [1]. Worldwide, 314,000 new cases and 207,000 deaths were estimated in 2020, constituting 1.6 % of all newly diagnosed cancers and 2.2% of cancer deaths for women < 75 years of age [2].

Importantly, the estimated 5-year survival for women with ovarian cancer is only 30% and due to the rarity of symptoms, screening and treatment options, recurrence rates are high [3-5]. Two large-scale screening trials, one from the U.S., the Prostate, Lung, Colorectal and Ovarian (PLCO) trial with approximately 80,000 U.S. women and a second trial from the United Kingdom (UK) with >200,000 women have shown no improvement in mortality from ovarian cancer after screening with either transvaginal ultrasound or Cancer Antigen 125 (CA-125), a specific protein in the blood that may be higher in women with ovarian cancer, compared to no screening [6, 7]. As such, more than 75% of patients are diagnosed with advanced-stage disease. The complexity of ovarian cancer also contributes to the difficulty in early detection. Ovarian tumors have etiologically distinct histotypes (serous, endometrioid, mucinous, clear cell or other rare or undifferentiated neoplasms). Current literature suggests that serous tumors, the most frequent and aggressive subtype, can be further classified into high- and low-grade neoplasms and that the different histotypes have different tissues of origin. High-grade and low-grade serous tumors arise from the fallopian tube, while endometrioid, mucinous and clear cell tumors

originate from endometriosis tissue, germ cells, or transitional cells [8]. Most women, approximately 60%, will be diagnosed with a serous tumor in an advanced stage [9, 10].

The high proportion of ovarian tumors diagnosed at an advanced stage and the resultant poor outcomes highlight the need for a deeper understanding of factors that could prevent the onset of disease. Parity, oral contraceptive use, aspirin use, and tubal ligation have all been associated with a decreased risk of ovarian cancer [10], while menopausal hormone therapy, endometriosis and pelvic inflammatory disease (PID) increase risk [10-14]. Recent data have also suggested that metabolic conditions such as diabetes, hypercholesterolemia and metabolic syndrome increase the risk of ovarian cancer [15-17]. A large-scale systematic review and meta-analysis of 16 studies reported a positive association between obesity (BMI >30 kg/m²) and ovarian cancer risk (pooled odds ratio (OR), risk ratio (RR), standardized incidence ratio (SIR) 1.30, 95% CI 1.12-1.50) [18]. In addition, a pooled analysis from the Ovarian Cancer Association Consortium (OCAC) reported that being overweight or obese was associated with an increased risk of ovarian cancer for most invasive cancer types [15].

Obesity-related morbidities, like elevated cholesterol levels, have also been linked to increased risk of ovarian cancer, but evidence is mixed [17]. In a meta-analysis of 19 studies assessing the association of diabetes with ovarian cancer, an increased risk was observed (OR 1.17, 95% CI 1.02-1.33 [19]. However, a study of Taiwanese women demonstrated no increased risk of epithelial ovarian cancer for diabetics compared to non-diabetics (HR 1.06, 95% CI: 0.92-1.42) [20]. Metabolic syndrome (MetS), an amalgam of factors that comprises diseases like diabetes, stroke and other cardiovascular complications, has similarly been linked to an increased risk of ovarian cancer in some studies, but not others [17, 21].

Understanding the association between lipid-dysregulating conditions and ovarian cancer is of increasing importance due to their consistently high prevalence. More specifically, the prevalence of obesity and obesity-related morbidities has steadily increased in US adults, rising from 30.5% in 1999 to over 43% in 2018 among women [22, 23]. Subsequently, metabolic conditions have seen dramatic increases. An estimated 16 million women in the US were estimated to have diabetes in 2018, an increase from 9% in 1999 to over 12% in 2016 [24]. Similarly, the prevalence of MetS has also increased alongside the prevalence of obesity. Data from the National Health and Nutrition Examination Survey (NHANES) reported that the prevalence of metabolic syndrome increased steadily from 2011-2016, rising from 31.7% to 36.9% among US women [25].

The possible associations identified between obesity, diabetes and MetS and risk of ovarian cancer, coupled with the increased prevalence of these conditions, suggest that these factors could be a prime target for expansion of chemoprevention options and etiologic research into potential carcinogenic pathways related to metabolic dysregulation. There are several prevention approaches to obesity and obesity-related conditions that may be useful as prevention approaches to other diseases such as cancer. Some studies suggest a potential link between metabolic conditions and the risk of ovarian cancer [26], but the mechanism by which these therapies affect carcinogenesis is unknown. One theory posits that lipid dysregulation, a common feature of obesity and related conditions, may be a key component of the carcinogenic process [27]. Cancers of all types are known to exhibit profound lipid irregularities that promote rapid cellular proliferation and growth. In addition, lipid dysregulation may trigger various inflammatory pathways. The subsequent increase in systemic inflammation has been proposed

as a potential pathway by which metabolic-disturbances exhibit their effects on ovarian carcinogenesis [27, 28].

Evidence in support of the hypothesis that metabolic dysregulation influences carcinogenesis via inflammatory pathways originates from studies on the treatment of metabolic conditions [29]. Typically, insulin resistance, diabetes, metabolic syndrome and hypercholesterolemia are managed by use of pharmaceutical interventions such as metformin or statin drugs. These interventions have specific mechanisms of action [30], but in addition to their primary indications (glucose regulation and cholesterol reduction, respectively) these drugs could also decrease inflammation and potentially influence ovarian cancer risk [31, 32]. Results from a few studies have suggested that metformin [31] and statin [32] drugs could be used for either primary prevention or improved survival in gynecologic cancers [33, 34]. However, the findings are mixed, and more evidence is needed to determine whether lipid-modifying drugs are associated with epithelial ovarian cancers.

The rarity of ovarian cancer is problematic for epidemiologic studies, because low accrual of case numbers may limit the detection of associations with risk factors [3]. The existing epidemiologic literature is limited by few studies for the association of metformin [35-40] and statins [36, 41-49] with ovarian cancer risk, most with small sample sizes. Additionally, prior studies vary regarding exposure ascertainment methods (prescription versus self-report), the ability to control for confounding by indication and limited information on dose and frequencies of medication use. The wide range of risk estimates across populations for both statins and metformin and the lack of homogeneity in analytic designs indicate the need for additional well-powered studies to elucidate true effects.

The overall **objective** of this dissertation was to better understand the association between lipid-related metabolic disturbances and ovarian cancer. The dissertation evaluated two pharmaceutical interventions that act on different biological pathways commonly prescribed to manage diabetes, insulin resistance, and high cholesterol, namely metformin and statin drugs, and their associations with risk of ovarian cancer. This dissertation also evaluated specific tissue markers of aberrant lipid metabolism that may be influenced by statins or metformin to determine their expression status in ovarian cancer tissues (S1PR1 and 3). The central **hypothesis** was that pharmaceutical interventions that regulate metabolic disturbances would have an inverse association with risk of ovarian cancer, and that ovarian cancer cells would express lipid markers heterogeneously by histotype. The **rationale** for this study was that identification of medications inversely associated with ovarian cancer risk and of tissue markers influenced by these medications and expressed in ovarian cancer cells, would provide critical insight regarding ovarian cancer etiology and inform much-needed strategies for prevention. The specific aims of this dissertation are:

1. To examine the relationship between metformin and ovarian cancer risk in the Clinical Practice Research Datalink (CPRD). Additionally, to evaluate whether the association is modified by important factors such as duration of metformin use, indication and concomitant medication use.

Hypothesis: Metformin use will lower risk of ovarian cancer for women who have higher cumulative exposure, and in ovarian cancer risk factor groups such as diabetes or polycystic ovarian syndrome (PCOS).

2. To examine the relationship between statin medications and ovarian cancer risk in the Clinical Practice Research Datalink (CPRD). Additionally, to assess whether the

association is modified by important factors such as duration of statin use, indication or concomitant medication use.

Hypothesis: Statin use will lower risk of ovarian cancer for women who have higher cumulative exposure and in ovarian cancer risk factor groups, but the effect may differ by statin pharmacokinetic classes or individual statins.

3. To quantify tissue expression of key sphingosine-1phosphate (S1P) pathway markers, using immunohistochemistry (IHC) on ovarian cancer tissue microarrays, and to evaluate heterogeneity by subtype and survival by expression status in the Polish Ovarian Cancer Study (POCS).

Hypothesis: Expression of S1P receptors S1PR1 and S1PR3 will be higher in ovarian cancer subtypes, for which obesity is a risk factor, and in late stage or higher-grade disease. Elevated expression of S1PR1 and S1PR3 will be predictive of poorer survival.

Aims 1 and 2 utilized data from the Clinical Practice Research Datalink (CPRD). The CPRD is one of the largest databases of medical records from primary care practices in the world, established in 1987. It houses information on more than 11 million residents of the United Kingdom (UK), including historical records. These individuals are a representative sample of the entire UK population. CPRD contains prospectively collected exposure data for both metformin and statin drugs with up to 18 years of follow-up data for patients [50]. Multivariable logistic regression was used to test the associations between metformin and statins in all women with ovarian cancer, and in distinct subgroups.

Aim 3 utilized data from the Polish Ovarian Cancer Study (POCS) and stains for two S1P pathway markers in epithelial ovarian cancer tissue samples [51]. Formalin fixed paraffin-embedded (FFPE) samples from 166 ovarian cancer cases in Warsaw and Lodz, Poland were

used during the years 2001-2003 [52]. S1P is a sphingolipid that has been associated with angiogenesis, cellular proliferation and cell growth [53-55] and has been proposed as a potential target of metformin [56]. This aim tested whether expression of key S1P pathway markers differed by histotype and whether survival from ovarian cancer differed by marker status.

Significance and innovation

This dissertation used one of the largest primary care databases in the world to evaluate the associations between lipid-modifying drugs and ovarian cancer risk, adding substantially to the literature on this topic with the most comprehensive evaluation on either drug to date. Results from Aims 1-2 provided valuable evidence as to whether metformin or statin drugs could be repurposed in the prevention of ovarian cancers and improved upon existing studies by including analyses by indication, concomitant medication use and accounting for birth cohort effects. A biochemical pathway that may be targetable by either metformin or statin drugs was evaluated for Aim 3 that contributes information on lipid-related pathways in ovarian cancer carcinogenesis.

Chapter 2: Background

Classification of ovarian cancer

A primer on staging and grading of ovarian cancers

Two common tumor characteristics are used to describe the spread and morphology of cancers: stage and grade. Disease stage is determined by assessment of the following: 1) the location of the tumor, 2) tumor size, and 3) lymph node involvement or whether the tumor has metastasized to another body location. Tumor staging (typically by way of the TNM staging system) is conducted by a pathologist upon diagnosis or during surgery to remove the tumor. With this system, the physician determines the T = size of the main tumor, N = number of lymph nodes with cancer and M = metastatic status [57]. A stage I ovarian cancer is considered completely contained in either the ovaries or the fallopian tube, where a stage II tumor has pelvic extension. This means that there is evidence of cancer on an additional site – either the uterus, fallopian tubes or ovaries. Stage III ovarian cancer has spread outside of the pelvis and may be found in lymph nodes. Stage IV is defined by a distant metastasis, indicating that the cancer has spread far from its origin in the fallopian tube or ovary. Stage IV is the highest stage [58].

The grade of a tumor is described by a pathologist based on the appearance of the cells that comprise the cancerous tissue. If cells from a cancerous tissue resemble differentiated cells that surround it, this may be a low-grade cancer, as opposed to cells that are undifferentiated, or do not resemble somatic cells surrounding the tumor, which would likely indicate a high-grade disease [57]. Occasionally, grade is considered as part of cancer stage, but this is not consistent in cancer staging.

Epithelial ovarian cancer histotypes and cells of origin

Epithelial ovarian cancer is comprised of four primary histotypes. Serous tumors account for approximately more than 60% of invasive tumors, while non-serous histotypes (endometrioid, mucinous and clear cell) account for 30-35%. Serous tumors may be further classified into high- and low-grade cancers, with high-grade disease as more common [10]. The remaining <10% of non-epithelial ovarian tumors are a group of rare histotypes that include mixed cell, germ cell, sex cord-stromal, and tumors without specific morphology [59]. Just over a decade ago, these ovarian histotypes were treated as subgroups of a single disease. Today, ovarian carcinogenesis is better understood, and these histotypes are considered to be, in part, separate diseases that may arise from distinct cells of origin [60]. High-grade serous and low-grade serous (approximately 10% of all serous tumors) tumors may arise from the fallopian tube [8, 61]. The cells of origin for non-serous tumors may form from inclusion cysts or implants in the ovarian surface epithelium (OSE). Cells implanted into the OSE for endometrioid and clear cell tumors may arise from endometriotic tissue in the pelvis [8, 60]. Mucinous tumors may develop from germ or transitional cells of the ovary, but there remains some ambiguity in the cells of origin and full agreement has not been reached [8].

A grouping was proposed to accommodate both molecular and genetic features of the separate histotypes (Figure 1). Non-serous cancers are usually early stage and low-grade with rare TP53 mutations and are slow-growing. High grade serous tumors present at a late stage and grow rapidly. These aggressive tumors often have TP53 mutations, which affect a cell's ability to control rapid cellular proliferation and apoptosis of cancer cells, resulting in cell survival and proliferation [8, 62].

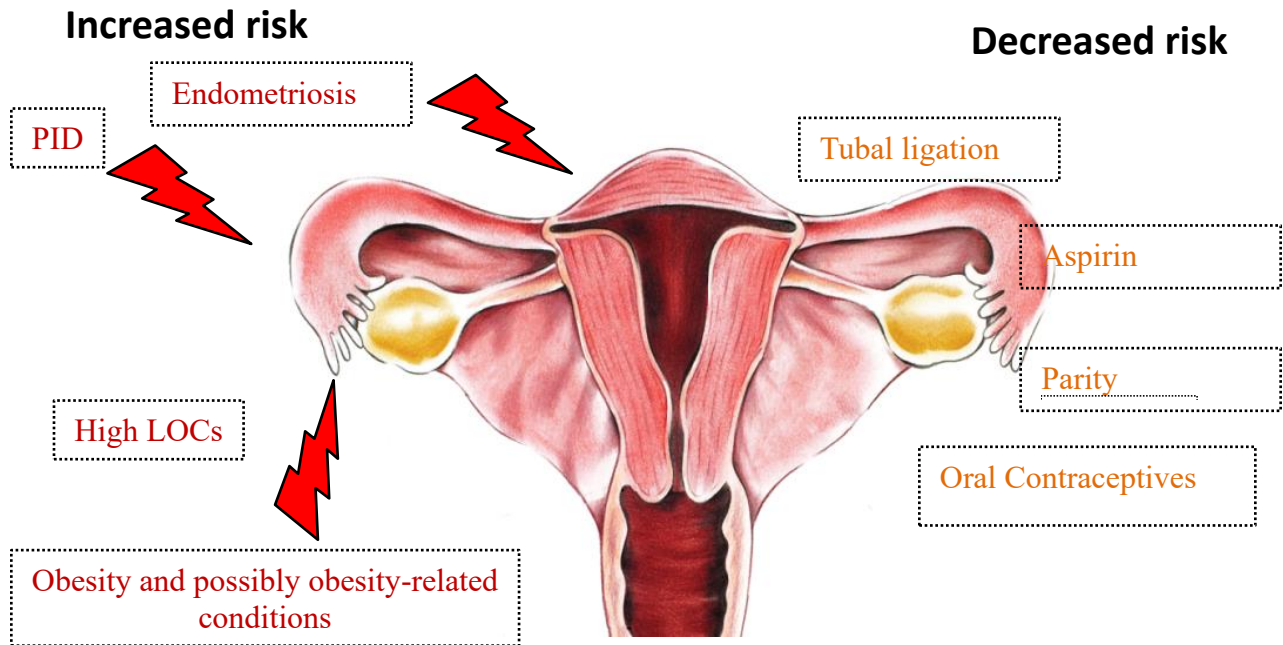


Figure 1 Ovarian cancer risk factor associations accounting for cell of origin

Epidemiology of ovarian cancer

In 2020, there were 314,000 incident cases of ovarian cancer worldwide, and 207,000 deaths [2]. Ovarian cancer incidence is highest in developed countries, and more specifically, in North America and Europe [63]. The distribution of histotypes among the total number of incident epithelial ovarian cancers varies by region, but frequencies of serous cancers tend to be highest (60%) followed by endometrioid (10%), mucinous and clear cell tumors (6-7%) [64].

Mortality and morbidity

Serous ovarian cancer has the poorest 5-year survival estimate at only 43% for any stage at diagnosis. When considering stage at the time of diagnosis, 26% of women with stage IV ovarian cancer survive five years after diagnoses versus 86% diagnosed at stage I, and the stage distributions vary considerably by cancer histotypes. Although more women survive to 5 years

after a diagnosis of endometrioid, mucinous or clear cell cancers at any stage (82%, 71% and 66%, respectively), only a small proportion of women with a stage IV diagnosis survive to 5 years (29%, 13% and 16%, respectively) compared to a stage I diagnosis (95%, 93% and 85%, respectively) [65]. Symptoms of ovarian cancer are either not present or non-specific in early stages. Some women experience unexplained weight loss, back pain, pain and bloating in the abdomen, constipation or urinary complications, but because of the lack of specificity of symptoms and lack of screening tools, many women are diagnosed with metastatic disease at the time of diagnosis [66, 67].

Risk factors for epithelial ovarian cancer

Risk factors for ovarian cancer by histotype are presented in Table 1. Reproductive risk factors have been identified for ovarian cancer, with observed differences in risk factor associations by histotype [10]. Parity is a strong protective factor and a reduction in ovarian cancer risk appears to occur after the first birth and continues with every successive pregnancy until the fourth birth [10]. Oral contraceptive use also reduces the risk of ovarian cancer in all histotypes and a stronger protective effect occurs with increasing duration of use. Some risk factors that may be associated with a reduction in inflammation, such as tubal ligation procedures, also reduce risk of ovarian cancer in all histotypes [12]. Regular low-dose use of aspirin has been demonstrated to reduce the risk of serous tumors and may also reduce risk of endometrioid and mucinous tumors [11]. Smoking may increase risk of mucinous tumors, and decrease risk of clear cell tumors but is not associated with risk of serous or endometrioid types. Inflammatory conditions in the pelvis, like endometriosis, increase the risk of ovarian cancer in all histotypes [13]. While polycystic ovarian syndrome (PCOS) and pelvic inflammatory disease (PID) are inflammatory conditions of the pelvis, the associations between PCOS and PID with

risk of ovarian cancer in any histotype has not been consistent [14]. This may be due to an underestimation of the prevalence of these conditions as they are usually ascertained via self-report, and women may not be aware that they have them due to inconsistent clinical testing. Other factors, such as a 5-year increase in age at menopause increases the risk of ovarian cancer in most histotypes (except for mucinous tumors). Age at first menstrual cycle (menarche) is not associated with ovarian cancer histotypes. First-degree family history of ovarian and/or breast cancer increases the risk of serous ovarian cancers [10].

Select lifestyle and demographic factors have also been associated with ovarian cancer. Both obesity ($\text{BMI} \geq 40 \text{ kg/m}^2$) and extremely low body mass ($\text{BMI} < 18.5 \text{ kg/m}^2$) may increase the risk of ovarian cancer compared to normal weight ($18.5\text{-}24.9 \text{ kg/m}^2 \text{ BMI}$)[18]. Evidence from the literature suggests that the association between BMI and ovarian cancer differs for serous versus non-serous cancers [15]. $\text{BMI} \geq 40 \text{ kg/m}^2$ was associated with mucinous and endometrioid tumors, while no association was observed for serous or clear cell tumors [10, 15]. Finally, an increase of 0.5 cm in adult height is associated with increased risk of ovarian cancer in all histotypes with the strongest associations observed in serous and clear cell tumors [68].

Conditions related to obesity– diabetes, metabolic syndrome and hypercholesterolemia may also impact the risk of ovarian cancer, but their effects are not consistent in the literature [19, 69]. The clinical definition of diabetes is having one or more of the following diagnostic criteria [70]:

1. Hemoglobin A1C levels $\geq 6.5\%$ (correlate of blood glucose levels within 3 months of blood draw)
2. Fasting Plasma Glucose (FPG) $\geq 126 \text{ mg/dl}$
3. A 2-hour plasma glucose $\geq 200 \text{ mg/dl}$ after a 75g oral glucose tolerance test

4. Random plasma glucose ≥ 200 mg/dl in a patient with hyperglycemia symptoms

A systematic review and meta-analysis of 19 studies evaluating the association between diabetes and risk of ovarian cancer found an increased risk of all epithelial ovarian cancers for women who had diabetes (OR 1.17, 95% CI: 1.02-1.33) with significant heterogeneity across studies ($I^2=0.04$) [19]. Effect sizes ranged from 0.59-2.42, and this variability was evident regardless of study design. All studies from this 2013 review [19] and additional studies identified in a 2017 review [71] are presented in Table 2. Furthermore, only one study [72] reported risk estimates for the association of diabetes with ovarian cancer histotypes (RR_{serous} 0.71, 95% CI: 0.41-1.23; $RR_{\text{nonserous}}$ 1.41, 95% CI: 0.70-2.85).

Metabolic syndrome (MetS) is a condition defined by the occurrence of multiple cardiovascular disease (CVD) risk factors [25]. In order to be diagnosed with MetS, an individual must have at least three of the following five conditions [73, 74]:

1. Waist circumference >102 cm in men, or >88 cm in women
2. Plasma triglycerides ≥ 150 mg/dl
3. High-density lipoprotein (HDL) cholesterol <40 mg/dl in men or <50 mg/dl in women
4. Blood pressure $\geq 130/85$ mm Hg
5. Fasting plasma glucose ≥ 110 mg/dl

To date, only two studies have examined MetS in relation to ovarian cancer with disparate results observed (Table 2). The risk of ovarian cancer for women with MetS was borderline in one study (RR 1.26, 95% CI: 1.00-1.59) [21] while a second study reported a null estimate for risk (RR 0.99, 95% CI: 0.86-1.15) [17]. The lack of studies that reported an association between metabolic syndrome and risk of ovarian cancer may partly be explained by

the lack of universal reporting of MetS in medical records, or that components of the metabolic syndrome definition (glucose and cholesterol measures, for instance) are not often available in population-based studies. To date, no studies have reported estimates for the association between MetS and ovarian cancer histotypes.

Hypercholesterolemia is defined as a low-density lipoprotein (LDL) cholesterol level above 190 mg/dL [75]. Estimates for the association between hypercholesterolemia and risk of ovarian cancer are reported in only a limited number of studies, and the estimates vary widely (RRs 0.95-1.60) [76-78] (Table 2). No studies have assessed hypercholesterolemia and risk of histotype-specific ovarian cancer.

The variability identified in the literature for the aforementioned metabolic factors highlights several important issues. Although the effect estimates vary widely by each condition, findings from the limited studies generally support an increased risk of ovarian cancer with the presence of these conditions [79-81]. The subsequent sections will provide an in-depth discussion of ovarian cancer etiologic hypotheses, including the potential role of lipid dysregulation in the development of ovarian cancer.

Table 1 Summary of studies related to reproductive and demographic risk factors of ovarian cancer

	Adjusted OR HR (95% CI) Overall	Adjusted OR HR (95% CI) Serous (all)	Adjusted OR or HR (95% CI) HG Serous	Adjusted OR or HR (95% CI) LG Serous	Adjusted OR or HR (95% CI) Endometrioid	Adjusted OR or HR (95% CI) Mucinous	Adjusted OR HR (95% CI) Clear Cell
Lifestyle and demographic risk factors							
Body Mass Index (BMI) [15]							
<18.5	1.08 (0.84-1.39)	0.93 (0.72-1.20)	1.07 (0.90-1.27)	1.33 (0.67-2.62)	1.47 (0.98-2.21)	2.48 (1.03-4.51)	2.69 (1.34-5.41)
18.5-24.9 (ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)
25-29.9	1.00 (0.92-1.09)	0.93 (0.86-1.02)	0.89 (0.81-0.99)	1.02 (0.65-1.59)	1.12 (0.96-1.30)	1.19 (0.95-1.50)	1.05 (0.79-1.40)
30-34.9	1.06 (0.97-1.16)	0.94 (0.84-1.05)	0.88 (0.76-1.02)	0.86 (0.44-1.67)	1.37 (1.14-1.64)	1.48 (0.92-2.37)	1.14 (0.79-1.63)
35-39.9	1.21 (1.07-1.38)	1.06 (0.90-1.23)	0.89 (0.73-1.09)	1.16 (0.52-2.60)	1.74 (1.36-2.23)	2.03 (1.10-3.77)	1.59 (1.14-2.24)
≥ 40	1.22 (1.05-1.41)	0.89 (0.74-1.08)			1.86 (1.42-2.24)	2.70 (1.76-4.16)	1.58 (1.04-2.40)
Height (per 0.5 cm) [10]	1.06 (1.04-1.08)	1.06 (1.03-1.09)	1.06 (1.03-1.10)	1.05 (0.93-1.18)	1.06 (1.00-1.13)	1.04 (0.95-1.13)	1.08 (0.98-1.19)
Aspirin use [11]	0.91 (0.84-0.99)	0.89 (0.80-0.99)	-	-	0.90 (0.74-1.09)	0.89 (0.58-1.38)	1.09 (0.84-1.41)
First-degree family history of breast cancer [10]	1.09 (1.00-1.19)	1.13 (1.02-1.26)	1.13 (1.00-1.29)	1.23 (0.71-2.14)	1.47 (1.15-1.87)	0.73 (0.47-1.13)	0.75 (0.46-1.22)
First-degree family history of ovarian cancer [10]	1.48 (1.26-1.75)	1.61 (1.32-1.97)	1.60 (1.26-2.03)	0.90 (0.22-3.71)	0.97 (0.52-1.82)	1.33 (0.59-3.00)	0.96 (0.36-2.57)
Smoking [82]	0.99 (0.94-1.05)	0.99 (0.92-1.06)	0.95 (0.87-1.04)	1.11 (0.76-1.63)	0.93 (0.79-1.09)	1.27 (1.01-1.59)	0.95 (0.74-1.21)
Reproductive risk factors							
Parity [10]							
0	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)
1	0.82 (0.43-0.91)	0.86 (0.75-1.00)	0.85 (0.72-1.00)	0.83 (0.41-1.65)	0.78 (0.60-1.03)	0.59 (0.38-0.92)	0.67 (0.46-0.98)
2	0.74 (0.68-0.81)	0.87 (0.78-0.97)	0.87 (0.76-1.00)	0.87 (0.51-1.50)	0.49 (0.39-0.62)	0.61 (0.44-0.86)	0.38 (0.27-0.53)
3	0.67 (0.62-0.74)	0.82 (0.73-0.92)	0.81 (0.71-0.93)	0.87 (0.51-1.49)	0.41 (0.32-0.54)	0.52 (0.36-0.74)	0.29 (0.19-0.43)
≥4	0.58 (0.53-0.64)	0.72 (0.63-0.81)	0.67 (0.58-0.78)	0.45 (0.23-0.89)	0.34 (0.25-0.45)	0.55 (0.38-0.80)	0.14 (0.08-0.25)

	Adjusted OR HR (95% CI) Overall	Adjusted OR HR (95% CI) Serous (all)	Adjusted OR or HR (95% CI) HG Serous	Adjusted OR or HR (95% CI) LG Serous	Adjusted OR or HR (95% CI) Endometrioid	Adjusted OR or HR (95% CI) Mucinous	Adjusted OR HR (95% CI) Clear Cell
Oral Contraceptive (OC) use (years) [10]							
Never	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)
≤ 1	0.98 (0.87-1.05)	0.99 (0.88-1.12)	0.99 (0.84-1.16)	1.80 (0.98-3.31)	1.01 (0.76-1.35)	0.98 (0.66-1.45)	0.68 (0.42-1.09)
>1-5	0.86 (0.78-0.92)	0.85 (0.77-0.95)	0.88 (0.77-1.01)	1.13 (0.65-1.94)	0.82 (0.64-1.05)	0.84 (0.58-1.21)	0.88 (0.62-1.24)
>5-10	0.77 (0.67-0.84)	0.72 (0.64-0.83)	0.78 (0.67-0.92)	0.94 (0.48-1.83)	0.85 (0.64-1.13)	0.91 (0.61-1.37)	0.80 (0.54-1.20)
>10	0.67 (0.58-0.75)	0.64 (0.54-0.74)	0.68 (0.56-0.83)	0.56 (0.22-1.42)	0.64 (0.44-0.93)	1.18 (0.77-1.81)	0.51 (0.29-0.87)
Tubal Ligation [12]	0.71 (0.66-0.77)	0.81 (0.74-0.89)	0.80 (0.73-0.89)	0.89 (0.65-1.22)	0.48 (0.40-0.59)	0.68 (0.52-0.89)	0.52 (0.40-0.67)
Hysterectomy [10]	0.96 (0.89 (1.03)	1.03 (0.94-1.13)	0.87 (0.53-1.43)	1.02 (0.91-1.14)	0.84 (0.66-1.07)	0.72 (0.51-1.02)	0.57 (0.36-0.88)
Age at menarche (per 1-year increase) [10]	0.99 (0.97-1.00)	0.99 (0.97-1.02)	1.01 (0.98-1.03)	1.01 (0.91-1.11)	1.00 (0.94-1.05)	1.00 (0.93-1.07)	0.92 (0.85-0.99)
Age at menopause (per 5-year increase) [10]	1.06 (1.02-1.10)	1.05 (1.01-1.10)	1.03 (0.98-1.09)	0.54 (1.23-1.92)	1.19 (1.05-1.34)	0.95 (0.81-1.11)	1.37 (1.15-1.64)
Hormone therapy use (HRT) [10]	1.36 (1.28-1.46)	1.41 (1.30-1.53)	1.48 (1.34-1.65)	1.87 (1.17-2.97)	1.67 (1.36-2.05)	1.00 (0.75-1.34)	0.90 (0.64-1.28)
Pelvic Inflammatory Disease (PID)	0.99 (0.83-1.19)	0.95 (0.82-1.11)	0.89 (0.74-1.08)	1.48 (0.92-2.38)	0.84 (0.56-1.25)	1.15 (0.83-1.60)	1.05 (0.65-1.70)
Polycystic ovarian syndrome (PCOS) [83]	0.87 (0.65-1.15)	-	0.88 (0.62-1.26)	1.07 (0.43-2.66)	1.03 (0.63-1.67)	0.61 (0.22-1.66)	0.77 (0.34-1.78)
Endometriosis [13]	1.46 (1.31-1.63)	-	1.13 (0.97-1.32)	2.11 (1.39-3.20)	2.04 (1.67-2.48)	1.02 (0.69-1.50)	3.05 (2.43-3.84)

Table 2 Summary of estimates for associations between metabolic conditions and risk of ovarian cancer

Obesity-related condition	Adjusted OR RR* (95% CI)
Diabetes	
Inoue et al., 2006 [79]	2.42 (0.96-6.09)
Khan et al., 2006 [84]	1.86 (0.43-8.05)
Rapp et al., 2006 [85]	1.33 (0.71-2.50)
Stattin et al., 2013 [86]	0.89 (0.52-1.54)
Chodick et al., 2010 [80]	2.39 (1.43-4.00)
Bjorge et al., 2013 [17]	0.59 (0.23-1.48)
Johnson et al., 2011 [87]	1.10 (0.84-1.44)
Lambe et al., 2011 [88]	1.23 (0.76-1.97)
Wotton et al., 2011 [89]	0.93 (0.40-1.88)
Lo et al., 2012 [90]	1.05 (0.93-1.20)
Gapstur et al., 2012 [72]	1.05 (0.93-1.20)
Attner et al., 2012 [91]	1.00 (0.49-2.01)
O'Mara et al., 1985 [92]	1.01 (0.57-1.77)
Adler et al., 1996 [93]	0.90 (0.50-1.40)
Parazzini et al., 1997 [94]	0.80 (0.54-1.19)
Kuriki et al., 2007 [95]	1.23 (0.53-2.84)
Reis et al., 2010 [96]	2.13 (1.40-3.23)
Bosetti et al., 2012 [97]	1.21 (0.80-1.82)
Harding et al., 2015* [81]	1.23 (1.15-1.32)
Hemminki et al., 2010* [98]	1.84 (1.59-2.12)
Weiderpass et al., 2002* [99]	0.97 (0.87-1.08)
Harding et al., 2015† [81]	1.38 (1.01-1.89)
Hsu et al., 2015† [100]	3.89 (1.72-8.81)
Shu et al., 2010† [101]	1.32 (0.64-2.71)
Hypercholesterolemia	
Melvin et al., 2012 [76]	0.95 (0.53-1.71)
Hu et al., 2011 [77]	1.33 (0.85-2.07)
Kabat et al., 2018 [78]	1.60 (0.89-2.86)
Metabolic Syndrome	
Esposito et al., 2012 [21]	1.26 (1.00-1.59)
Bjorge et al., 2011 [17]	0.99 (0.86-1.15)

* Type 2 Diabetes only

† Type 1 Diabetes only

Ovarian cancer etiologic hypotheses

Multiple theories of ovarian cancer tumorigenesis have been proposed [102]. Three of the most prominent theories involve the effect of sex steroid hormones on the ovary (gonadotropin hypothesis), the role of inflammation, and the lifetime exposure to ovulatory cycles (incessant ovulation hypothesis) [103]. Each of these theories is supported by evidence from risk factor associations but no theory completely explains ovarian carcinogenesis.

Gonadotropin signaling and incessant ovulation

The two original hypotheses of ovarian cancer carcinogenesis, pituitary gonadotropin signaling and incessant ovulation, relate to associations of reproductive and lifestyle factors with ovarian cancer (detailed above in Table 1) [103, 104]. As noted above, parity and oral contraceptive use decrease the risk of ovarian cancer for all histotypes. During pregnancy, gonadotropins from the pituitary gland -- luteinizing hormone (LH) and follicle-stimulating hormone (FSH) are lowered by dramatic increases of estrogen and progesterone steroid hormones, and ovulation is suppressed. Oral contraceptives similarly stabilize LH and FSH and suppress ovulation [104]. Because of the commonalities identified with these risk factors, early investigators hypothesized that the process of ovulation and the gonadotropin signaling that is necessary for ovulation must signal protective changes in the ovarian epithelium [105]. In recent studies of women who receive fertility treatments that are known to increase levels of LH and FSH, results have been mainly null for ovarian cancer risk [106-108]. Other studies have reported disproportionate reductions in risk for women who are pregnant and breastfeeding versus age at menarche. Pregnancy reduces a woman's risk of ovarian cancer far beyond the one-

year less ovulatory time for a woman who begins ovulating at 12 years vs 13 years old [109, 110].

Inflammation and lipid dysregulation

A seminal paper by Ness et al. published in 1999 [102] proposed that inflammation may be a strong contributing factor for ovarian cancer carcinogenesis given the increased risk observed for women who have endometriosis and pelvic inflammatory disease. Recent evidence from the Ovarian Cancer Association Consortium (OCAC) has also shown that low-dose long-term use of aspirin, which has established anti-inflammatory action, is associated with a decreased risk of serous ovarian cancer [11]. Other recent evidence from OCAC has replicated the increased risk of ovarian cancer for women who have a history of endometriosis and pelvic inflammatory disease [13, 14] and decreased risk after tubal ligation procedures [12].

A more recent theory of carcinogenesis posits that the dysregulation of systemic lipids may promote ovarian carcinogenesis [111, 112]. Many cancers have been shown to exhibit extensive alterations in multiple lipid types [113, 114], including ovarian cancer [112, 115]. Cancer cell proliferation requires dramatic changes in lipid metabolism in order to form cellular membranes, perform increased cellular signaling and to adapt to an environment of hypoxia (decreased oxygen) as lipid molecules form the cellular membrane and are active messengers in signal transduction within and between cells [29, 116]. Risk factor associations for conditions known to be associated with aberrant systemic lipid profiles (obesity, diabetes, hypercholesterolemia and MetS, outlined in previous sections) suggest that an underlying lipid imbalance may initiate tumorigenesis, but the mechanism(s) by which lipid abnormalities may exhibit these effects on epithelial ovarian cancer have not been elucidated. Lipid dysregulation may have strong connections to inflammation [29]. The connection between lipid abnormalities

and inflammation appears to be bidirectional – the cellular changes in metabolism spark immunologic alterations, and vice versa [117, 118].

Description of the exposures: metformin

Metformin drug characteristics

Metformin is a biguanide anti-hyperglycemic medication that is typically used to treat pre-diabetes or manifest Type 2 diabetes [119]. The mechanism(s) of action of metformin is not well understood, but it is known to primarily suppress glucose production in the liver and potentially increases glucose uptake and glycogenesis in skeletal muscle, leading to reduced blood glucose levels [120]. Metformin reduces cellular adenosine triphosphate (ATP), a key source of cellular energy, by inhibiting mitochondrial complex I and increasing cellular adenosine monophosphate (AMP). The increase in cellular AMP activates pathways that lead to improved insulin sensitivity, and inhibition of gluconeogenesis by blocking an enzyme (adenylate cyclase) in hepatocellular membranes that leads to increase in glycolysis (glucose breakdown) and inhibition of gluconeogenesis (glucose production)[120].

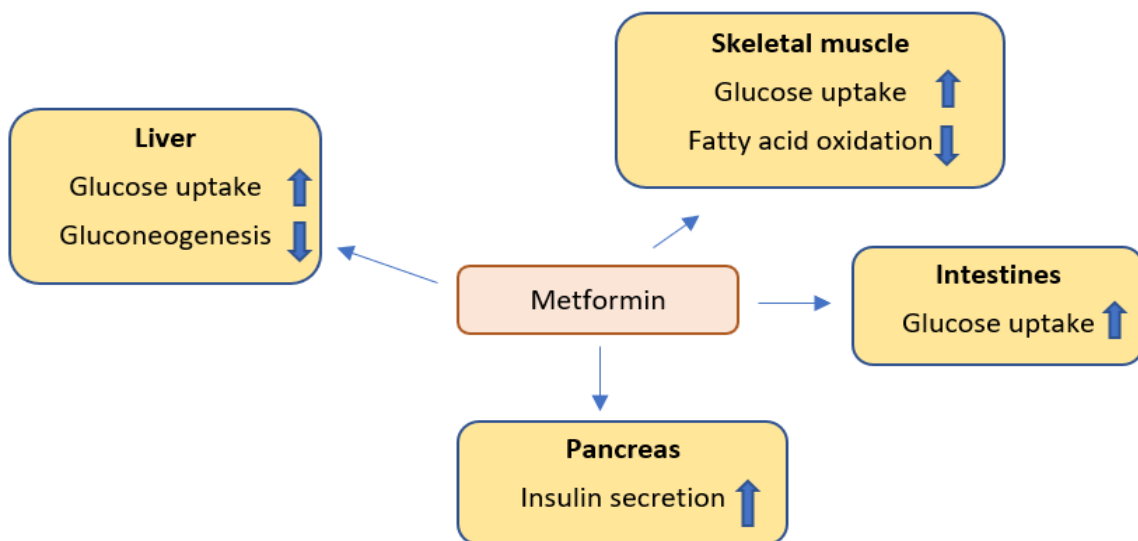


Figure 2. Mechanisms by which metformin regulates glucose

Metformin was introduced in Europe and the UK in 1958 to treat diabetes and pre-diabetes, and was approved by the FDA for use in the United States in 1995 [121]. The use of metformin among adults in the United States rose from 51.6 million in 2008 to 61.6 million in 2012 and it was the 8th most consistently prescribed drug during that period [122]. Between the years 2000 and 2012, the prevalence of metformin use increased from 55% to >80% in the UK for patients with Type 2 diabetes, and this trend was similar in men and women [123]. Currently, it is estimated that 84% of individuals with Type 2 diabetes are prescribed metformin [123].

Metformin is considered one of the World Health Organization's (WHO) essential medicines (EML) [124]. The WHO selects medications that are determined to be health care priorities based on the WHO's review of the literature pertaining to "public health relevance, evidence of efficacy and safety, and comparative cost-effectiveness" of the drugs. Essential

medicines must be available “within the context of functioning health systems at all times, in adequate amounts, in the appropriate dosage forms, with assured quality and adequate information, and at a price the individual and community can afford” [125].

Additional indications for metformin

In addition to the primary indication of metformin for the treatment of Type 2 diabetes and pre-diabetes, women with polycystic ovarian syndrome (PCOS) may also be prescribed metformin. Insulin resistance occurs in women with PCOS, and it is hypothesized that increased circulating levels of insulin may influence increased production of insulin-like growth factor (IGF-1). Insulin and IGF1 may stimulate ovaries to produce more testosterone and may halt follicle growth mid-cycle leading to ovarian cysts that would normally rupture at ovulation but persist [126]. Women with PCOS may be anovulatory (having menstrual cycles without ovulation) and may experience symptoms of excess androgen hormones [127]. Metformin has proven to be an effective treatment for inducing ovulation by increasing insulin sensitivity [128].

Recommendations for metformin prescriptions

Metformin is the only biguanide drug marketed in the US and UK and includes the brand names, Glucophage, and Glucophage XR (extended release)[129]. Glucophage is available in 500, 800 or 1000 mg of the compound metformin hydrochloride. Glucophage XR is available in either 500 or 750mg tablets. Dosage of metformin at 2,000 mg per day was demonstrated to be most efficacious in lowering blood glucose levels, but may be prescribed as high as 3,000 mg/day depending on the patient [130].

Existing knowledge

To date, seven studies reported risk estimates for the association between metformin use and risk of non-histotype specific ovarian cancer (Table 3). Five of these reported a risk estimate on the order of 0.60 or below, but only one of these studies conducted in the National Health Insurance Database in Taiwan showed a significantly reduced risk of ovarian cancer [35, 38-40]. Importantly, one of these studies [40], reporting a reduced risk of ovarian cancer (OR: 0.61, 95% CI: 0.3-1.25) was conducted in the General Practice Research Database (GPRD), a data source from which the current CPRD was developed. Two studies reported null findings, one conducted in the Finland Diabetes Database (FinDM) and another in the Womens' Health Initiative (WHI) [36, 37]. Findings from this dissertation analysis builds upon this prior CPRD study by contributing an additional 10 years of follow-up data, including indications for each drug, thoroughly investigating potential dose-response relationships using different methods for dose ascertainment (DDDs and number of prescriptions), and accounting for potential birth cohort effects.

Table 3 Studies reporting effect measures for the association between metformin use and risk of ovarian cancer

Authors	Title	Time period	Data Source	Follow-Up	Case number	Control number or baseline cohort size	Association measure HR or OR
Randomized Controlled Trials (RCTs)							
Home et al., 2010 (ADOPT) [35]	Experience of malignancies with oral glucose-lowering drugs in the randomized controlled ADOPT (A Diabetes Outcome Progression Trial) and RECORD (Rosiglitazone Evaluated for Cardiovascular Outcomes and Regulation of Glycaemia in Diabetes) clinical trials	2000-2002	NA	4.4y	0	1,454	HR 0.28 (0.01-5.51)
Home et al., 2010 (RECORD) [35]				5.5y	1	1,122	HR 0.66 (0.16-2.14)
Cohort Studies							
Urpilanian et al., 2018 [36]	The role of metformin and statins in the incidence of epithelial ovarian cancer in type 2 diabetes: a cohort and nested case-control study Diabetes, metformin and incidence of and death from invasive cancer in postmenopausal women: results from the Women's Health Initiative (WHI) Type 2 diabetes mellitus and medications for type 2 diabetes mellitus are associated with risk for, and mortality from cancer in a German primary care cohort	1996-2011	Finland Diabetes Database (FinDM)	>40 years	303	137,643	HR 1.02 (0.72-1.45)
Gong et al., 2016 [37]		2005-2015	Women's Health Initiative (WHI)	15y	553	145,826	HR 1.06 (0.64-1.74)
Baur et al., 2011 [38]		2003-2008	Targets and Essential Data for Commitment of Treatment (DETECT)	5y	251	6,826	HR 0.69 (0.43-1.10)

Authors	Title	Time period	Data Source	Follow-Up	Case number	Control number or baseline cohort size	Association measure HR or OR
Tseng et al., 2015 [39]	metformin reduces ovarian cancer risk in Taiwanese women with type 2 diabetes mellitus	1998-2002	National Health Insurance of Taiwan	7y	3,201	479,475	HR 0.66 (0.59-0.73)
Case-Control Studies							
Bodmer et al., 2011 [40]	Use of metformin and the risk of ovarian cancer: a case-control analysis	1995-2009	General Practice Research Database (GPRD)	NA	1,611	10,781	OR 0.61 (0.30-1.25)

Abbreviations: OR = odds ratio, HR = hazard ratio

Description of the exposures: statins

Statins are a group of drugs designed to block the primary rate-limiting step in cholesterol synthesis and are commonly prescribed as a primary prevention for cardiovascular disease (CVD), which includes outcomes such as stroke and hypertension, and often coexists with diabetes. As a class, statins are collectively known as 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase inhibitors or HMG-CoA reductase inhibitors [131, 132]. These compounds inhibit the enzyme HMG-CoA reductase which impedes formation of L-mevalonate in the liver, a key component step in the formation of cholesterol and the production of low-density lipoproteins (LDL) (Figure 3).

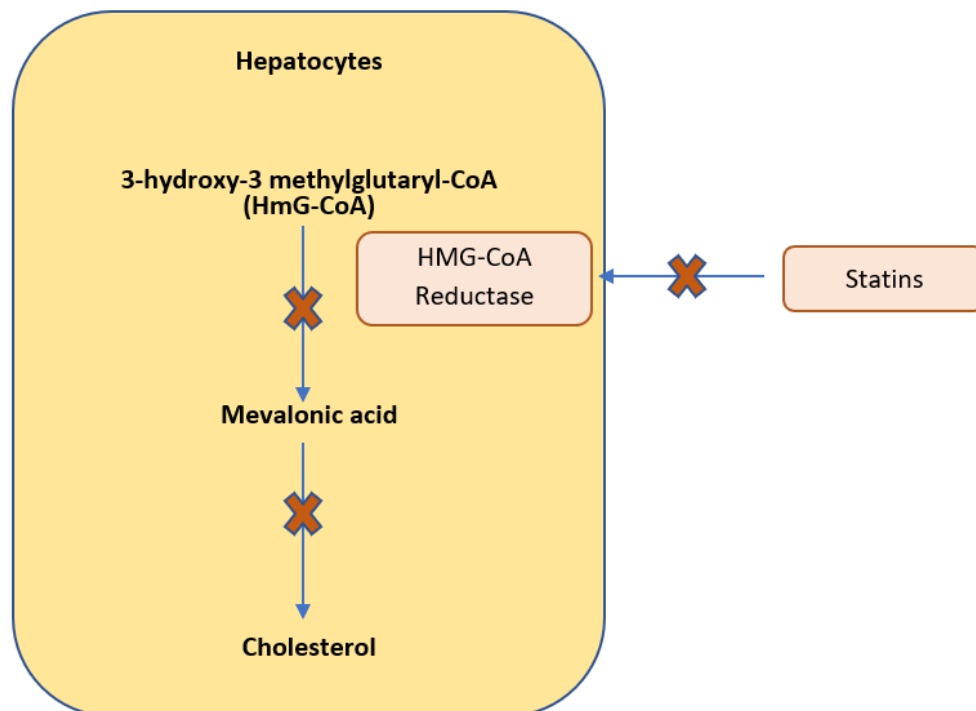


Figure 3 Statin mechanism of action

There are currently seven FDA-approved statins in the US (atorvastatin/Lipitor, fluvastatin/Lescol, lovastatin/Mevacor, pravastatin/Pravachol, rosuvastatin/Crestor, simvastatin/Zocor, pitavastatin/Livalo, and five are available in the UK (excluding lovastatin and pitavastatin). The use of statins in the US has steadily increased since their introduction in 1987. The Centers for Disease Control and Prevention (CDC)'s National Health and Nutrition Examination Survey (NHANES) reported that the overall prevalence of statin use among US adults over age 40 increased from 16% to over 26% between the years of 2003-2014 [133], and was about 23-24% in the two latest NHANES cycles covering the years 2015-2018 (Figure 4. In the UK, prescriptions for statins during the period 1995-2013 were evaluated in The Health Improvement Network (THIN) database. These UK data showed that overall statin prescriptions rose sharply from 1995-2006, from 0.51/1,000 person-years in 1995 to 19.83/1,000 person-years in 2006. Prescriptions for new statins decreased to 10.76/1,000 person-years by the year 2013. Trends in statin initiation during this period were comparable for men and women [134]. Prevalence of statin use also varies by type in both the US and the UK, with simvastatin and atorvastatin as the most commonly prescribed. In the US, the NHANES survey data from 1999-2016 demonstrates the high prescription rates for atorvastatin and simvastatin, while pravastatin, rosuvastatin, and lovastatin are prescribed less frequently, and fluvastatin and cerivastatin are prescribed infrequently [133] (Figure 5). In the UK, availability of data regarding the prevalence of individual statin types over time is limited, but one study that examined the distribution of statin use by type in 2005 using the General Practice Research Database (GPRD) suggests that the most frequent prescriptions for atorvastatin and simvastatin mimic what is observed in the US, collectively accounting for 90% of statin prescriptions. The remaining three statins

(fluvastatin, pravastatin and rosuvastatin) contribute to the remaining 10% of statin prescription [48, 135].

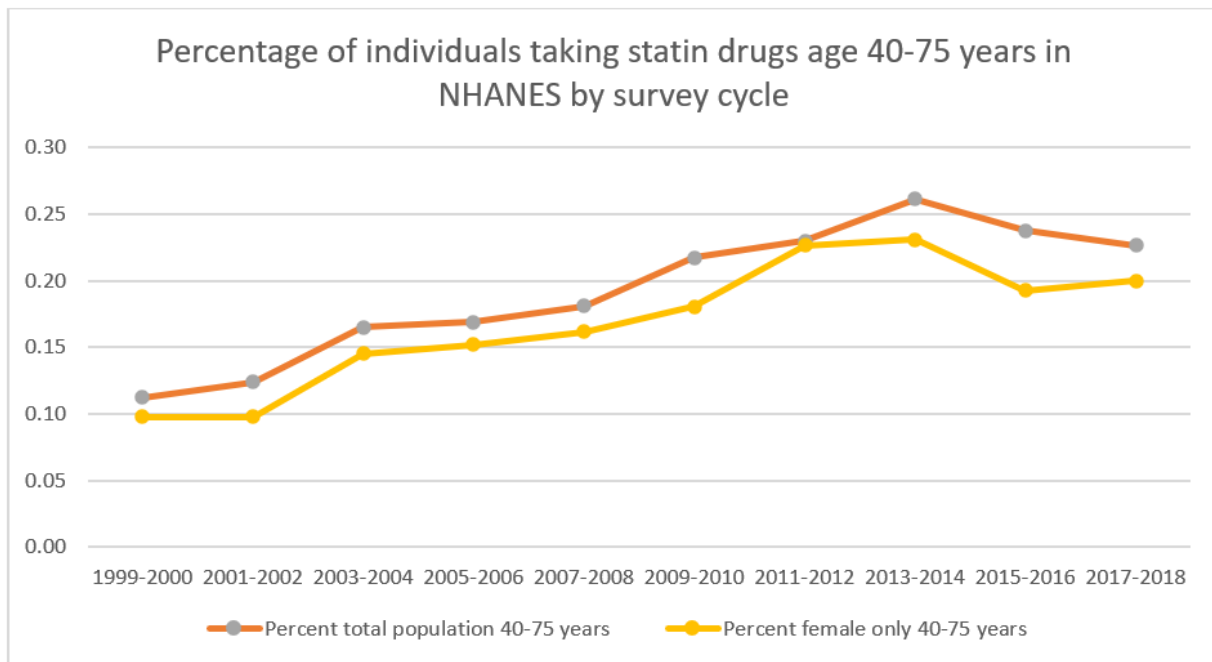


Figure 4 Prevalence of statin use in the general population among women in the United States in NHANES 1999-2018

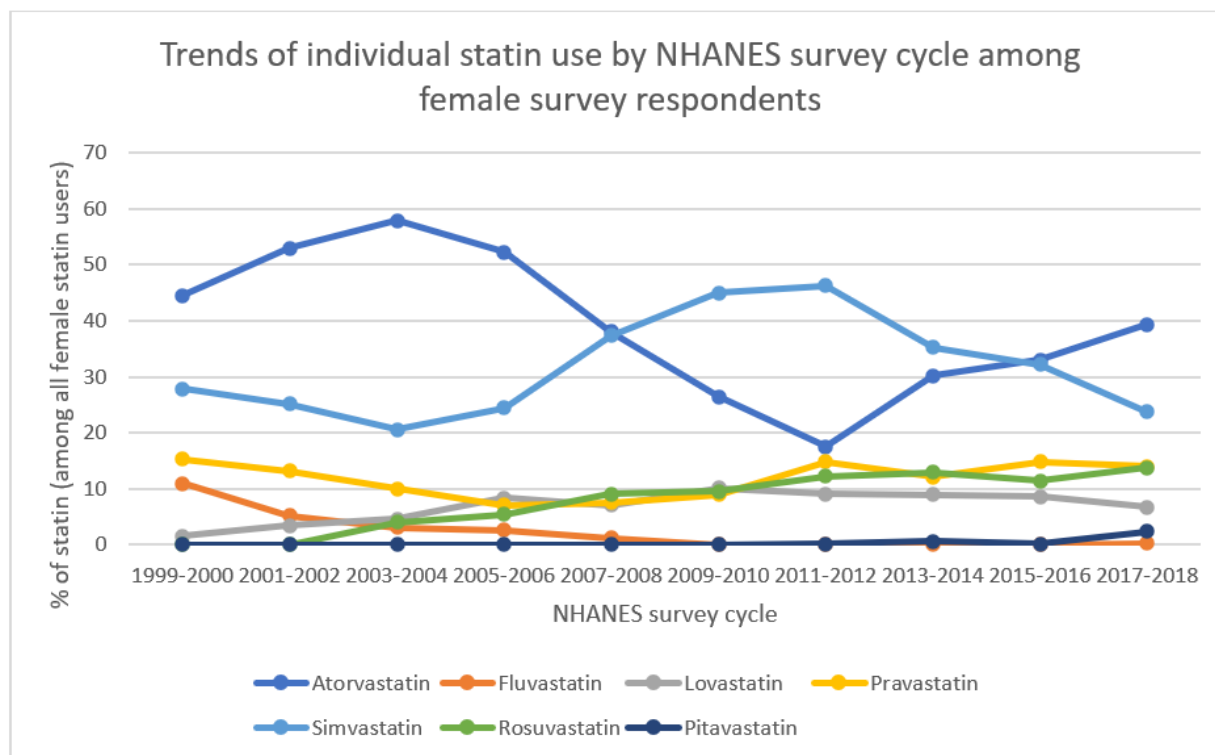


Figure 5. Prevalence of statin use by type in women from the United States in NHANES 1999-2018.

Statin formulations

All statin drugs currently available for use in both the US and the UK can be classified based on their origin (fungal or synthetic) and solubility (hydro- or lipo-philicity). Type 1 statins include those that are derivatives of fungi -- lovastatin, simvastatin and pravastatin – and Type 2 statins are synthetic – atorvastatin, rosuvastatin, fluvastatin and pitavastatin [136]. Statins act to inhibit HMG-CoA reductase by directly binding to the enzyme and prevent the binding of the substrate HMG-CoA and so all statins are considered “competitive inhibitors” of the enzyme. Four statins – lovastatin, simvastatin, atorvastatin and fluvastatin are lipophilic, and are considered to be more widely available to other tissues outside the liver than the hydrophilic statins – pravastatin and rosuvastatin. The lipophilic nature of the former set of compounds

makes possible the transport out of liver cells and into other tissues. Pitavastatin is considered lipophilic but does not exhibit the high bioavailability of the other lipophilic statins [132, 136].

Statins are also classified by their intensity, which is a measure of the percent cholesterol reduction and includes dose as part of its definition. High-intensity statins reduce cholesterol levels between 42-55% and have dose ranges of 40-80 mg and 20-40 mg/day respectively. Moderate-intensity statins reduce cholesterol levels by 32-38% while low-intensity statins reduce cholesterol levels by 21-29% [137, 138]. Typically, atorvastatin, rosuvastatin and simvastatin are prescribed as a high-intensity statin but can be prescribed at lower doses that vary by country. In these instances, atorvastatin, rosuvastatin and simvastatin would be considered moderate-intensity statins. In order to remove some of the difficulties in comparing statin use across countries and populations, a measure known as “defined daily doses” or DDD of statin use can be calculated based on the drug and its dose level for an individual. A further description of this measure and its calculation is described in the methods sections of manuscripts 1 and 2, as it will be used to calculate and define the intensity of statin use for the subjects in the CPRD.

Recommendations for statin prescriptions

Statin drugs are currently used for both the primary and secondary prevention of cardiovascular disease (CVD). The United States Preventative Services Task Force (USPSTF) has issued a statement for the use of statin drugs in the primary prevention of cardiovascular disease. For adults aged 40-75 without a history of cardiovascular disease, a low-or moderate dose statin is recommended if CVD risk factors are present, but for adults over 75 years, little data is available to determine the risks vs benefits of statin use [139]. In the UK, the National Institute for Health Care Excellence (NICE) recommends the initiation of atorvastatin at a dose

of 20 mg daily for primary prevention of CVD in individuals who are at a 10% or greater 10-year risk of CVD. For secondary prevention in the US, a high-intensity statin is typically prescribed with a goal of lowering cholesterol by at least 50%, usually atorvastatin or simvastatin [140]. In the UK, secondary prevention of CVD with high-intensity atorvastatin is recommended [138].

Existing knowledge

A systematic review of nine studies in peer-reviewed literature (Table 4) supports that statins may reduce the risk of ovarian cancer, but the association was dependent on the specific statin drug and formulation [49]. Ever-use of statins was found to be associated with a decreased risk of ovarian cancer in seven [36, 41, 43-47] of the studies, but only significantly decreased in one of these [47]. The association between statin use and ovarian cancer risk was null in the remaining two studies [42, 48]. The summary estimate from the meta-analysis suggested a reduction but was not statistically significant (RR 0.87, 95% CI: 0.74-1.03) [49]. The risk of ovarian cancer after exposure to lipophilic statins (simvastatin, atorvastatin, lovastatin, fluvastatin or cerivastatin) was non-significantly reduced (RR 0.88, 95% CI: 0.69-1.12) while risk was higher after hydrophilic statin (pravastatin, rosuvastatin) use (RR 1.06, 95% CI: 0.72-1.57) (Figure 7). After excluding the Women's Health Initiative (WHI) where one statin (pravastatin) was found to increase risk (HR 1.89, 95% CI: 1.24-2.88) and >20% of the population was taking pravastatin a meta-analysis showed that the risk of ovarian cancer was significantly reduced (OR 0.83, 95% CI: 0.70-0.99) (Figure 6) [42]. Although only three studies [42, 45, 47] were identified that reported histotype-specific effects, risks of mucinous and endometrioid tumors were reduced for ever-users of statins (Table 8). Serous and clear cell tumors showed inconsistent effects.

Finally, only one study has evaluated statin use by indication, to determine whether indication for prescription modifies the effect of statin use with risk of ovarian cancer [47]. This study found that women without conditions like coronary artery disease, hypertension, diabetes and obesity had reduced risks of ovarian cancer compared to women who had these conditions.

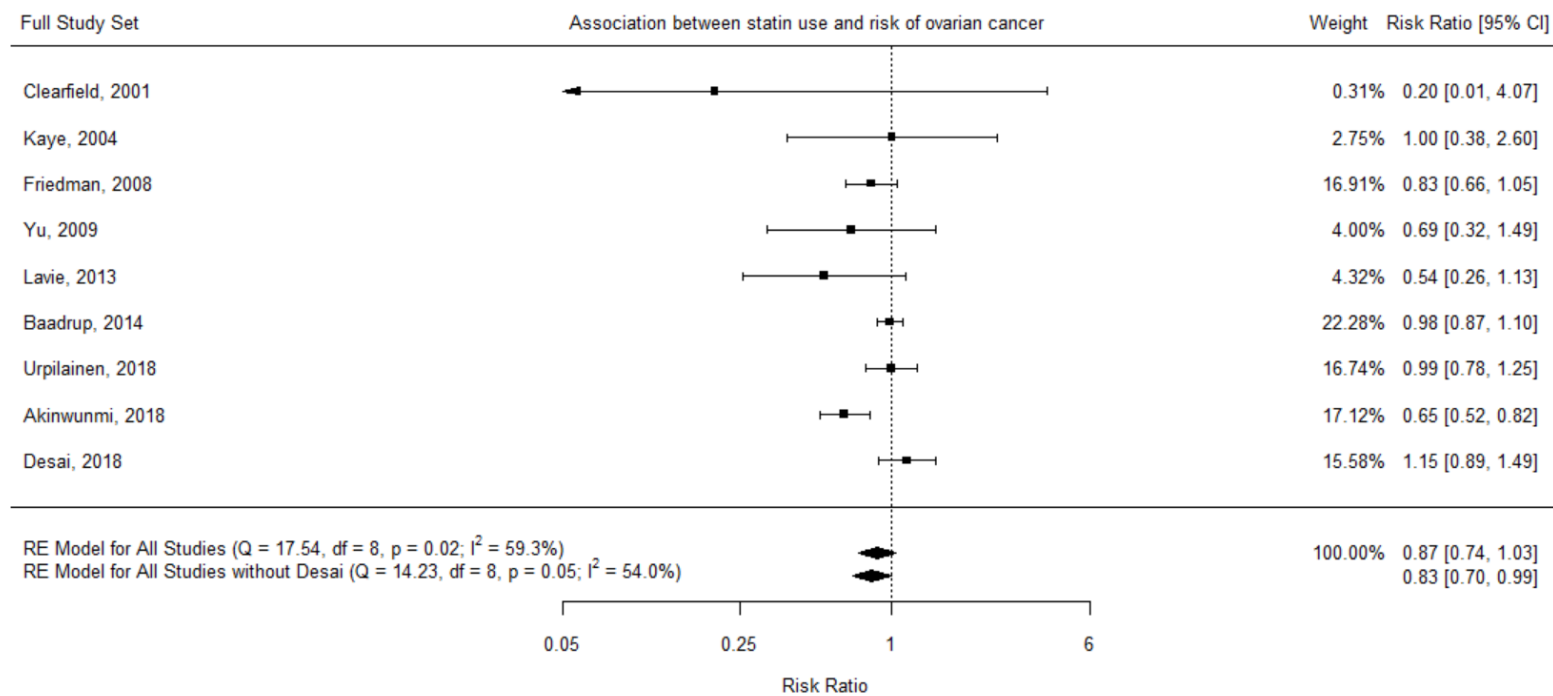


Figure 6. Association between ever-use of statins and ovarian cancer risk

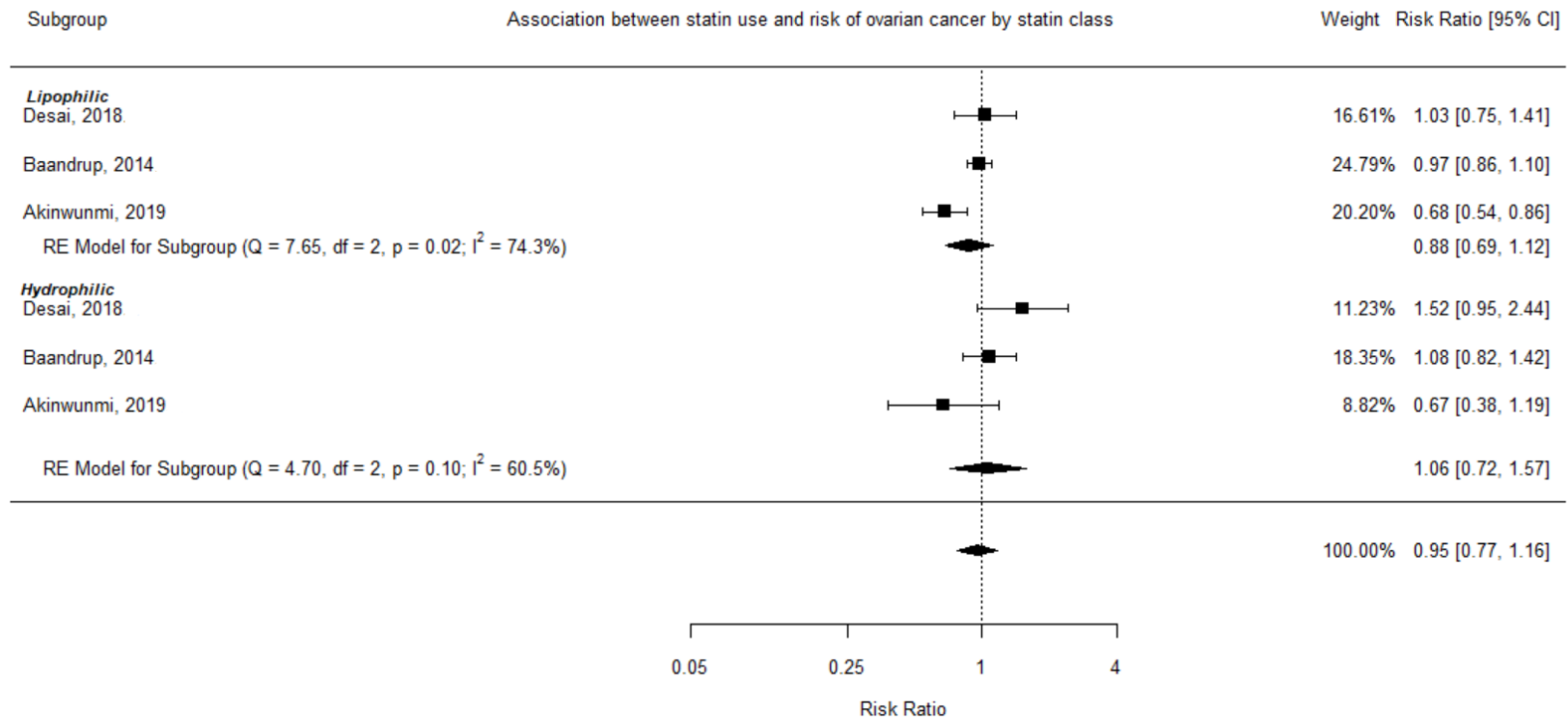


Figure 7. Association between statin use and risk of ovarian cancer by statin class

It has been proposed that pravastatin use may be associated with cancer risk because it is unable to block the mevalonate pathway in tissues outside the liver [141]. Up-regulation of the mevalonate pathway has been demonstrated in breast, liver, pancreatic and prostate cancers, as well as leukemia and lymphoma [142], and in some studies of epithelial ovarian cancer cells [143]. Multiple theories exist regarding the role of the mevalonate pathway in tumorigenesis, including increased cell proliferation due to increased availability of cholesterol and regulation of cellular energy [144]. Hydrophilic pravastatin is less likely to move through cellular membranes in extrahepatic tissues due to a lack of a sodium-independent bile-acid transporter that is necessary for its transport across cell membranes [145]. Because of this, pravastatin is only able to block the mevalonate cascade in the liver itself, and a subsequent up-regulation in mevalonate synthesis may result in extra-hepatic tissues to accommodate the decrease in cholesterol synthesis by the liver [141]. The ovary is susceptible to this potential up-regulation of this pathway, as it has been shown to express mevalonate [146]. Lipophilic statins, however, move more freely and can block the mevalonate pathway in tissues outside of the liver while simultaneously inhibiting this pathway intra-hepatically [147]

Table 4 Studies assessing the association between statin use and ovarian cancer risk

Authors	Title	Time Period	Data Source	N Cases	N Controls or Baseline Cohort	Association Measure
Randomized Controlled Trials (RCTs)						
Clearfield et al. 2001[41]	Air Force/Texas Coronary Atherosclerosis Prevention Study (AFCAPS/TexCAPS): efficacy and tolerability of long-term treatment with lovastatin in women.	NR	The Air Force/Texas Coronary Atherosclerosis Prevention Study (AFCAPS)	2	997	Ever use: RR 0.20 (0.01-4.15)
Cohort Studies						
Desai et al. 2018[42]	An analysis of the association between statin use and risk of endometrial and ovarian cancers in the Women's Health Initiative.	1993-2012	Women's Health Initiative (WHI)	763	126,253	Time-Dependent use: HR 1.30 (1.04-1.62) Baseline use: HR 1.15 (0.89-1.50)
Yu et al. 2009[43]	Statin use and female reproductive organ cancer risk in a large population-based setting.	1990-2004	Group Health of Washington State	326	93,619	Ever use: HR 0.69 (0.32-1.49)
Friedman et al. 2008[44]	Screening statins for possible carcinogenic risk: up to 9 years of follow-up of 361,859 recipients.	1994-2003	Kaiser Permanente Medical Care Program norther California (CPMCP)	79	NR	Time-dependent use: RR 0.83 (0.66-1.05)
Urpilainen et al. 2018[36]	The role of metformin and statins in the incidence of epithelial ovarian cancer in type 2 diabetes: a cohort and nested case-control study.	1996-2011	Finland diabetes database (FinDM)	303	137,643	Ever use: HR 0.99 (0.78-1.25)

Authors	Title	Time Period	Data Source	N Cases		N Controls or Baseline Cohort	Association Measure
Case-Control Studies							
Baandrup et al. 2015[45]	Statin use and risk for ovarian cancer: a Danish nationwide case-control study.	2000-2011	Danish Cancer Registry	4103	58,706	Ever use: All epithelial: OR 0.98 (0.87-1.10) Serous OR 1.03 (0.90-1.19) Endometrioid OR 0.80 (0.58-1.10) Mucinous OR 0.63 (0.39-1.00) Clear: OR 1.48 (0.92-2.38)	
Lavie et al. 2013[46]	The effect of statins on risk and survival of gynecological malignancies.	2003-2010	Hiafa case-control study; CITOUS	126	126	>1-year use: OR 0.56 (0.33-0.94)	
Akinwunmi et al. 2019[47]	Statin therapy and association with ovarian cancer risk in the New England Case Control (NEC) study.	3 time periods 1992-1998 1998-2002 2003-2008	New England Case control study	2040	4140	Ever use: All epithelial: OR 0.65 (0.52-0.82) Serous: OR 0.69 (0.53-0.90) Endometrioid: OR 0.70 (0.45-1.09) Mucinous: OR 0.39 (0.18-0.82) Clear cell: OR 0.67 (0.32-1.38)	
Kaye and Jick 2004[48]	Statin use and cancer risk in the General Practice Research Database	1990-2002	General Practice Research Database (GPRD)	91	1,066	Current use: RR 1.0 (0.4-2.70)	
Reviews and Meta-Analyses Summarizing Studies Assessing the Association of Statin use with Risk of Ovarian Cancer							
Liu et al. 2014[34]	Effect of statin on risk of gynecologic cancers: a meta-analysis of observational studies and randomized controlled trials.	multiple	multiple	624	-	Ever use: RR 0.79 (0.64-0.98) >5-year use: RR 0.48 (0.28-0.80)	
Boudreau et al. 2010[148]	Statin use and cancer risk: a comprehensive review.	multiple	multiple	411	-	Measures of association presented for Kaye, Friedman and Yu separately; estimates match study-specific estimates in this table.	

Abbreviations: HR = hazard ratio, RR = risk ratio, OR = odds ratio

Table 5. Ovarian cancer histotype-specific estimates by study

	Akinwunmi [47] OR (95%CI)	Baandrup [45] OR (95%CI)	Desai [42] OR (95%CI)
<i>Subtype (ever vs. never use)</i>			
Serous	0.69 (0.53-0.90)	1.03 (0.90-1.19)	1.22 (0.87-1.70)
Endometrioid	0.70 (0.45-1.09)	0.80 (0.58-1.10)	-
Mucinous*	0.39 (0.18-0.82)	0.63 (0.39-1.00)	-
Clear cell	0.67 (0.32-1.38)	1.48 (0.92-2.38)	1.39 (0.70-2.79)

Abbreviations: OR = odds ratio

* Estimates for mucinous tumors from Akinwunmi paper include borderline mucinous

Description of lipid pathway exposures: S1P pathway markers

The third aim of this dissertation investigated a lipid pathway that may be involved in ovarian carcinogenesis, and may also be targetable by either metformin or statins. The sphingosine-1-phosphate (S1P) bioactive lipid pathway facilitates cellular signal transduction, growth, and proliferation and is thought to influence the pathogenesis and propagation of multiple cancer cell types, including ovarian. The S1P pathway is under investigation as a potential target for cancer prevention agents in breast and colon cancers but has not been well-characterized in ovarian cancer. The third aim of this dissertation comprehensively interrogated the S1P biochemical pathway in ovarian cancer tissue in order to determine the implications of pathway activity in ovarian cancer histotypes and on histotype-specific survival.

Cancer cells exhibit alterations in lipid metabolism that may promote tumorigenesis by improving proliferation capacity and adaptability to changing intra- and intercellular environments. Bioactive lipid involvement in ovarian cancer pathogenesis is suspected based on evidence that lipid-dysregulating conditions like obesity, diabetes, and metabolic syndrome (MetS) [15, 29], and lipid-stabilizing medications like statins [49] may be associated with an increased risk of ovarian cancer. S1P is a bioactive lipid mediator that enhances cellular proliferation, signaling and angiogenesis. Involvement of S1P signaling is suggested in obesity and related conditions, and as a promotor of multiple cancer types, including breast, colon, pancreatic, gastric and ovarian cancers [149]. Key components of the S1P molecular pathway include kinases that catalyze the formation of S1P (SPHK1/SPHK2), and five G-coupled protein receptors for S1P (S1PR1-5) whose activation produce cancer-promoting pleiotropic effects [150] (Figure 8). Activated S1P receptors stimulate genetic expression of various pro-tumorigenic pathways, such as PI3K/AKT, PI3K/Rac, MAPK, Ras/ERK, PLC and others [150].

In ovarian cancer, SPHK1 and S1P receptor expression has been demonstrated for S1P kinases and receptors [151], and higher expression was associated with tumor progression [56].

Mounting evidence of S1P pathway involvement in cancer has prompted expansive research into SPHK/S1PR inhibitor development, and the repurposing potential of existing medications for use in cancer prevention, like metformin [152]. This third dissertation aim compared expression levels of critical S1P pathway markers (S1PR1,3) across ovarian cancer histotypes and determined whether expression status varied by ovarian cancer risk factors, by histotype or survival by marker status. A description of each marker, its specific function and location are captured in Table 6.

Existing knowledge

To date, few studies [56, 151, 153] indicate that the S1P pathway activity may play a role in ovarian cancer. Specific ceramides, which are precursors to S1P, as well as S1P were elevated in the serum of women with ovarian cancer compared to healthy women [153]. A recent metabolomics evaluation conducted in the Nurses' Health Study (NHS) [154] assessed circulating levels of precursors to S1P, showing that total concentrations as well as specific variations of these molecules may be associated with ovarian cancer risk overall. Differences in odds ratios by histotype from this evaluation were not significant for either lipid group, but odds ratios were suggestive of potential differences between serous and non-serous tumors [154]. Elevated tissue expression of SPHK1 and S1P receptors has been found in ovarian cancers compared to normal ovary tissue, and higher expression of S1P was associated with angiogenesis and tumor progression [151]. Chapter 6 of this dissertation provides additional description of S1P and its associated pathway, including precursor molecules and expectations for pathway expression components.

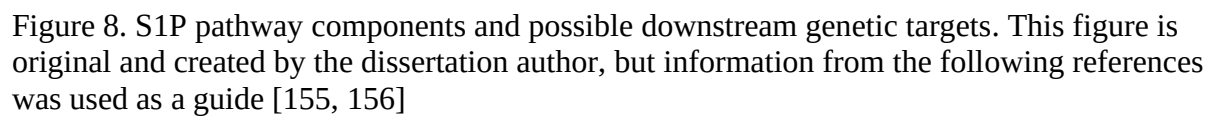


Table 6. Key S1P pathway markers

Marker	Function	Location	Gene
S1P kinases			
SPHK1	Phosphorylates sphingosine to sphingosine-1-phosphate (S1P)	Cytosol; activated SPHK1 is present on the cellular membrane	SPHK1
SPHK2	Phosphorylates sphingosine to sphingosine-1-phosphate (S1P)	Nucleus	SPHK2
S1P receptors			
S1PR1	Activates numerous pro-tumorigenic signaling pathways such as Ras/ERK and PI3K/AKT among others	Cellular membrane	S1PR1
S1PR2	Activates numerous pro-tumorigenic signaling pathways such as Ras/MAPK; may activate anti-tumorigenic JNK pathway	Cellular membrane	S1PR2
S1PR3	Activates numerous pro-tumorigenic signaling pathways such as PLC and Rho	Cellular membrane	S1PR3

Abbreviations: S1P = sphingosine-1-phosphate; SPHK1 = sphingosine kinase 1; SPHK2 = sphingosine kinase 2; S1PR1-3 = sphingosine phosphate receptors 1-3

Chapter 3: Methods

Aims 1 and 2: CPRD

Aims 1 and 2 utilize data from the Clinical Practice Research Datalink (CPRD) Gold, a prospective and longitudinal general practice research database in the United Kingdom (UK) that is one of the largest in the world. The CPRD started as the Value-Added Medical Products (VAMP) dataset in 1987, and became the General Practice Research Database (GPRD) in 1993 [157]. Finally, in 2012, it took on its current designation as the CPRD [158]. The CPRD provides prospectively collected exposure data on a representative sample of the population of the UK from 674 medical practices, with over 11 million participants, including historical records. England, Wales, Scotland and Northern Ireland are regions covered by the CPRD. Over 1000 peer-reviewed research articles have been published using the CPRD. CPRD uses a program called the Quality and Outcomes Framework that was initiated in 2004 to improve high quality recording of medical practice data [159].

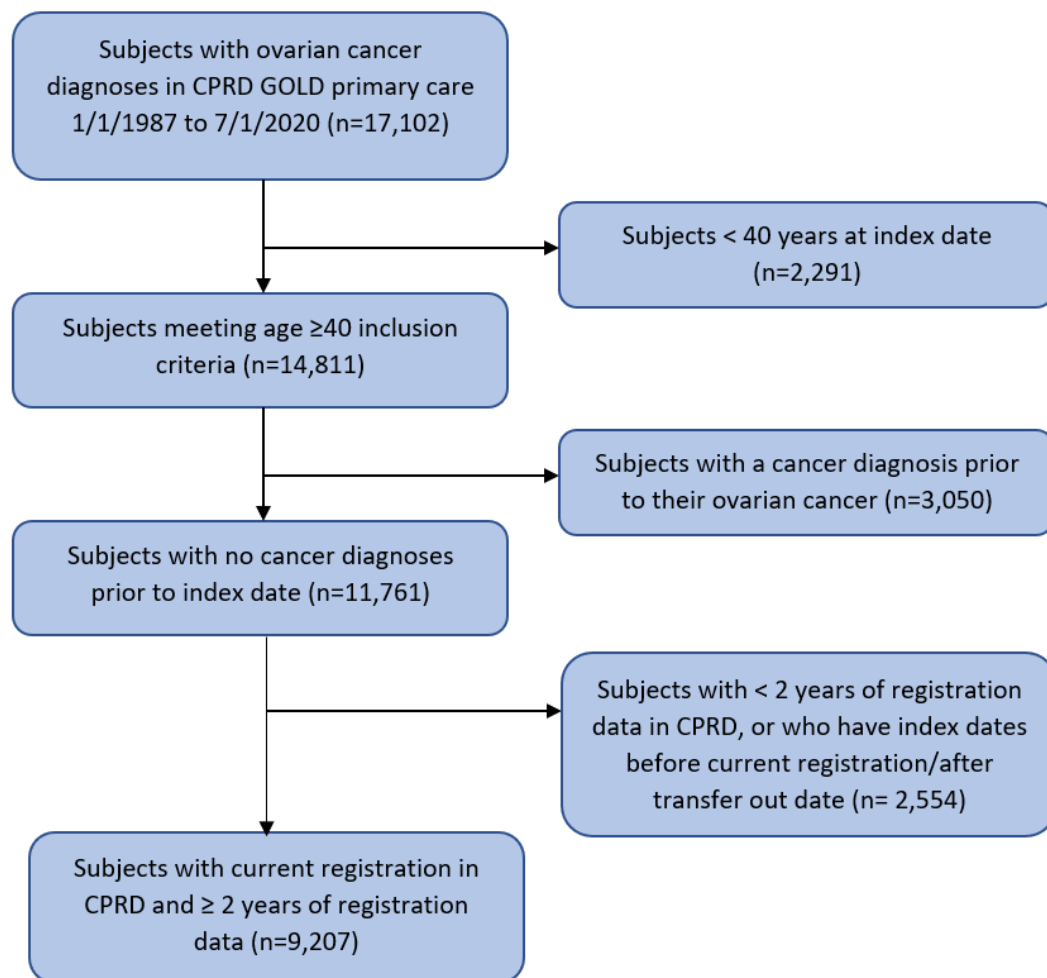
Study design

Using a nested case-control design, the associations of metformin use (Aim 1) and statin use (Aim 2) with risk of ovarian cancer were examined in the CPRD. The source population for aims 1 and 2 is the general population of the UK, as the CPRD is considered a representative sample of the entire UK. The nested case-control design allows for efficient sampling by selecting a subset of controls for comparison. Up to ten controls per case were chosen along with all cases diagnosed between 1987-2020 in CPRD [160]. As ovarian cancer is a rare outcome, the effect (odds ratios) approximates the risk ratios (RRs) calculated from the entire cohort.

Study population

The study population for Aims 1 and 2 included women ≥ 40 years old in the CPRD between the years 1987- 2020, without a history of prior cancer, except non-melanoma skin cancer and who had at least 2 years of data history in the CPRD. According to the most recent release of data from the Surveillance, Epidemiology and End Results Program (SEER), only 4-5% of ovarian cancers are diagnosed below age 40, and these cancers may be primarily non-epithelial in origin [161]. Ovarian cancer cases were identified by use of Read (not an acronym) or medical codes. These cases were matched 10:1 with at-risk, incidence density sampled controls in CPRD who have not had a surgical removal of both ovaries (oophorectomy). When more than 10 eligible controls per case were identified, 10 controls were randomly selected from the eligible control women. Frequency matching was performed on age, the general practice where the case was identified and the number of years the case was in the CPRD prior to diagnosis. Controls were alive and disease-free at the time of their matched cases' diagnosis, or index date, and had no prior cancer diagnoses except for non-melanoma skin cancer [162]. Frequency-matched factors were reviewed, and found to be balanced between cases and controls (Appendix E, Supplemental Table 2). The case selection is detailed in Figure 9. The final analytic population included 9,207 cases and 92,070 controls.

Figure 9. Case selection from CPRD, Manuscripts 1 and 2



Data management

The CPRD GOLD primary care data is a resource comprised of ten datasets that are available for request and can be linked to additional UK databases (linkage discussed in the following section). The 10 datafiles are 1) Patient, 2) Practice, 3) Staff, 4) Consultation, 5) Clinical, 6) Additional clinical, 7) Referral, 8) Immunization, 9) Test, and 10) Therapy file. Each patient characteristic or medical event is entered by CPRD-participating practices in the form of a READ (not an acronym) code. Depending on the type of READ code entered, the patient information is stored in one of the 10 files. Assembly of the CPRD dataset for Aims 1 and 2 of this dissertation began by sampling ovarian cancer cases and frequency-matched controls who had READ codes for the specified inclusion criteria contained in their patient history, captured across the primary care files (e.g., ovarian cancer case READ codes were captured in the clinical or referral files). READ codes defining cases, and frequency matching information for controls are captured in Appendices 1 and 2. Each variable used in the analysis for Aims 1 and 2 was created by assembling all READ codes available to providers for a specific characteristic or condition using the CPRD code browser, comprehensive lists of READ codes were compiled and following assembly, the lists of READ codes defined all variables. Variables were created based on their pre-defined structures. Binary variables were created by categorizing a patient as a “1” if a READ code was present anywhere in a patient file prior to the index date, and 0 if the READ code was not present. Categorical variables were similarly created. For example, alcohol status (current, former or never use) was defined using an additional component, the index date, to determine if a woman was currently drinking alcohol at the time, she entered the study. Dose variables for the Defined Daily Doses (DDD) required the creation of an algorithm as the

calculation of the DDD was complex, and DDDs are not specifically captured in CPRD GOLD (Appendix 3). Medication strength for each prescription of either metformin or statins was extracted from the READ code, and following extraction the algorithm for DDD calculation was applied (Appendix 4). As parity is a strong risk factor for ovarian cancer, the total number of pregnancies per woman was obtained via linkage to the newly-created pregnancy register [163].

Following variable creation, each variable was checked against a comparator estimate from published data to determine whether the observed descriptive characteristics of that variable (frequency or mean, depending on the variable type) was similar or dissimilar from expected values. For instance, the prevalence of statin use obtained from CPRD primary care data was compared to prevalence estimates for statin use obtained from the National Health and Nutritional Examination Survey (NHANES) estimates for statin use in the US for women >40 years (i.e., comparable to the CPRD population used for Aims 1 and 2).

Selection, information and confounding bias

Patients can opt out of including their personal data in the CPRD, and the opt out rate reported by the National Health Service (NHS) as of September 2021 is 5% of the UK population that registered with a general practice. Patient opt-outs were also balanced across general practices in the UK. It is therefore unlikely that selection bias was introduced into the study populations for Aims 1 and 2 by study by patient opt-outs in the UK [164]. As such, the very large (> 92,000 women) control population for the nested case-control studies for Aims 1 and 2 was sampled from non-cases who met the control sampling criteria, and unlikely to be selected based on the patient-opt out status of the larger cohort. For each nested case-control study for Aims 1 and 2, control selection was also conducted at random, independently of exposure status. It is therefore unlikely that the probability of selecting a control was influenced

by prospectively collected metformin or statin use. Selection bias may also occur in a case-control study when missing values for covariates lead to differences in analytic populations. To account for this possibility, missing values were retained in analyses by using missing categories for each covariate to ensure that all women were retained in the models. Finally, because control sampling was driven by the identification of cases and multiple matching factors were used, it is possible that the control set was not representative of the person-time experience of women in the CPRD, or in the underlying UK population [165]. However, as the controls were matched to cases 10:1, the control set was large and the size of the control group is expected to have aided in minimizing selection bias that may have resulted from selection driven by cases.

Regarding the representativeness of the source population by the controls and the external validity of Aims 1 and 2, all controls were selected from the CPRD, which is representative of the entire UK population (i.e., not hospitals or clinics, or a source that would produce “sicker” controls). Therefore, the results for the nested case-control studies produced from this analytic population are generalizable to populations with similar demographic characteristics, such as the US and other European populations. The temporal relationship between medication use and the ovarian cancer outcome, as well as the random selection of a large control set contributed to the internal validity of the nested case-control studies for Aims 1 and 2.

Information bias can occur due to measurement error or misclassification of exposures or outcome. In the CPRD primary care data, an assumption when creating most variables from READ codes is to classify a subject in the non-event category when a READ code is missing from their patient record. It is often not possible to differentiate whether the lack of a READ code is an indication of a non-event or if the data is simply missing (or was never recorded by a practitioner). This potential misclassification of the non-event group may have led to information

biases in the analysis that are acknowledgeable but not possible to rectify. With regard to the exposures for Aims 1 and 2, prospectively collected metformin and statin use is not expected to be differential by outcome status. This non-differential misclassification of exposure would be expected to cause a bias toward the null value. However, prevalence estimates for metformin and statin use were evaluated in the CPRD controls, and were similar to prevalence estimates in the US, a similar population to the UK, signaling that any potential misclassification of exposure status was likely low in the analytic population. For covariates, such as OC use, individuals in the non-event category may disproportionately represent that category with respect to the individuals with the event, leading to differential misclassification by exposure status and biasing the effect estimate away from the null. This complication is highlighted by the parity variable used for this analysis. Women without a value for the number of pregnancies they have experienced are grouped together. These women could be either nulliparous, or missing pregnancy data. It is not possible to determine true nulliparity from true missing.

Confounding bias may occur if not accounted for in the study design phase by procedures like matching, or if statistical models are not appropriately adjusted for all variables that are associated with both the exposure, the outcome, and are not on the causal pathway between the exposure and outcome. In the analyses of association between metformin and statin use with ovarian cancer risk, confounders were chosen a priori and tested using the 10% change in estimate rule to determine whether they were confounders in the CPRD data. Strong risk factors for ovarian cancer, like OC use and parity are established covariates in ovarian cancer risk factor analyses and models for Aims 1 and 2 included both of these factors. Others, such as BMI, diabetes status, alcohol and smoking were evaluated with the 10% change in estimate rule. Although important adjustment factors for Aims 1 and 2 were captured, there is a possibility of

residual confounding from some covariates. However, OC use and parity, although strongly associated with ovarian cancer, were not associated with either metformin or statin use in the CPRD analytic population, and so if misclassification in these variables occurred, it would not be expected to influence the association between metformin or statins and ovarian cancer risk. BMI, however, was a confounder in these data, and if residual confounding occurred from the BMI variable, due to missing BMI variables, the effect estimate may be biased away from the null value, as BMI was differential by medication status but not by outcome status. Final models were adjusted for OC use, parity and BMI, as other theoretical confounders did not meet the 10% change in estimate criterion.

Statistical analysis

A detailed description of statistical methods used for Aims 1 and 2 of this dissertation are found in the respective manuscripts, Chapters 4 and 5, “Statistical analysis”. Briefly, multivariable logistic regression was used to assess the associations (OR, 95% CI) between metformin and statins with ovarian cancer risk overall and in important subgroups. Final models were determined by incorporating a priori confounders (OC use and parity) and confounders determined by utilizing the 10% change in estimate rule (BMI).

Power analysis

The power calculations for three prevalence estimates of metformin or statin use in the controls for Aims 1 and 2 (5%, 10% and 20%) are shown below. Minimal detectable odds ratios range from 0.86-0.93 and are based on a fixed sample of 9,200 ovarian cancer cases and 92,000 controls. Probability levels for the exposures in the controls were chosen based on a range of

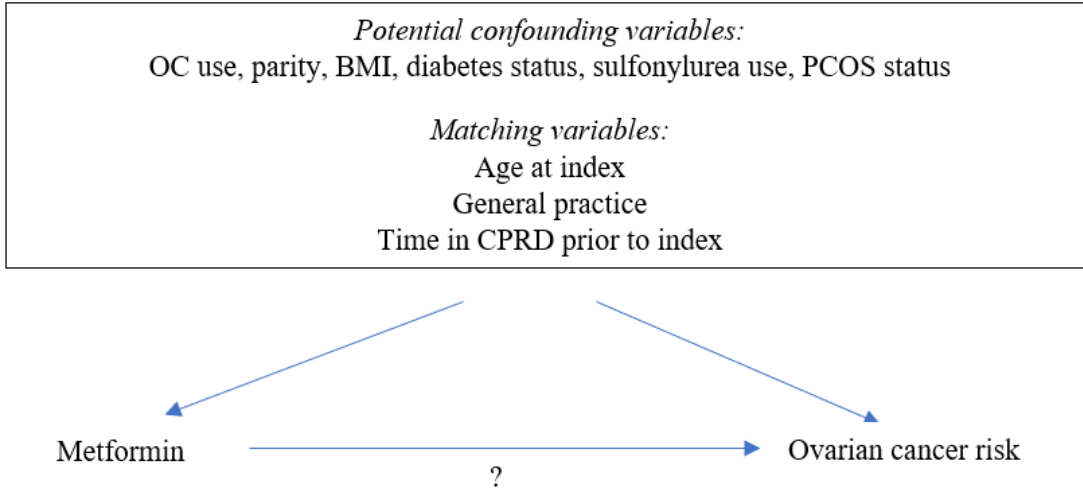
expected probabilities from published literature using similar populations and on nationwide data from the US (NHANES) [123, 166].

Figure 10 Power analysis for Manuscripts 1 and 2

Probability of Exposure in Controls	Minimal Detectable OR Ovarian Cancer
5%	0.86
10%	0.90
20%	0.93

Theoretical directed acyclic graphs (DAGs) Aims 1 and 2

Metformin:



Statins:

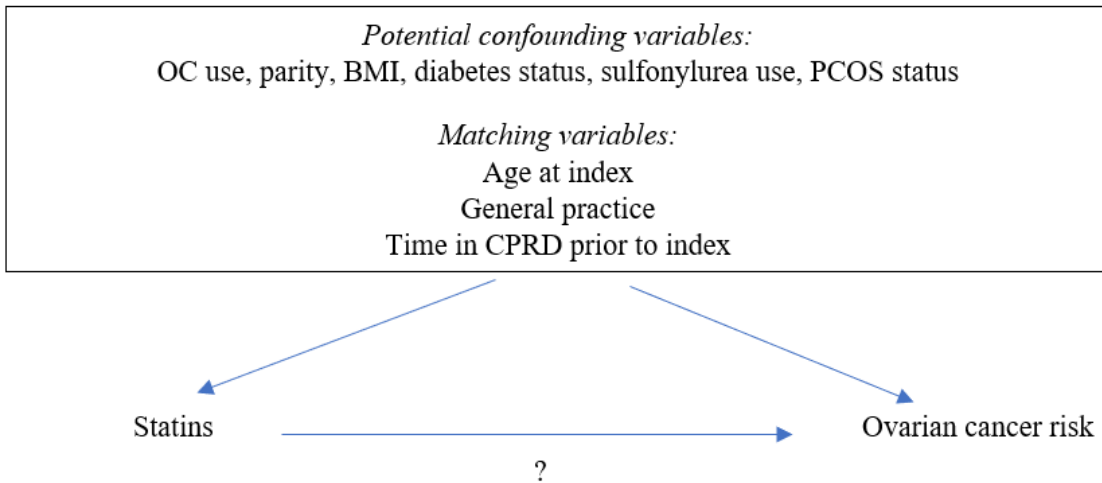


Figure 11. Directed Acyclic Graphs for Manuscripts 1 and 2

Outcome, exposure and covariate window considerations in electronic medical record (EMR)-derived data

A major strength of the CPRD primary care data is the availability of prospectively collected information. In the present analysis, the 30-year follow-up time window allows for ascertainment of critical information from the initiation of the CPRD precursor datasets (Value Added Medical Products, VAMP) and the General Practice Research Database (GPRD) to the present-day version of the CPRD. Because the study window spans multiple decades, it was important to consider that the prevalence of some factors in the current analysis were different according to decade of birth. The following figures (Figures 12-16) depict how ascertainment of ovarian cancer, exposures and covariates may have differed over time for birth cohorts. Five birth cohorts are used in each figure as a representative example of the concept, but eight birth cohorts were evaluated in analyses for Manuscripts 1 and 2, defined by decade of birth.

The capture of ovarian cancers in the oldest women in the dataset was lower than later years, as it is possible that some women were diagnosed with ovarian cancer and died prior to the initiation of the UK primary care data collection. Descriptions of time-varying capture of selected CPRD variables are shown in Figures 12-16.

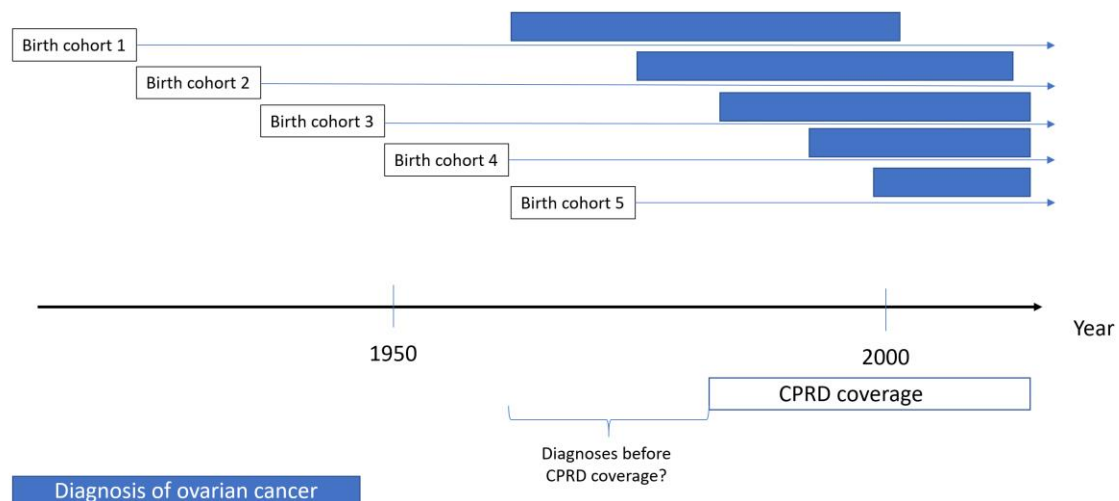


Figure 12. Ovarian cancer capture by birth cohort

With regard to the exposures for the Aims 1-2, metformin and statins, each drug was introduced for use in the UK at a different time. Metformin was available beginning in the 1950s, prior to the initiation of data collection in the earliest version of CPRD (1987). Because of the early introduction of metformin as a diabetes control agent and the age of index dates for women in this analysis (> 40 years), women in each birth cohort in the present analysis would have access to metformin if they presented with diabetes. Of note, metformin and statin use windows do not extend to the end of each graph for all birth cohorts, as it is expected that women in the earliest birth cohorts will have died in the earliest years of the study window for the dissertation analyses.

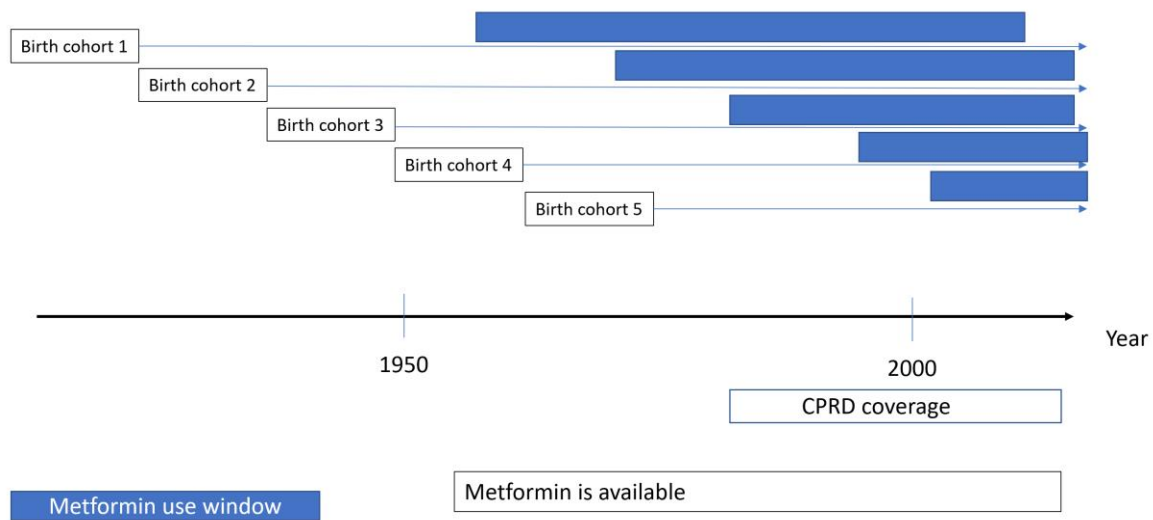


Figure 13. Metformin use windows by birth cohort

The first statin was introduced in the UK in 1987, early in the history of primary care data collection in the UK. Other statin types were variably introduced over the following years. Women in the earliest birth cohorts in this analysis were likely to have lower use of statins compared to later birth cohorts, these women were older when statins were introduced, and the uptake of statin use was initially slow.

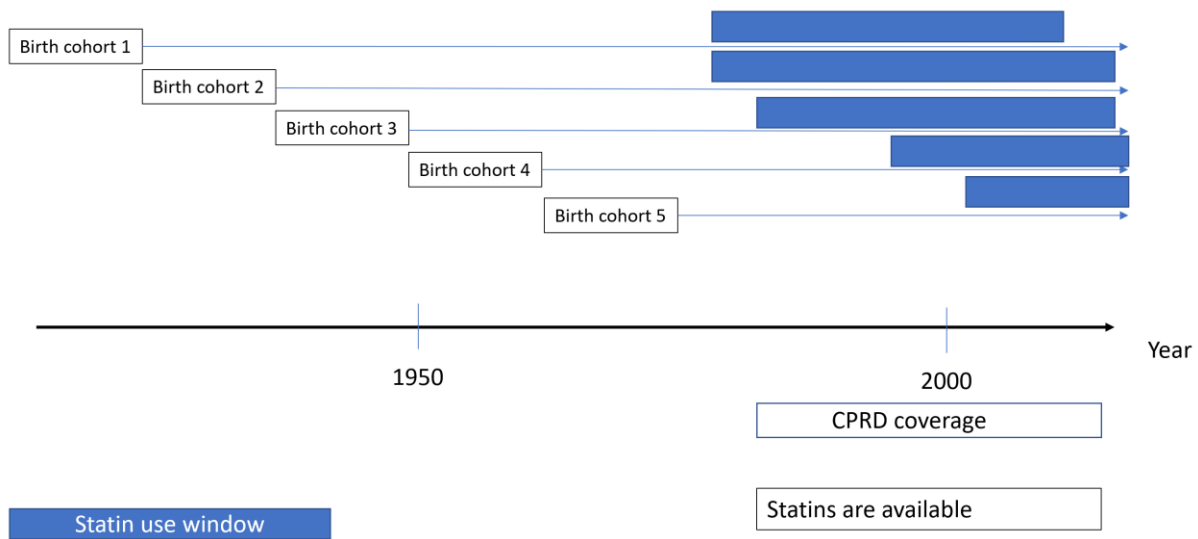


Figure 14. Statin use windows by birth cohort

Oral contraceptives were introduced in the 1950's, but did not become widely available until the 1970's. Women in the earliest birth cohorts would not have had access to OCs during their reproductive years, making OC use in the oldest women in the analytic population very low. Women who experienced their reproductive years prior to the start of CPRD would also have not had their pregnancy information recorded, resulting in a low capture of pregnancy events for women in the earliest birth cohorts, and an increase in the capture of pregnancy events over time as the reproductive years of more women overlapped with the study window. As the prevalence of OC use was different by birth cohort, this may also have been a source of misclassification of exposure in Aims 1 and 2. However, as OC use was only associated with ovarian cancer and not with either exposure for Aims 1 and 2, misclassification of OC use is not expected to have biased the effect estimates.

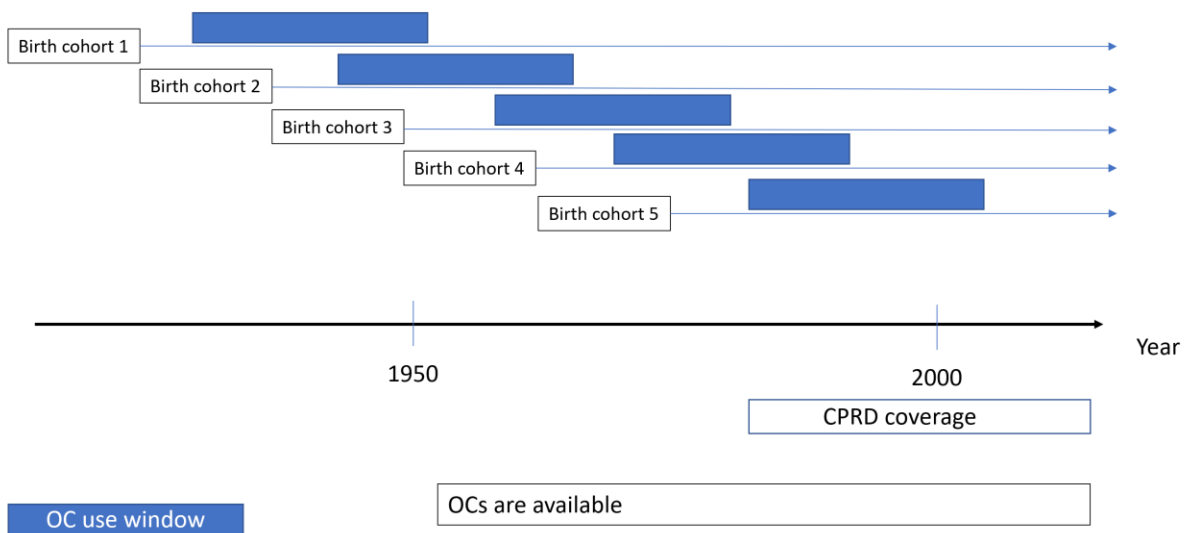


Figure 15. Oral contraceptive use in CPRD according to birth cohort

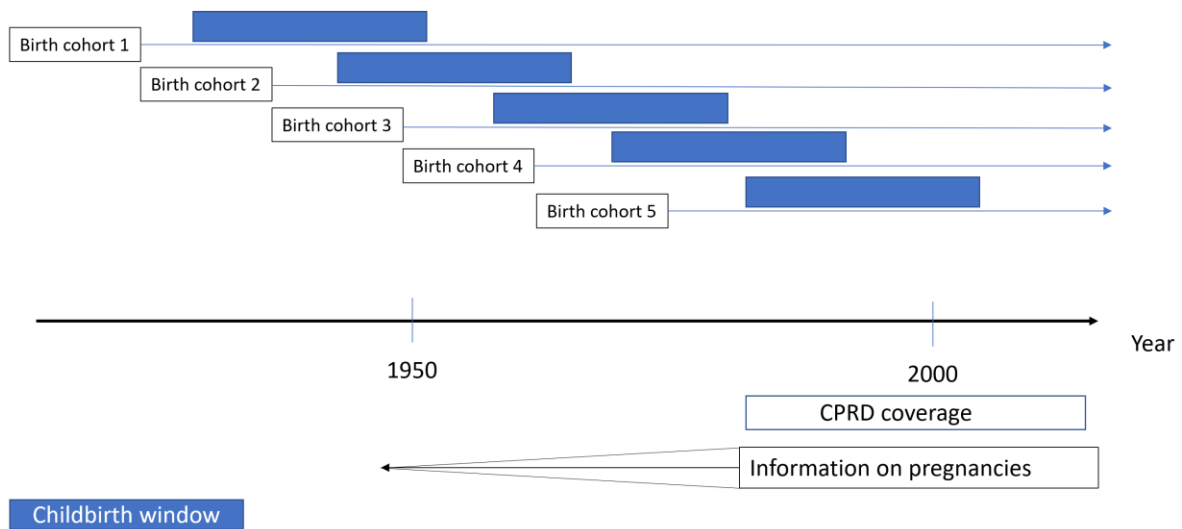


Figure 16. Reproductive window overlap with CPRD study window

To address the complexities in capturing critical variables over the study period, the present analysis evaluated the prevalence of study variables, including the outcome, exposures, reproductive and behavioral factors over time. The manuscripts for Aims 1-2 also introduced birth cohort evaluations as well as analyses in subsets according to historical CPRD dataset changes and the density of case capture to ensure that variations in each of these factors did not affect the primary association under study in each paper.

Aim 3: The Polish Ovarian Cancer Study (POCS)

Aim 3 is a case-only analysis performed in a population of Polish women with ovarian cancer who participated in the Polish Gynecologic Cancer Case-Control Study conducted in Warsaw or Lodz, Poland between June 1, 2001 and December 30, 2003. The Polish Gynecologic Cancer Case-Control Study is a case-control study of ovarian and endometrial cancer with a total of 317 incident ovarian cancer cases. Women with ovarian cancer were recruited from five Polish hospitals: the Cancer Center, M. Curie Skłodowska Institute of Oncology, the Polish Oncological Foundation, Dr. Madurowicz Memorial Hospital, and the Polish Mother's Health Memorial Hospital by rapid identification system [167]. Additionally, the cancer registries of Warsaw and Lodz were reviewed for any ovarian cancer case that was not determined by the hospital identification system. Demographic, reproductive and other lifestyle factor information were collected by interviewer administered questionnaires and medical records provided pathology and treatment information [167, 168].

Women with ovarian cancer (cases), recruited from Warsaw or Lodz, Poland from June 1, 2001 to December 30, 2003 between the ages of 20-74 years were included in this study. Cases had a diagnosis of ovarian cancer and an available tissue sample for analysis. As this is a

case-only study, no control subjects were utilized for comparison. The analysis considered an outcome defined as positive versus negative tissue expression of two lipid pathway markers, and exposures were ovarian cancer risk factors and tumor characteristics, defined thoroughly in Chapter 6 of this dissertation.

Data assembly

A set of 175 (55%) of ovarian cases had tumor tissue available for the analysis. Immunohistochemistry (IHC) methods were used develop assays for S1P pathway markers. Pilot testing for each antibody was conducted on control tissue for each marker, as well as on a commercial ovarian tissue TMA that included all ovarian cancer histotypes. Following establishment of the IHC stains, study TMAs were stained for each marker using the antibody concentrations developed during the pilot phase. Two S1P pathway markers were chosen for analysis (S1PR1 and S1PR3) based on the quality of pilot staining and data availability. There were six TMA cores that were not able to be scored due to low tissue quality, and three women with missing covariate information for OC use and menopausal hormone therapy that could not be included, leaving an analytic population of n=166. A pathologist manually scored each stained slide for the number of cells stained positive for each marker, recording this value as both a percentage of stained cells and a staining intensity.

Selection, information and confounding bias

The samples contained on the study TMAs (n=166) were compared to those that were not contained on the study TMAs (n=142) (Appendix G, Supplemental Table 2). With regard to potential selection biases, the total ovarian cancer cases occurring in Warsaw and Lodz during the study years are expected to have been captured in POCS as both the hospital systems and

cancer registers were utilized for case capture, and so encompass the total number of possible ovarian cancers that occurred during those years. Additional checks for selection bias with regard to selection occurring when assembling the TMA were conducted, as all women were not captured on the TMA due to lack of tissue availability. Distributions of demographic, anthropometric and reproductive characteristics between the set contained on the TMA and the set not contained on the TMA were similar when assessed by univariate (chi square) statistics (all $p > 0.05$). Tissue samples contained on the study TMAs were more complete with respect to stage and grade information, but had fewer endometrial cancers than the set that was not included on the TMAs. However, the frequency of endometrial tumors (10%) on the TMAs is the expected prevalence of endometrial tumors with respect to all epithelial types, and study estimates are not expected to be influenced by any differences between tumors not included on TMAs and tumors included on TMAs.

Potential information bias was assessed also assessed for Aim 3. Each ovarian cancer sample was evaluated by a pathologist and categorized to a histology group. However, 63 POCS women had mixed cell or were not categorized in the four primary histotype groups. To address any misclassification bias that may have occurred with regard to the other histologies, this dissertation reviewed these tumors separately from the primary histotypes (serous, endometrioid, mucinous and clear cell). To account for any possibility of misclassification among the tumors with other histologies, this study included separate analysis among well-differentiated and poorly differentiated tumors, as well-differentiated tumors are more often non-serous, and poorly differentiated tumors are often high grade serous. With regard to study variables, women in POCS were provided questionnaires upon study entry, and a trained interviewer collected self-reported information. Recall bias is therefore possible in the POCS data, but is expected to be

non-differential by outcome and exposures status, as expression data for the present analysis were determined after variable collection. Finally, a trained pathologist scored each tumor core on the study TMAs to determine the extent of IHC staining for each S1P pathway marker of interest. The manual scoring technique was preferred over automated scoring for this project, as the markers under evaluation are newly established in ovarian cancer. The manual scoring technique also removed the possibility that the scores for the tumor cores were incorrectly attributed.

Confounding bias was assessed in the POCS data for demographic, anthropometric and reproductive characteristics. Final models were adjusted for both a priori covariates identified from the literature specific to the POCS data [52] as well as additional factors identified via the 10% change in estimate strategy. As final models were adjusted for critical risk factors for ovarian cancer as well as factors potentially related to S1P expression, model adjustment likely corrected for any imbalances by outcome status. However, because S1P expression markers are not well-defined in ovarian cancer, it is possible that an unmeasured confounder was not tested for based on a lack of a priori evidence for its association with expression status, tumor characteristic or a risk factor. This unmeasured confounding would bias the effect estimate toward or away from the null in a manner specific to the potential confounder.

Statistical analysis

The statistical analysis for Aim 3 can be found in Chapter 6, “Statistical analysis”. As two cores per woman were included on TMAs, correlations between each core per woman were evaluated and found to be $\rho > .80$ by Spearman’s rank coefficient. The core with the highest percentage of cells stained positive (area) was chosen for analysis [52]. The primary analyses for this Aim were to evaluate differences across histotype for S1PR1 and S1PR3, associations with

tumor characteristics and survival by marker status. Categories of marker expression were defined as dichotomous (0 = below median, 1 = above median) and as a four-level variable (0 = none, 1 = <25%, 2 = 50-75% and 3 = $\geq$ 75%) based on the distribution of expression status. Logistic regression models were used to determine associations between tumor characteristics and risk factors for ovarian cancer with marker expression status. Survival by marker expression status was evaluated by Kaplan-Meier survival curves, plotting the proportion of women alive against time since interview in days. Log-rank tests were used to determine statistical differences between expression groups for each marker.

Power calculations

For manuscript 3, the power calculation is based on a fixed study sample of n=166 women for the comparison of higher versus lower expression of S1P markers. For the comparison of covariate and tumor characteristics with exposure probabilities of 0.2, 0.3 and 0.4, representing a range of probabilities for both ovarian cancer risk factors and histology prevalence estimates. There will be 80% power to detect odds ratios between 0.21-0.38. Probability levels for exposures were chosen based on a range of expected percent positive cells stained from previous literature [56, 151].

Figure 17. Power analysis for Manuscript 3

Probability of Exposure	Minimal detectable OR
20%	0.21
30%	0.32
40%	0.38

Chapter 4: Manuscript 1

Association between metformin use and ovarian cancer risk in the Clinical Practice Research Datalink (CPRD)

Abstract

Introduction Epithelial ovarian cancer (EOC) is the most fatal gynecologic malignancy. Several metabolic factors, including obesity, diabetes, and circulating lipids are associated with EOC. Drugs targeting these conditions, such as metformin, have drawn interest as potential chemoprevention agents for various cancers, yet few studies have been conducted for associations with risk of EOC. Importantly, subgroup associations according to formulation, cumulative exposure, and indication are critical to evaluating possible benefit groups. Metformin use was comprehensively evaluated with risk of EOC in the Clinical Practice Research Datalink (CPRD).

Methods A nested case-control study was conducted in CPRD, which is the primary care database of the United Kingdom and contains data collected prospectively from >10 million residents and >2,000 general practices (GPs). Cases were women diagnosed with EOC between 1987–2020, aged >40 years, with ≥ 2 years of CPRD data, and no prior cancer (except non-melanoma skin cancer) (n=9,207). Controls had the same inclusion criteria, were incidence-density sampled, were cancer free at the time of selection, and matched to cases 10:1 on age, GP, and time in CPRD (n=92,070). Medication use was defined as having ≥ 1 prescription up to one year prior to date of diagnosis or control selection. Logistic regression was used to determine associations via odds ratios (ORs) and 95% confidence intervals (CIs) between metformin use

and EOC risk overall, according to medication use characteristics (prescription count, duration of use), drug indications and concomitant medication use.

Results Metformin (OR 1.01, 95% CI 0.91-1.13) was not associated with EOC risk overall, or according to increasing prescription counts, higher duration of use, or among women using concomitant medications (a proxy for disease severity). Metformin was not associated with EOC in any body mass index category (ORs 0.84-1.03). Associations between metformin and EOC did not differ for primary indications (p_{int} 0.16-0.69). A non-significantly lower risk of EOC for metformin users was observed in a small group of women with but not without polycystic ovarian syndrome, a secondary metformin indication (p_{int} 0.01).

Conclusion In this comprehensive study using the largest series of EOC cases to-date, metformin use was not associated with EOC risk overall, according to drug characteristics or indications for use among women indicated for use.

Impact Potential use of metformin to lower the risk of EOC has received considerable attention, but findings from this large population-based analyses did not support a role in chemoprevention.

Introduction

Ovarian cancer is the most fatal gynecologic malignancy worldwide with over 314,000 incident cases and >200,000 deaths estimated to have occurred for the year 2020 [2]. Lack of symptoms early in the disease process leads to > 70% of cases identified at a late stage, when 5-year survival estimates are only 30% [5]. No effective screening approaches exist despite extensive efforts to identify biomarkers for detecting ovarian cancers [6, 7]. Identification of agents for primary prevention may reduce ovarian cancer incidence and may help in elucidating pathways involved in carcinogenesis.

Metformin is one among several candidate drugs that have been considered for ovarian cancer prevention. Studies have shown that critical components of cellular metabolism which rely on the availability of glucose, such as fatty acid synthesis, may be required for oncogenesis [169]. Metformin is an oral biguanide commonly prescribed for treatment of type 2 diabetes and polycystic ovarian syndrome (PCOS), two conditions associated with high body mass index (BMI) and potentially linked to increased ovarian cancer risk [102, 113]. Metformin's primary mechanism of action is to reduce glucose levels by suppressing gluconeogenesis in the liver and reducing glucose uptake in the intestines [170], which may result in modifications of cellular lipid composition and changes in cellular metabolism [171]. With these characteristics, metformin has recently become a candidate drug for chemoprevention in several cancer sites, including ovarian cancer [33, 172-174].

Observational studies evaluating the association between metformin and ovarian cancer risk have shown mixed results, with some studies pointing to a possible risk reduction for metformin users [175-177], and others showing no associations [36, 37]. Importantly, existing research has been primarily conducted among patients with diabetes rather than with other

indications such as PCOS [36, 38, 39] and follow-up times were limited in some studies (five to seven years) for assessing metformin exposure windows as well as ovarian cancer incidence [38, 39]. Most prior studies included a few hundred ovarian cancer cases or were limited with respect to exposed cases, affecting the ability to investigate potential benefit groups such as drug indication, duration, dose, and the use of concomitant medications [36-38, 40]. Finally, the existing literature varied in the capture of important adjustment factors critical for ovarian cancer analyses such as oral contraceptive use or parity. To address these gaps, this research project utilized the Clinical Practice Research Datalink (CPRD), the nationally representative primary care database of the United Kingdom (UK). With over 9,000 ovarian cancer cases identified from six million women during 1987-2020, this analysis represents the largest evaluation of the association between metformin use and ovarian cancer risk to date.

Methods

Study population

This nested case-control study was conducted in the CPRD GOLD, a primary care database containing a representative sample of the UK population. Cases were identified as women diagnosed with ovarian cancer between January 1987 and July 2020, who were at least 40 years of age at diagnosis [161, 178], with at least two years of history in the primary care data, and without a prior diagnosis of cancer except non-melanoma skin cancer (Figure 9). Ten controls per case were selected, via incidence density sampling without replacement, from women who were at risk of ovarian cancer (i.e., without history of bilateral oophorectomy prior to the diagnosis date of an incident case). Control subjects were required to have no prior diagnosis of any cancer except for non-melanoma skin cancer and were selected from individuals

with at least two years of data history in the CPRD. Controls were frequency matched to cases on year of birth (+/- 1 year), the general practice of the case, and the number of years of data in CPRD prior to the case's index date [135]. Data recording routines of general practices may vary, and an individual's length of CPRD data history influences the ascertainment of the exposure windows of interest, making it important to match on these criteria. CPRD anonymizes and protects all patient data and ensures that individuals providing data for research purposes are aware of their personal health information (PHI) rights. All patients in CPRD-participating practices are provided the option to opt-out of data-sharing with CPRD. The analytic population for the current analysis included 9,207 cases and 92,070 controls and the selection process for cases is detailed in Chapter 3, Figure 9. The frequency matching for controls is shown in Appendix B. The current analysis was approved by the Institutional Review Board (IRB) of the University of Maryland (UMD) (Appendix H).

Outcome definition

Ovarian cancer cases were defined as any first-time occurrence of a malignant neoplasm or cancer of the ovary, fallopian tube, broad ligament, parametrium or uterine adnexa in the CPRD Gold between the years 1987-2020 (Figure 9). The date of diagnosis of each case (index date) was defined as the date the ovarian cancer was recorded in CPRD. The outcome definition was assembled using Read codes for ovarian cancer in the CPRD primary care data and can be found in Appendix A.

Exposure definition

Metformin use was defined as having at least one prescription (use) for metformin recorded in their CPRD GOLD history at least one year prior to the index date. Metformin use

was also assessed by evaluating the number of metformin prescriptions in quartiles based on the distribution of metformin prescriptions in the controls ($Q1 \leq 11$, $Q2 \leq 30$, $Q3 \leq 62$, $Q4 \leq 656$). Metformin duration was defined by the cumulative years a woman was exposed (≤ 1 year, 1-3 years, 3-5 years, 5-10 years, and >10 years) based on the distribution in the controls. Metformin dose was evaluated using Defined Daily Doses (DDD) defined by the World Health Organization (WHO) as the average maintenance dose per day for a drug [179]. The DDD assignment for metformin is 2g. DDDs for each prescription were calculated by creating a DDD conversion factor (WHO-assigned DDD/ strength for the prescription in milligrams) and multiplying this conversion factor by the number of tablets for that prescription. Cumulative DDDs were calculated by summing the total DDDs across metformin prescriptions in a patient record. Read codes were used to define metformin use in the CPRD primary care data and can be found in Appendix A.

Study covariates

All study variables in addition to the outcome and exposure were assembled using Read codes, by an algorithm applied to the CPRD primary care data or by linkage. Oral contraceptive (OC) use (ever versus never use), diabetes status (ever versus never), polycystic ovarian syndrome (PCOS) (ever versus never), smoking status (current, former, never) and alcohol use (current, former, never), cardiac glycosides (ever versus never), insulin (ever versus never), sulfonylurea use (ever versus never), beta blockers (ever versus never) and statins (ever versus never) were assembled by Read codes. BMI (kg/m^2) was determined by an algorithm applied to the CPRD data using height and weight measures, and the BMI measure at least 5 years prior to ovarian cancer diagnosis date (or five years prior to index date for controls) was determined for each woman. BMI was categorized as ($<18.5 \text{ kg}/\text{m}^2$ underweight, $18.5\text{-}24.9 \text{ kg}/\text{m}^2$ normal

weight, 25-30 kg/m² overweight and 30+ kg/m² obese or missing). The number of pregnancies per woman was obtained via linkage to the CPRD Pregnancy Register which applies a complex algorithm to all women with primary care records to determine how many pregnancies she had recorded in CPRD. Parity was categorized as (1, 2, 3, 4+ and nulliparous/missing) [163].

All variables coded ever versus never had no missing values, as an underlying assumption in CPRD for most variables is that the absence of a Read code indicates a non-event. Other variables, such as parity, BMI, smoking status and alcohol use require a different inference. For example, a missing BMI value cannot be interpreted as a non-event, as all study participants are expected to have a BMI value across the lifespan. As such, the percent of missing information in controls was determined to be 40% for BMI, 66% for parity, 15% for smoking status, and 23% for alcohol use.

Statistical analysis

Univariate statistics comparing cases and controls were determined via chi-square tests for categorical variables and t-tests of mean differences between groups for continuous variables. Multivariable logistic regression was used to estimate odds ratios (ORs) and 95% confidence intervals (CIs) for the association between metformin use, dose and duration with ovarian cancer risk. OC use, BMI, diabetes status, sulfonylurea use, PCOS, and parity were evaluated as potential confounders. Final models were determined by incorporating *a priori* ovarian cancer risk factors (OC use and parity) as well as covariates determined to be confounders by the 10% change-in-estimate strategy (BMI). Age, general practice, and time in CPRD were used as frequency-matching factors in the design phase of the study and not included in final models. The primary effects of BMI, diabetes, and PCOS with ovarian cancer risk were assessed, followed by the effect of metformin use within indication group. Statistical heterogeneity in

effect sizes between indication groups was determined by fitting a model with an interaction term between metformin and each indication. Indications for metformin often coexist with other comorbidities like cardiovascular disease which may be treated with concomitant medications. Additionally, type 2 diabetes may be treated with multiple medications even in the absence of comorbid conditions. Thus, the joint effects of metformin and use of medications commonly prescribed for diabetes control or cardiovascular disease (sulfonylureas, insulin, cardiac glycosides, beta blockers, and statins) were evaluated by creating a categorical variable indicating the presence or absence of each medication: metformin and a second medication (++), metformin alone (+ -), no metformin but a different medication (- +), or neither medication (- -). The reference category for joint effect analyses was the category using neither medication (- -). Categorical analyses containing counts in any strata less than five were not modeled. As both metformin use as well as intake of important adjustment factors, like OC use, may differ for women across time, the association between metformin and ovarian cancer risk was evaluated by birth cohorts (1899-1910, 1911-1920, 1921-1930, 1931-1940, 1941-1950, 1951-1960, 1961-1970, 1971-1979). Because the parity variable obtained from the Pregnancy Register included a category that combined women missing parity information and nulliparous women which were not possible to separate, a sensitivity analysis was conducted in the youngest women born after 1960 whose missing pregnancy information was anticipated to align more strongly with nulliparity rather than missing values. Additional analyses were conducted dividing the analytic population by specific dates related to historical CPRD expansions (Value Added Research Products (VAMP) 1987-1993, General Practice Research Datalink (GPRD) 1993-2012, and CPRD 2012-present) as well as according to timepoints capturing the density of cases across the study period. Pharmacoepidemiologic literature specific to CPRD analyses of drug-cancer

associations [180]. A p value less than 0.05 was used to determine significance in all tests of association. All analyses were performed in SAS statistical software version 9.4, (Cary, NC).

Results

Characteristics of study participants are presented in Table 7. The average birthyear for all participants was 1940 (1899-1979). As age was a matching factor, ages at index date were 64.7 ± 12.0 years. With regard to reproductive factors, cases were less likely than controls to use OCs and were more likely to have a family history of ovarian cancer ($p_{\text{chisq}} < 0.01$). Controls were more likely to have multiple pregnancies ($p_{\text{chisq}} < 0.01$). With regard to behavioral and anthropometric factors, the proportion of smoking and alcohol use were similar between cases and controls, but cases were more likely to be obese (BMI $30+ \text{ kg/m}^2$). All cases were diagnosed between 1987-2020, with 2-4% of the total case number diagnosed in each study year. (Appendix E, Supplemental Table 1).

Study characteristics by 10-year birth cohort

The characteristics of study participants were additionally evaluated by birth cohort to assess changes in reproductive, anthropometric, and behavioral factors over the study period (Appendix E, Supplemental Table 2). Mean year at diagnosis was 1992 (1987-1999) in birth cohort 1899-1910 and 2016 (2011-2020) in birth cohort 1971-1979. Considering reproductive factors, the use of OCs increased in later birth cohorts. The youngest women in the study sample (birth years 1961-1979) had the highest use of OCs (34-42%). OC use was $<16\%$ among women born between 1951-1960 and $<5\%$ among those born 1899-1950. Generally, OC use was higher in controls than in cases in all birth cohorts. The CPRD Pregnancy Register does not allow to distinguish nulliparity from missing parity information, and misclassification of the non-parous

state provided by the Pregnancy Register depends on the extent of the missing parity data, which changed over time. Pregnancy events occurred rarely among women born between 1899-1960, with the frequency of the missing/nulliparous category ranging from 99% in the earliest birth cohort (1899-1910) to approximately 50% in women born between 1951-1960. For women born after 1960, missing/nulliparous pregnancy data reduced to approximately 30%, with higher frequencies of multiple pregnancies among controls. Family histories of ovarian cancers were recorded in less than 1% of cases and controls in all birth cohorts. History of behavioral factors, such as smoking and alcohol use, were similar between cases and controls in all birth cohorts. However, the increased prevalence of obesity among cases compared to controls was evident among women born between 1930-1979. Missing data for behavioral and anthropometric factors decreased with increasing birth cohort.

Primary association between metformin and ovarian cancer risk

There were 379 case subjects (4.1%) and 3,490 controls (3.8%) in the study population who ever used metformin (Table 8). There was no association between ever-use of metformin and ovarian cancer risk overall after adjustment for BMI, OC use, and parity (OR 1.01, 95% CI 0.91-1.13) (Table 2). Across birth cohorts, metformin use varied from 0.8%-5.5% among cases and 0.5%-5.2% among controls. No association between metformin use and ovarian cancer risk was observed in any birth cohort (Table 9). Because a decline in the frequency of the missing/nulliparous category of pregnancy events with increasing birth cohort was observed, the missing values in later birth cohorts may align more strongly with nulliparity rather than true missing values. As such, a sensitivity analysis treating the missing/nulliparous category as nulliparous in the later birth cohorts (1960-1971, 1971-1979) was conducted, but the association between metformin and ovarian cancer risk remained null (Table 9). There was also no

association between metformin and ovarian cancer risk in sensitivity analyses dividing the CPRD population according to dates of historical CPRD changes, as well as according to dates corresponding to the density of case capture across the study period (Appendix E, Supplemental Table 3-4).

An evaluation of the effect of higher cumulative metformin exposure, measured in number of prescriptions, showed that risk of ovarian cancer was elevated in the group of women with the lowest cumulative prescription number (1-11) compared to women who never used metformin (OR 1.31, 95% CI 1.05-1.64). However, no difference in risk was observed for women in groups with quartiles of higher cumulative prescription counts users (12-30, 31-62, 62+ prescriptions) compared to never users of metformin. There was also no association between metformin and ovarian cancer risk according to duration of use (Table 8). There was no association between metformin in any DDD quartile with ovarian cancer risk (1-240, 241-749, 750-1802, 1803+) (Table 8).

Association between metformin use and ovarian cancer risk by BMI

Analyses of the primary effect of BMI on ovarian cancer risk showed that overweight and obese women had a significantly higher risk of ovarian cancer ≥ 5 years before ovarian cancer diagnosis compared to women with normal BMI (overweight: 1.10, 95% CI 1.03-1.18; obese: OR 1.19, 95% CI 1.11-1.28) (Appendix E, Supplemental Table 5). Underweight status (BMI $< 18.5 \text{ kg/m}^2$) was associated with a significantly decreased risk of type 2 diabetes (OR 0.57, 95% CI 0.44-0.74), but was associated with a significantly increased risk of PCOS (OR 1.83, 95% CI 1.04-2.32). Overweight status (BMI 25-29.9 kg/m^2) was significantly associated with an increased risk of type 2 diabetes (OR 2.49, 95% CI 2.28-2.72), and risk was further increased for obese women (OR 7.18, 95% CI 6.60-7.82). Overweight status was not associated with PCOS,

but being obese was (overweight: OR 1.14, 95% CI 0.79-1.64; obese: OR 2.32, 95% CI 1.65-3.28) (Appendix E, Supplemental Table 5). Finally, although BMI was associated with both ovarian cancer risk and metformin indications, no association between metformin and ovarian cancer risk was observed in any category of BMI (low, normal weight, overweight or obese) (Table 10).

Evaluation of metformin and ovarian cancer risk by indication group

There were 610 cases diagnosed with type 2 diabetes with a mean age of 63 ± 11 years, and 5907 controls, diagnosed at a mean age of 62 ± 11 years. The average duration of type 2 diabetes prior to ovarian cancer diagnoses was 8 ± 6 years for cases and 8 ± 7 years prior to index date for controls. Type 2 diabetes was not associated with ovarian cancer risk overall (OR 0.98, 95% CI 0.89-1.07) in a model adjusted for OC use, parity and BMI. A sensitivity analysis removing subjects who were diagnosed with diabetes 0-3 months prior to their ovarian cancer diagnosis, to limit the possibility of an undiagnosed ovarian cancer occurring prior to diabetes onset also showed no association (OR, 0.97, 95% CI 0.89-1.06) (Appendix E, Supplemental Table 7). Most metformin users were diabetics (96.2%) and the average duration of metformin use among diabetics was longer than that of non-diabetics (5 ± 4 years vs. 2 ± 1 year). In this analysis, no difference was observed between groups defined by presence or absence of type 2 diabetes (type 2 diabetes: OR 1.09, 95% CI 0.92-1.30; non-type 2 diabetes: OR 1.07, 95% CI 0.59-1.95; p_{int} 0.16) (Table 11).

The prevalence of PCOS was low in the study population (cases: 37 (0.4%); controls: 239 (0.3%) (Table 11). The average age at PCOS diagnosis was 35 ± 11 years for cases and 32 ± 8 years for controls, and the average duration of PCOS for cases was 13 ± 10 years and for controls was 18 ± 9 years prior to their index dates. PCOS was significantly associated with

ovarian cancer risk (OR 1.55, 95% CI 1.10-2.20). Of the 276 women with PCOS in the study population, five cases and 46 controls (18.5%) had a recorded history of metformin use.

Interaction between groups defined by PCOS indication was observed (PCOS: OR 0.60, 95% CI 0.21-1.74; no PCOS: OR 1.03, 95% CI 0.92-1.15, $p_{\text{int}} = 0.01$). All models were adjusted for OC use, parity, and BMI.

Joint effects of metformin and concomitant medications

Evaluations of metformin use concomitantly with statins, sulfonylureas, insulin, beta blockers or cardiac glycosides showed no difference in ovarian cancer risk for women who used metformin alone (ORs 0.96-1.01) or used a second medication (ORs 0.68-1.14) (Table 12). Because sulfonylureas are antidiabetic agents often prescribed alongside metformin as a second-line option when metformin alone fails to control a patient's diabetes, this research project additionally evaluated the timing of sulfonylurea prescriptions among women who took both medications. Timing of metformin and sulfonylurea prescriptions, respective to each other, did not impact ovarian cancer risk compared to never-users of either drug (Appendix E, Supplemental Table 8).

Discussion

In this nested case-control study of over 100,000 women enrolled in the CPRD, ever-use of metformin was not associated with ovarian cancer, overall, by decade of birth, duration of use, or number of prescriptions. Metformin use was also not associated with ovarian cancer risk by categories of increasing cumulative DDDs, a measure that captured both dose and duration of use. High BMI prior to ovarian cancer diagnosis, a risk factor for metformin indications (Type 2 diabetes and PCOS), was associated with ovarian cancer risk, but Type 2 diabetes, the most

prevalent indication for metformin, was not. The association between metformin and ovarian cancer risk did not differ by presence or absence of type 2 diabetes. In a small analysis evaluating the effect of metformin use in women with and without PCOS, this study found that women with PCOS may have a higher risk of ovarian cancer, and that ovarian cancer risk may be lower for women with PCOS who took metformin. Finally, evaluations of joint interactions with medications commonly prescribed alongside metformin (beta blockers, cardiac glycosides, statins, sulfonylureas, and insulin) did not indicate associations with ovarian cancer risk. Overall, the findings from this analysis do not support a risk-reducing role for metformin in ovarian cancer.

This analysis includes the largest number of ovarian cancer cases evaluating metformin use with ovarian cancer risk. The results of this project are consistent with two large prospective cohort evaluations in the Finland Diabetes Database (FinDM) [36] and the Women's Health Initiative (WHI) [37], following over 100,000 women for 40 and 15 years, respectively, and observing no association between metformin use and ovarian cancer risk (FinDM: HR 1.02, 95% CI 0.72-1.45; WHI: HR 1.06, 95% CI 0.64-1.74). Results from the current analysis differed from three other observational studies from the German Targets and Essential Data for Commitment of Treatment (DETECT) study [38], the Taiwanese national health insurance database [39], and the General Practice Research Datalink (GPRD) [40], an early precursor to CPRD prior to the year 2010. These studies reported an over 30% reduction in risk of ovarian cancer for women who took metformin as well as reduced risks of ovarian cancer with increasing metformin prescription counts. The current research project improved upon some methodological limitations noted in earlier studies by including a long (30+ year) study window, subsequently capturing a large case number and group of participants who used metformin and well-suited to

address evaluations of metformin duration and cumulative prescription counts. An analysis of the association between metformin and ovarian cancer risk in 10-year birth cohorts allowed the current study to assess how ovarian cancer risk factors such as OC use and parity impacted the association over time. To further address potential etiologic pathways in ovarian cancer carcinogenesis, this project evaluated the associations between metformin and ovarian cancer risk according to metformin indications and medications commonly prescribed alongside metformin, two novel assessments previous studies did not pursue.

Results from this research project align with large observational studies also reporting null associations, and our results provide additional evidence that no association exists between metformin and ovarian cancer risk overall among women indicated for use. Although this study was not restricted to diabetics, like previous studies, and was able to evaluate effect modification in groups defined by indication (Type 2 diabetes and PCOS), almost all women who used metformin were diabetics, and no association between metformin and ovarian cancer risk was observed either among diabetics or non-diabetics. This study observed a reduced risk of ovarian cancer for metformin users among women with PCOS. The prevalence of PCOS in the study population was low ($< 0.4\%$ overall) and the number of women with PCOS who used metformin and also developed ovarian cancer was very low (5 women, 0.05% of cases). Given that the population prevalence of PCOS may lie between $0.4\text{--}2.3\%$ in US and Canadian populations similar to the UK [83], and the disease is often hard to diagnose [181], PCOS may be underreported in CPRD primary care data. However, because PCOS was diagnosed at an average age of 32-34 years in this analysis, well before a woman became a case or was selected as a control, the frequency of PCOS misclassification is likely to be similar by disease status. PCOS misclassification is therefore likely to be nondifferential, and bias the association between PCOS

and ovarian cancer risk toward the null. Similarly, the association between metformin and ovarian cancer risk would be biased toward the null in groups defined by PCOS status. A possible mechanism explaining the association observed among women with PCOS may relate to insulin sensitivity, as metformin reduces hyperinsulinemia observed in PCOS and insulin-like growth factors that activate tumorigenic cellular pathways [182].

This research study has several strengths. The current analysis was conducted in one of the largest primary care databases in the world with over 4.4 million active participants and more than 11 million total records. Over 9,000 cases were included, representing the largest set of ovarian cancers in any study of metformin and ovarian cancer risk to date. CPRD data are representative of the UK population allowing generalizability to similar populations. The captured data represented over 30-years of follow-up time, providing a large window to ascertain both metformin use and adequate time for ovarian cancer development. Important covariates were created for the study from Read codes that allowed for appropriately adjusted statistical models. The size of the CPRD allowed for the selection of a large case set and 10:1 matching of controls, a large analytic population for evaluation of many independent subgroups. The CPRD utilizes the Quality and Outcomes Framework (QoF) for improving and monitoring data quality [183]. The QoF is an initiative that started in 2004 to incentivize general practitioners to capture and report high-quality data to their electronic records. The National Health Service monitors the QoF participation by general practitioners and audits the framework regularly. Finally, linkage to the Pregnancy Register, a newly developed CPRD linkage, provided information on the total number of pregnancies for many women in the study, and is critical risk-reducing factor for ovarian cancer.

Although working with data collected via EMR systems has many strengths in capturing a wealth of patient data over long periods of time, limitations exist in the quality of data capture that are important to address in analyses using these data. This project noted challenges in the collection of important population characteristics in the CPRD primary care data. Histologic subtype information (serous, endometrioid, mucinous, and clear cell) is not available in the CPRD GOLD primary care data. It has been established that ovarian cancer subtypes are etiologically distinct diseases, with different risk factor patterns, survival probabilities, and clinical characteristics [10, 184]. This project was unable to determine whether metformin heterogeneously influenced ovarian cancer subtypes, but results fall in-line with large observational studies reporting null associations that also did not evaluate histology-specific associations. The outcome, exposures and covariates for this analysis were created using Read codes applied to the GOLD primary care database, and an underlying assumption of studies conducted in the CPRD is that if a Read code is absent, then the disease (or risk factor) is absent. While this may have led to some misclassification of metformin use or other covariates as unexposed, this misclassification is expected to be nondifferential by case or control status as metformin use as well as important ovarian cancer risk factors were captured in both cases and controls prior to their index dates. Laboratory data on glycosylated hemoglobin levels (a marker of diabetes and diabetes control) for study participants were not usable because laboratory test results are not comprehensively uploaded into the UK physician EMR systems after receipt. The null findings of the present study are not expected to be affected by lack of laboratory data to classify women by the severity of their Type 2 diabetes, because no association was observed between metformin use and ovarian cancer risk among all type 2 diabetics, or among women taking concomitant medications like sulfonylureas that could also signal more severe diabetes.

Vital statistics and anthropometrics are not routinely collected at UK general primary care practices, and a high level of missing data for some critical measures such as BMI were observed. This research project found no association between metformin and ovarian cancer risk among the subjects missing BMI values, and included a missing category for the BMI variable in adjusted models to ensure that all women in the analysis, regardless of missing BMI, contributed to the models. Finally, as all women in this analysis were indicated for metformin use, results of the present analysis are interpretable in the context of metformin indications rather than the general population of women at-risk for ovarian cancer.

In summary, this national-level analysis that was conducted in one of the largest primary care databases in the world found no association between metformin use and ovarian cancer risk overall. Diabetes status did not affect the association, nor drug characteristics like the number of metformin prescriptions or the duration of use. Among type 2 diabetics, metformin use was also not associated with ovarian cancer risk, but some limited evidence suggested that metformin may reduce ovarian cancer risk among women with PCOS. Results from this work have important implications for future researchers investigating the topic of metformin in possible ovarian cancer prevention by demonstrating that for the general population, there may be no preventive benefit. Future work should determine the association in ovarian cancer histotypes as well as consider data sources that can minimize the limitations observed by use of EMR-derived data.

Tables for manuscript 1

Table 7. Characteristics of ovarian cancer cases and controls in CPRD

Characteristic	Cases N=9,207	Controls N=92,070	P*
Birth year	1940 ± 13.7 (1900-	1940 ± 13.7 (1899-	0.73
Age at index (year)	64.7 ± 12.0 (40-90)	64.7 ± 12.0 (40-95)	0.69
OC use			
Ever	593 (6.4%)	7211 (7.8%)	< 0.0001
Never	8641 (93.8%)	84859 (92.1%)	
Family history of ovarian cancer			
Yes	37 (0.4%)	75 (0.08%)	< 0.0001
No	9170 (99.6%)	91995 (99.9%)	
Parity			
1 pregnancy	1082 (11.7%)	11022 (11.9%)	< 0.0001
2 pregnancies	905 (9.8%)	10168 (11.0%)	
3 pregnancies	504 (5.5%)	5710 (6.2%)	
4 pregnancies	233 (2.5%)	2688 (2.9%)	
4+ pregnancies	128 (1.4%)	2126 (2.3%)	
Missing	6355 (69.0%)	60365 (65.5%)	
BMI kg/m ² 5 years prior to			
<18.5=Low	92 (1%)	1024 (1.1%)	< 0.0001
18.5-24.99=Normal	2157 (23.4%)	23357 (25.4%)	
25-29.99=Overweight	1927 (21.0%)	18945 (20.6%)	
30+ =Obese	1487 (16.5%)	13212 (14.3%)	
Missing	3544 (38.5%)	35532 (38.6%)	
Smoking status			
Never	2169 (23.5%)	20996 (22.8%)	0.83
Former	4211 (45.7%)	41292 (44.8%)	
Current	1610 (17.4%)	15887 (17.3%)	
Missing	1217 (13.2%)	13895 (15.1%)	
Alcohol use			
Never	1244 (13.5%)	12468 (13.5%)	0.67
Former	653 (7.1%)	6296 (6.8%)	
Current	5298 (57.5%)	51755 (56.2%)	
Missing	2012 (21.9%)	21551 (23.4%)	

*Chi-square tests were used to determine differences between categorical variables, and t-tests were used to determine differences in mean values between cases and controls. p<0.05 was considered statistically significant

Table 8. Associations between metformin use and ovarian cancer risk in CPRD

	Cases N=9,207 n (%)	Controls N=92,070 n (%)	OR* (95% CI)
metformin use			
Ever	379 (4.1%)	3490 (3.8%)	1.01 (0.91-1.13)
Never	8828 (95.9%)	88580 (96.2%)	Ref
metformin prescriptions			
Range	1-321	1-656	
1 - 11	118 (1.3%)	894 (0.1%)	1.31 (1.05-1.64)
12 - 30	71 (0.1%)	892 (0.1%)	0.79 (0.60-1.04)
31 - 62	92 (0.1%)	845 (0.1%)	1.06 (0.83-1.36)
62+	98 (0.1%)	859 (0.1%)	1.12 (0.88-1.43)
Never use	8828 (95.9%)	88580 (96.2%)	Ref
metformin duration			
≤ 1 year (low use)	84 (0.9%)	698 (0.8%)	1.15 (0.92-1.45)
1-3 years	72 (0.8%)	792 (0.9%)	0.87 (0.68-1.11)
3-5 years	63 (0.7%)	590 (0.6%)	1.00 (0.76-1.30)
5-10 years	110 (1.2%)	925 (1.0%)	1.11 (0.91-2.36)
>10 years	50 (0.5%)	485 (0.5%)	0.97 (0.72-1.30)
Never use	8828 (95.9%)	88580 (96.2%)	Ref
metformin Defined Daily Doses			
1-240	102 (1.1%)	857 (0.9%)	1.13 (0.92-1.40)
241-749	94 (1.0%)	875 (1.0%)	1.01 (0.82-1.26)
750 -1802	82 (0.9%)	883 (1.0%)	0.86 (0.69-1.08)
1803+	101 (1.1%)	866 (0.9%)	1.07 (0.87-1.32)
Never use	8828 (95.9%)	88589 (96.2%)	Ref

Abbreviations: CPRD = Clinical Practice Research Datalink; OR = odds ratio

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Table 9. Association between metformin use and ovarian cancer risk in CPRD by decade of birth

	Cases N=9,207	Controls N=92,070	OR (95% CI)*	OR (95% CI)†	OR (95% CI)‡
YOB 1899-1910					
metformin					
Ever	1 (0.8%)	6 (0.5%)	-	-	-
Never	121 (99.2%)	1144 (99.5%)	-	-	-
Total	122	1150	-	-	-
YOB 1911-1920					
metformin					
Ever	7 (1.3%)	98 (1.9%)	1.00 (0.41-2.45)	1.00 (0.41-2.45)	-
Never	514 (98.7%)	5061 (98.1%)	Ref	Ref	-
Total	521	5159	-	-	-
YOB 1921-1930					
metformin					
Ever	61 (3.8%)	618 (3.8%)	1.10 (0.76-1.59)	1.10 (0.75-1.58)	-
Never	1555 (96.2%)	15516 (96.1%)	Ref	Ref	-
Total	1616	16134	-	-	-
YOB 1931-1940					
metformin					
Ever	125 (5.5%)	1180 (5.2%)	0.93 (0.75-1.24)	0.93 (0.71-1.24)	-
Never	2145 (94.5%)	21548 (94.8%)	Ref	Ref	-
Total	2270	22728	-	-	-
YOB 1941-1950					
metformin					
Ever	94 (3.8%)	960 (3.8%)	0.93 (0.66-1.30)	0.93 (0.66-1.30)	-
Never	2393 (96.2%)	24014 (96.2%)	Ref	Ref	-
Total	2487	24974	-	-	-
YOB 1951-1960					
metformin					
Ever	57 (3.9%)	453 (3.1%)	1.20 (0.74-1.96)	1.21 (0.75-1.98)	-
Never	1415 (96.1%)	14245 (96.9%)	Ref	Ref	-
Total	1472	14698	-	-	-

YOB 1961-1970					
metformin					
Ever	29 (4.7%)	144 (2.3%)	1.26 (0.66-2.39)	1.28 (0.67-2.45)	1.28 (0.67-2.43)
Never	582 (95.3%)	6035 (97.7%)	Ref	Ref	-
<i>Total</i>	<i>611</i>	<i>6179</i>		-	-
YOB 1971-1979					
metformin					
Ever	5 (4.6%)	31 (3.0%)	1.51 (0.43-5.25)	1.60 (0.45-5.69)	1.67 (0.47-5.96)
Never	103 (95.4%)	1017 (97.0%)	Ref	Ref	-
<i>Total</i>	<i>108</i>	<i>1048</i>		-	-

Abbreviations: YOB = year of birth; CPRD = Clinical Practice Research Datalink; OR = odds ratio; Ref = reference

*Adjusted for: OC use, BMI

†Adjusted for: OC use, BMI, parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

‡Adjusted for: OC use, BMI, parity (linear term, with missing/nulliparous = 0)

Table 10. Association between metformin and ovarian cancer risk by BMI categories in CPRD

	Cases N=9,207		Controls N=92,070		OR* (95% CI)	p _{int} [†]
	Non- users n (%)	Metformin users n (%)	Non- users n (%)	Metformin users n (%)		
BMI (kg/m²)						
< 18.5	254 (2.8%)	1 (0.3%)	3051 (3.4%)	22 (0.6%)	-	0.68
18.5 – 24.9	1938 (21.0%)	28 (7.4%)	20750 (23.4%)	350 (10.0%)	0.84 (0.57-1.24)	
25 – 29.9	1866 (20.3%)	88 (23.2%)	18063 (20.4%)	865 (24.8%)	0.97 (0.77-1.21)	
30+	1280 (13.9%)	207 (54.6%)	11561 (13.1%)	1806 (51.7%)	1.03 (0.88-1.20)	
Missing	3487 (42.0%)	55 (14.5%)	35155 (39.7%)	447 (12.8%)	1.24 (0.94-1.65)	
<i>Total</i>	<i>8828</i>	<i>379</i>	<i>88580</i>	<i>3490</i>		

Abbreviations: CPRD = Clinical Practice Research Datalink; BMI = body mass index; OR = odds ratio

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

values for interaction between groups were determined using a model with an interaction term between metformin and BMI categories

Table 11 Associations between metformin and ovarian cancer risk by indication group in CPRD

Cases N=9,207			Controls N=92,070			
	Non- users n(%)	Metformin users n(%)	Non- users n(%)	Metformin users n(%)	OR * (95% CI)	p _{int} [†]
Type 2 diabetes						
Ever	243 (2.8%)	367 (96.8%)	2530 (2.9%)	3377 (96.7%)	1.09 (0.92-1.30)	0.16
Never	8585 (97.2%)	12 (3.2%)	86050 (97.1%)	113 (3.2%)	1.07 (0.59-1.95)	
Total	8828	379	88580	3490		
PCOS						
Ever	32 (0.4%)	5 (1.3%)	193 (0.2%)	46 (1.3%)	0.60 (0.21-1.74)	0.01
Never	8796 (99.6%)	225 (98.7%)	88387 (99.8%)	3444 (98.7%)	1.03 (0.92-1.15)	
Total	8828	379	88580	3490		

Abbreviations: CPRD = Clinical Practice Research Datalink; PCOS = polycystic ovarian syndrome; OR = odds ratio

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

[†]p values for interaction between groups were determined using a model with an interaction term between metformin and the indication

Table 12 Joint effects of metformin use with other diabetes and cardiovascular disease medications in CPRD

	Cases N=9,207	Controls N=92,070	OR* (95% CI)
Beta-blocker			
--M / --B	8101 (88.0%)	81691 (88.7%)	Ref
+M / --B	343 (3.7%)	3193 (3.5%)	1.01 (0.90-1.14)
--M / +B	727 (7.9%)	6889 (7.5%)	1.08 (1.00-1.17)
+M / +B	36 (0.4%)	297 (0.3%)	1.14 (0.80-1.62)
Cardiac glycosides			
--M / --C	8632 (93.8%)	86850 (94.3%)	Ref
+M / --C	365 (4.0%)	3299 (3.6%)	1.04 (0.93-1.16)
--M / +C	196 (2.1%)	1730 (1.9%)	1.09 (0.94-1.27)
+M / +C	14 (0.2%)	191 (0.01%)	0.68 (0.39-1.27)
Statins			
--M / --St	7454 (81.0%)	75127 (81.6%)	Ref
+M / --St	82 (0.9%)	774 (0.8%)	0.99 (0.79-1.26)
--M / +St	1374 (14.9%)	13453 (14.6%)	0.99 (0.93-1.05)
+M / +St	297 (3.2%)	2716 (2.9%)	1.01 (0.89-1.15)
Sulfonylureas			
--M / --Su	8771 (95.3%)	87915 (95.5%)	Ref
+M / --Su	153 (1.7%)	1489 (1.6%)	0.96 (0.81-1.14)
--M / +Su	57 (0.6%)	665 (0.7%)	0.82 (0.63-1.08)
+M / +Su	226 (2.5%)	2001 (2.2%)	1.05 (0.91-1.21)
Insulin			
--M / --I	8777 (95.3%)	88129 (95.7%)	Ref
+M / --I	293 (3.2%)	2734 (3.0%)	1.00 (0.88-1.13)
--M / +I	51 (0.6%)	451 (0.5%)	1.10 (0.82-1.48)
+M / +I	86 (0.9%)	756 (0.8%)	1.06 (0.85-1.33)

Abbreviations: CPRD = Clinical Practice Research Datalink; OR = odds ratio

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Chapter 5: Manuscript 2

Statin therapy and risk of ovarian cancer in a national primary care database from the United Kingdom

Abstract

Introduction Epithelial ovarian cancer (EOC) is the 8th most common cancer among women worldwide, and the most fatal gynecologic malignancy. Metabolic factors, including obesity, diabetes, and circulating lipids such as cholesterol are associated with EOC. Statins, which lower cholesterol, have drawn interest as possible chemoprevention options for EOC, but only nine studies have been published to-date, with mixed results. Importantly, subgroup associations are lacking in published literature, including formulation, cumulative exposure, indication and pharmacokinetic action that are critical to evaluating possible benefit groups. Statin use in relation to risk of EOC was comprehensively evaluated in the Clinical Practice Research Datalink (CPRD).

Methods This nested case-control study was conducted in the CPRD, the primary care database of the United Kingdom (UK) that contains data collected prospectively from over 10 million residents and greater than 2,000 general practices. Cases were defined as women diagnosed with ovarian cancer between 1987–2020, aged greater than 40 years, had at least 2 years of CPRD data, and no prior cancer (except non-melanoma skin cancer) (n=9,207). Controls were incidence-density sampled and at-risk of ovarian cancer, had no prior history of any cancer except non-melanoma skin cancer, and matched to cases 10:1 on age, general practice of the case, and time in CPRD (n=92,070). Medication use was defined as having at least 1 prescription up to one year prior to date of diagnosis or control selection. Multivariable logistic regression, producing odds ratios (ORs) and 95% confidence intervals (CIs), was used to analyze

associations between statin use and ovarian cancer risk overall, according to medication use characteristics (prescription count, duration of use, type of statins), drug indications, formulations, and concomitant medication use.

Results Statin use was not associated with ovarian cancer risk overall (OR 1.02, 95% CI 0.96-1.08) or according to increasing prescription counts, higher duration of use, and use of concomitant medications (a proxy for disease severity). Statins were not associated with ovarian cancer in any BMI category (ORs 0.84-1.03, $p_{\text{int}} = 0.68$) and associations between statin use and ovarian cancer did not differ for primary indications ($p_{\text{int}} 0.16$ -0.69). Statin formulations additionally did not differ with respect to pharmacokinetic action, and there were no associations between individual statins or statin classes (lipo- or hydrophilic) with risk of ovarian cancer.

Conclusion In this comprehensive study using the largest series of ovarian cancer cases to-date, statin use was not associated with ovarian cancer risk overall, according to drug characteristics, indications for drug use or chemical formulation among women who were indicated for statins.

Impact Interest in statins as a chemopreventive agent for EOC has received considerable interest, but this population-based analysis did not identify a benefit group for statin use. Future studies should consider evaluation of ovarian cancer histotypes.

Introduction

Ovarian cancer is often diagnosed at a late stage due to lack of distinct symptomology and screening modalities [5] [6, 7]. Conditions such as obesity, diabetes and metabolic syndrome (MetS) increase the risk of ovarian cancer and are related to lipid abnormalities [15, 28]. Circulating levels of sphingomyelins, ceramides and cholesterol may also be associated with increased risk of ovarian cancer [154]. Repurposing existing pharmacologics designed to treat lipid-related disorders may provide primary prevention opportunities to reduce ovarian cancer incidence or give clues to biological pathways in carcinogenesis.

Statins are lipid-lowering agents indicated for the primary prevention of cardiovascular disease (CVD) among individuals with risk factors such as dyslipidemia, hypertension, diabetes and smoking [185], as well as for secondary prevention among those diagnosed with CVD [186]. Conditions that increase risk of CVD, such as diabetes, are also linked to increased risk of multiple cancer types, including ovarian. Several statin formulations exist, and all competitively inhibit 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase enzyme, blocking mevalonate formation in the liver, a key component in cholesterol synthesis [131]. Statins may also have pleiotropic effects such as reducing inflammation by HMG-CoA-independent pathways [187] [32, 131]. Importantly, lipid instabilities and inflammation are both suspected catalysts for carcinogenesis in many cancer types, leading to speculation that statins may be candidate drugs for repurposing as prevention agents for several cancer types, including gynecological cancers.

Recent evidence has suggested that statins may exhibit a protective effect on the development of ovarian cancer [43, 46-48], but few studies have analyzed these associations and accounted for important drug characteristics. Statins may act differently to influence ovarian cancer risk based on their solubility, as lipid-soluble lipophilic statins (simvastatin, atorvastatin, Fluvastatin) easily pass through cell membranes to reach the HMG-CoA target, but water-soluble hydrophilic statins (pravastatin and rosuvastatin) require active transport [188]. Importantly, to date, only three studies have evaluated statins by class, and only one study [42] reported statin-specific estimates of ovarian cancer risk [42]. A systematic review accounting for prevalence of individual statin use across included studies observed a reduced risk of ovarian cancer in populations where pravastatin use, a hydrophilic statin, was low [49], suggesting a need for additional patient-level analyses of statin use addressing both class and individual statin associations. Other factors, such as indication for use, dose or duration of use may also impact associations, but analyses incorporating subgroups are also sparse. This research project evaluated the association between statin use and ovarian cancer risk in the Clinical Practice Research Datalink (CPRD), a large primary care database in the United Kingdom (UK), addressing critical subgroups and drug characteristics. This dissertation analysis was approved by the Institutional Review Board (IRB) of the University of Maryland (UMD) (Appendix H)

Methods

This nested case-control study was conducted in the CPRD GOLD, a primary care database containing a representative sample of the UK population. Cases were identified as women diagnosed with ovarian cancer between January 1987 and July 2020, who were at least 40 years of age at diagnosis [161, 178], with at least two years of history in the primary care data and without a prior diagnosis of cancer except nonmelanoma skin cancer (Figure 9). Up to ten

controls per case were selected, via incidence density sampling without replacement, from women who were at risk of ovarian cancer (i.e., without history of bilateral oophorectomy prior to the diagnosis date of an incident case). Control subjects were required to have no prior diagnosis of any cancer except for non-melanoma skin cancer and were selected from individuals with at least two years of data history in the CPRD. Controls were frequency matched to cases 10:1 on year of birth (± 1 year), the general practice of the case, and the number of years of data in CPRD prior to the case's index date [135]. Data recording routines of general practices may vary, and an individual's length of CPRD data history influences the ascertainment of the exposure windows of interest, making it important to match on these criteria. CPRD anonymizes and protects all patient data and ensures that individuals providing data for research purposes are aware of their personal health information (PHI) rights. All patients in CPRD-participating practices are provided the option to opt-out of data-sharing with CPRD. The analytic population for the current analysis included 9,207 cases and 92,070 controls and the selection process for cases is detailed in Chapter 3, Figure 9. The frequency matching for controls is shown in Appendix B.

Outcome definition

Ovarian cancer cases were defined as any first-time occurrence of a malignant neoplasm or cancer of the ovary, fallopian tube, broad ligament, parametrium or uterine adnexa in the CPRD Gold between the years 1987-2020. The date of diagnosis of each case (index date) was defined as the date the ovarian cancer was recorded in CPRD. The outcome definition was assembled using Read codes for ovarian cancer in the CPRD primary care data and can be found in Appendix A.

Exposure definition

There were five statins available during the CPRD time window used in this analysis.: simvastatin, atorvastatin, fluvastatin, pravastatin and rosuvastatin. Statin use was analyzed as dichotomous (ever/never) with ever use was defined as having ≥ 1 prescription (use) for any of the five statins or 0 prescriptions (no statin use) recorded in their CPRD Gold history up to 1 year prior to the index date. Statin use was also analyzed by individual statin as dichotomous (ever versus never) for each of the five statins, and by statin class (ever versus never use of lipophilic: simvastatin, atorvastatin and fluvastatin or hydrophilic statins: pravastatin, rosuvastatin) [45]. Statin dose was also assessed by evaluating the number of prescriptions in quartiles based on the distribution of statin use in the controls ($Q1 \leq 10$, $Q2 \leq 28$, $Q3 \leq 59$, $Q4 > 60$). Statin duration, overall and by statin, was defined by the cumulative years a woman was exposed (≤ 1 year), 1-3 years, 3-5 years, 5-10 years and 10+ years) based on the distribution in the controls. Statin dose and duration was evaluated using cumulative Defined Daily Doses (DDD) defined by the World Health Organization (WHO) as the average maintenance dose per day for a drug [179]. The DDD recommendations for each statin range from 10 to 60 milligrams (mg) and are presented in Appendix F, Supplemental Table 1. DDDs for each prescription were calculated by creating a DDD conversion factor (WHO-assigned DDD/ strength for the prescription in milligrams) and multiplying this conversion factor by the number of tablets for that prescription. Cumulative DDDs were calculated by summing the total DDDs across statin prescriptions in a patient record overall and by individual statin.

Study covariates

The outcome, exposure and study variables that were tested as confounders or used in final adjusted models were assembled by Read codes, by an algorithm applied to the CPRD

primary care data or by linkage. Oral contraceptive (OC) use (ever versus never use), diabetes status (ever versus never), hypertension (ever versus never), dyslipidemia (ever versus never), heart disease (ever versus never), family history of ovarian cancer (ever versus never), smoking status (current, former, never) and alcohol use (current, former, never), cardiac glycosides (ever versus never), insulin (ever versus never), sulfonylurea use (ever versus never), beta blockers (ever versus never) and statins (ever versus never) were assembled by Read codes. BMI (kg/m^2) was determined by an algorithm applied to the CPRD data using height and weight measures, and the BMI measure at least 5 years prior to ovarian cancer diagnosis date (or five years prior to index date for controls) was determined for each woman. BMI was categorized as ($<18.5 \text{ kg}/\text{m}^2$ underweight, $18.5\text{-}24.9 \text{ kg}/\text{m}^2$ normal weight, $25\text{-}30 \text{ kg}/\text{m}^2$ overweight and $30+ \text{ kg}/\text{m}^2$ obese or missing). The number of pregnancies per woman was obtained via linkage to the CPRD Pregnancy Register which applies a complex algorithm to all women with primary care records to determine how many pregnancies she had recorded in CPRD. Parity was categorized as (1, 2, 3, 4+ and nulliparous/missing) [163].

All variables coded ever versus never had no missing values, as an underlying assumption in CPRD for most variables is that the absence of a Read code indicates a non-event. Other variables, such as parity, BMI, smoking status and alcohol use require a different inference. For example, a missing BMI value cannot be interpreted as a non-event, as all study participants are expected to have a BMI value across the lifespan. As such, the percent of missing information in controls for BMI was determined to be 40%, 66% missing parity information, 15% missing smoking status and 23% missing alcohol use.

Statistical methods

Univariate statistics comparing cases and controls were determined via chi-square tests for categorical variables and t-tests of mean differences between groups for continuous variables. Multivariable logistic regression was used to estimate odds ratios (ORs) and 95% confidence intervals (CIs) for the association between statin use and ovarian cancer risk overall, for individual statins (simvastatin, atorvastatin, pravastatin, fluvastatin and rosuvastatin) and by statin class. Statin intake on ovarian cancer risk was evaluated by the number of statin prescriptions, duration of statin use and DDDs. OC use, BMI, diabetes status, hypertension, dyslipidemia, heart disease, and smoking status were evaluated as potential confounders. Final models were determined by incorporating a priori ovarian cancer risk factors and factors associated with both ovarian cancer risk and statin use in the analytic population (OC use, parity and BMI). Age, general practice and time in CPRD were used as frequency matching factors in the design phase of the study and not included in final models. The primary effects of indications for statin use (hypertension, dyslipidemia and heart disease) as well as conditions related to indications (BMI, smoking, and Type 2 diabetes) with ovarian cancer risk were assessed, followed by the effect of statin use within indication and risk factor groups. Statistical heterogeneity in effect sizes between indication groups was determined by fitting a model with an interaction term between statin use and each indication. Indications for statins often coexist with other cardiovascular disease (CVD)-related comorbidities, which may be treated with concomitant medications. Thus, the joint effects of statin use and use of medications commonly prescribed for other CVDs and diabetes control (sulfonylureas, insulin, cardiac glycosides, beta blockers, and metformin) were evaluated by creating a categorical variable indicating the presence or absence of each medication: statins and a second medication (++), statins alone (+ -),

no statin but a different medication (- +), or neither medication (- -). The reference category for joint effect analyses was the category using neither medication (- -). Categorical analyses containing counts in any strata less than five were not modeled. Both statin use as well as intake of important adjustment factors, like OCs may differ for women across time, the association between statin use and ovarian cancer risk was evaluated by birth cohorts (1899-1910, 1911-1920, 1921-1930, 1931-1940, 1941-1950, 1951-1960, 1961-1970, 1971-1979). Because the parity variable obtained from the Pregnancy Register included a category that combined women missing parity information and nulliparous women which were not possible to separate, a sensitivity analysis was conducted in the youngest women born after 1960 whose missing pregnancy information was anticipated to align more strongly with nulliparity rather than missing values. Additional analyses were conducted dividing the analytic population by specific dates related to historical CPRD expansions (Value Added Research Products (VAMP) 1987-1993, General Practice Research Datalink (GPRD) 1993-2012, and CPRD 2012-present) as well as according to timepoints capturing the density of cases across the study period.

pharmacoepidemiologic literature specific to CPRD analyses of drug-cancer associations [180].

A p value less than 0.05 was used to determine significance in all tests of association. All analyses were performed in SAS statistical software version 9.4, (Cary, NC).

Results

Characteristics of study participants are presented in Table 13. The average birthyear for all participants was 1940 (range: 1899-1979). As age was a matching factor, ages at diagnosis date and index date for controls were 64.7 ± 12.0 years. All cases were diagnosed between 1987-2020, with 2-4% of the total case number diagnosed in each study year. (Appendix E, Supplemental Table 1). With regard to reproductive factors, cases were less likely than controls

to use OCs and were more likely to have a family history of ovarian cancer ($p_{\text{chisq}} < 0.01$). Controls were more likely to have multiple pregnancies ($p_{\text{chisq}} < 0.01$). With regard to behavioral and anthropometric factors, smoking and alcohol use were similar between cases and controls, but cases were more likely to be obese (BMI 30+ kg/m²).

Study characteristics by 10-year birth cohort

The characteristics of study participants were additionally evaluated by birth cohort and are summarized (Appendix E, Supplemental Table 2). Mean year at diagnosis was 1992 (1987-1999) in birth cohort 1899-1910 and 2016 (2011-2020) in birth cohort 1971-1979. With regards to reproductive factors, the use of OCs increased in later birth cohorts with the youngest women in our sample (birth years 1961-1979) having the highest use of OCs (34-42%). OC use was <16 % among women born between 1951-1960 and <5% among those born 1899-1950. Generally, OC use was higher in controls than in cases in all birth cohorts. Pregnancy events were identified rarely among women born between 1899-1960, with the frequency of the missing/nulliparous category ranging from 99% in the earliest birth cohort (1899-1910) to approximately 50% in women born between 1951-1960. For women born after 1960, missing/nulliparous pregnancy data reduced to approximately 30%, with higher frequencies of multiple pregnancies among controls. Family histories of ovarian cancers were recorded in < 1% of cases and controls in all birth cohorts. The proportion of behavioral factors, such as smoking and alcohol use, were balanced between cases and controls in all birth cohorts. However, the increased prevalence of obesity among cases compared to controls was evident among women born between 1930-1979. Missing data for behavioral and anthropometric factors decreased with later birth cohort (Appendix E, Supplemental Table 2).

Primary association between statin use and ovarian cancer risk

There were 1,671 (18.1%) cases and 16,169 controls who used statins in the study population (Table 14). No association was observed between ever-use of statins and ovarian cancer risk overall after adjustment for OC use, parity and BMI (OR 1.02, 95% CI 0.96-1.08) (Table 2). Across birth cohorts, statin use varied from 0% - 27 % among cases, and 0 - 26% among controls. In the 1951-1960 birth cohort, a higher risk of ovarian cancer was observed for statin users (OR 1.40, 95% CI 1.03-1.92). Statin use was not associated with ovarian cancer risk in any other birth cohort (Table 15). Additionally, there was no association between statin use and ovarian cancer risk in sensitivity analyses dividing the CPRD population according to dates corresponding to the density of case capture across study period or according to key dates in CPRD history (1987-1993, 1993-2012-2020) related to dataset updates and changes (Appendix F, Supplemental Tables 2-3).

Regarding individual statins and statin classes (Table 14), simvastatin and atorvastatin were used most frequently in the population (68% and 37% of all statin use, respectively). Pravastatin, rosuvastatin and Fluvastatin were less commonly used (9%, 10% and 2%, respectively). There was no association between individual statins and ovarian cancer risk. There were 552 (6.0%) cases and 5,207 (5.7%) controls (33% and 32% of all statin use) who used more than one statin. Ovarian cancer risk was not elevated for women who had ever used multiple statins compared to women who had never used statins (OR 1.01, 95% CI 0.92-1.11). There was no association between statin use and ovarian cancer risk by lipophilic or hydrophilic class (lipophilic: OR 1.00, 95% CI 0.88-1.13; hydrophilic: OR 1.02, 95% CI 0.96-1.08). Finally, there were 199 (2.2%) cases and 2,016 (2.2%) controls (11.9% and 12.4% of statin use), who used

more than one statin class. There was similarly no association between statin use and ovarian cancer risk in this group (OR 0.94, 95% CI 0.81-1.09).

Regarding assessments of cumulative statin use, there was no association between statin use and ovarian cancer risk in any quartile of cumulative prescription counts overall or by individual statin compared to never users of statins (ORs 0.95-1.03 for 1-10, 11-28, 29-59, or 60+ prescriptions). (Table 16, Appendix F, Supplemental Table 4). There was also no association between statin use and ovarian cancer risk according to duration of use, either for all statins combined for individual statins, or by statin class (ORs 1.00-1.07) (Table 16, Appendix F, Supplemental Table 5). Finally, there was no association between statins use in any DDD quartile with ovarian cancer risk overall (1-308, 309-938, 939-2203, 2203+) or by individual statin and statin class (Table 16, Appendix F, Tables 6-8).

Evaluation of statin use and ovarian cancer risk by indication groups

Three indications for statin use were evaluated to determine whether the effect of statin use differed according to the underlying conditions for prescription (hypertension, coronary heart disease and dyslipidemia). There were 2,834 (30.8%) cases and 27,837 (30.2%) controls with hypertension in the analytic population. Hypertension was not associated with ovarian cancer risk overall (OR 1.01, 95% CI 0.96-1.05) (Appendix F, Supplemental Table 9), and statin use was not associated with ovarian cancer risk among women with hypertension (OR 1.03, 95% CI 0.95-1.13), or those without (OR 0.98, 95% CI 0.89-1.08, $p_{\text{int}} = 0.59$). There were 665 (7.2%) cases and 6512 (7.1%) controls with coronary heart disease in the study population, and heart disease was not associated with ovarian cancer risk (OR 0.99, 95% CI 0.92-1.08). Statin use was not associated with ovarian cancer risk among women with a history of heart disease or those without (heart disease: OR 0.98, 95% CI 0.82-1.18; no heart disease: OR 1.01, 95% CI 0.94-

1.08, $p_{\text{int}} = 0.63$). Finally, there were 1,139 (12.4%) cases and 11,270 (12.2%) controls with a history of dyslipidemia. Dyslipidemia was not associated with ovarian cancer in the study population (OR 1.01, 95% CI 0.94-1.08), and statin use was not associated with ovarian cancer risk for women with dyslipidemia or those without (dyslipidemia: OR 1.09, 95% CI 0.95-1.25; no dyslipidemia: OR 0.99, 95% CI 0.92-1.07, $p_{\text{int}} = 0.52$) (Table 17). There was also no association between statin class and ovarian cancer risk for any of the three primary indications (Appendix F, Supplemental Table 10). Given the lack of association between individual statins and ovarian cancer risk in primary analyses, as well as the lack of association for each indication by statin class, individual statin analyses were not conducted for indications.

Evaluation of statin use and ovarian cancer risk by BMI, Type 2 diabetes and smoking status

Primary indications for statin use may coexist with conditions such as Type 2 diabetes, obesity or smoking. There were 610 (6.6%) cases and 5,907 (6.4%) controls with Type 2 diabetes in the analytic population (Table 17). Type 2 diabetes was not associated with ovarian cancer risk and statin use was not associated with ovarian cancer risk by Type 2 diabetes status overall. There was an association between hydrophilic statin use and ovarian cancer risk among women with Type 2 diabetes but not among those without Type 2 diabetes (Type 2 diabetes: OR 1.29, 95% CI 1.01-1.66; no Type 2 diabetes OR 0.92, 95% CI 0.79-1.07, $p_{\text{int}} = 0.04$) (Appendix F, Supplemental Table 11), but the associations between Type 2 diabetes and ovarian cancer risk among individual hydrophilic statins (pravastatin or rosuvastatin) were not significant ($p=0.20$ and $p=0.46$, respectively) (Appendix F, Supplemental Table 12). There was no association between lipophilic statins and ovarian cancer risk for either Type 2 diabetics or non-Type 2 diabetics (Type 2 diabetes: OR 0.98, 95% CI 0.82-1.18; no Type 2 diabetes OR 1.02, 95% CI 0.96-1.09) (Appendix F, Supplemental Table 11). There were 1,954 (21.2%) and 1,487 (16.2%)

of cases who were either overweight or obese in the study population, respectively, and there were 18,928 (20.6%) overweight or 13,367 (14.5%) obese controls. Analyses of the primary effect of BMI on ovarian cancer risk showed a higher risk for overweight and obese women 5 years or more before ovarian cancer diagnosis (overweight: 1.10, 95% CI 1.03-1.18; obese: OR 1.19, 95% CI 1.11-1.28). Statin use, however, was not associated with ovarian cancer risk in any BMI category. Finally, there were 5,821 (63.2%) cases and 57,179 (62.1%) controls who were either current or former smokers in the population. Current or former smoking was not associated with ovarian cancer risk compared to never smokers, and ovarian cancer risk did not differ by groups defined by smoking status (current, former, never) (Table 17). There was no association between hydrophilic or lipophilic statins and ovarian cancer risk by BMI category or smoking history (Appendix F, Supplemental Table 11).

Joint effects of statins and concomitant medications

Evaluations of statin use concomitantly with Beta-blockers, cardiac glycosides, metformin, sulfonylureas and insulin showed no difference in ovarian cancer risk for women who used statins alone or used a second medication. Odds ratios for all associations were not statistically significant and ranged from 0.96-1.26 (Table 18). An evaluation of concomitant medications by statin class also showed no associations between any concomitant medication use with ovarian cancer risk. For hydrophilic and lipophilic statins, odds ratios across all medications were not statistically significant and ranged from 0.92-1.29, respectively (Appendix F, Supplemental Tables 13-14).

Discussion

In this analysis of over 9,000 cases and 90,000 controls enrolled in the CPRD, no association between ever-use of statins and ovarian cancer risk overall or by decade of birth was observed. Additionally, no association with ovarian cancer risk was observed in additional analyses by type of statin, cumulative prescription counts, cumulative DDDs, or duration of use. Furthermore, the association between statins and ovarian cancer risk did not differ for women with hypertension, heart disease or dyslipidemia, three primary cardiovascular complications and indications for statin use or among women with other factors that may coexist with cardiovascular-related conditions (Type 2 diabetes, high BMI or smoking history). Hydrophilic statins, but not lipophilic statins may be associated with a higher risk of ovarian cancer risk among Type 2 diabetics. Overall, the findings from this research project do not provide evidence to support a risk-reducing role for statins in ovarian cancer.

This analysis includes the largest number of ovarian cancer cases of any study evaluating the effect of statin use on ovarian cancer risk. With a 30-year coverage, this study has incorporated a study period that captures statin use since they first became clinically available. The literature evaluating ovarian cancer risk after statin use is mixed. Results from the current evaluation were in-line with large observational studies reporting null results. These included a Danish nationwide case-control study using prescription records for statins with over 4,000 cases and almost 60,000 controls [45], a cohort study from the Finland diabetes database (FinDM) using a baseline population of almost 140,000 women (303 resulting cases) that utilized prescription records and included a follow-up of over 15 years [36], and an early case-control analysis in the GPRD with 91 ovarian cancer cases and 1,000 controls [48]. Four additional observational studies reported reduced but non-significant risks and were conducted in the Kaiser

Permanente and Group Health of Washington State insurance databases using prescription records, [43, 44], a small case-control study with 126 cases and controls from Israeli prescription records [46], and an incidental report from a small randomized controlled trial that was underpowered to assess cancer risk [41]. One cohort study, published in the Women's Health Initiative (WHI), reported an increased risk of ovarian cancer (HR 1.30, 95% CI 1.04-1.62) that they attributed to pravastatin use [42]. Finally, only one evaluation reported a significantly reduced risk of ovarian cancer after statin use, using data from The New England Case-Control (NECC), with over 2,000 cases and 4,000 controls (OR 0.65, 95% CI 0.52-0.82) [47].

Results from the current evaluation provide support that statin use may not be associated with ovarian cancer overall. However, the increased risk of ovarian cancer for the effect of hydrophilic, but not lipophilic statin use among Type 2 diabetics provides limited evidence that statins may not all act in a similar manner to influence risk, a hypothesis that has been previously suggested [49]. Hydrophilic statins require active transport across cellular membranes where they inhibit HMG CoA and downstream pathway products including mevalonate, cholesterol and protein isoprenylation enzymes [189]. This pathway is very active in liver cells, but is also present in many other tissues as it carries out cellular activities that include maintenance of cell membranes and functional activation of proteins. It has been suggested that pravastatin may be associated with an increased risk of multiple cancer types because it is unable to easily pass through cell membranes and block the synthesis of mevalonate in tissues outside the liver, which results in up-regulation of mevalonate pathway products in extra-hepatic cells. Increased cellular proliferation and tumorigenic activities in body tissues including the ovary may result. There was no effect of hydrophilic statin use on ovarian cancer risk in indication groups or groups defined by BMI or smoking status. Given this lack of consistency, it is possible that the association

between hydrophilic statin and ovarian cancer risk users among Type 2 diabetic patients is a chance finding.

This research study has several strengths. The current analysis was conducted in one of the largest primary care databases in the world with over 11 million patient records. Over 9,000 cases were included, representing the largest set of ovarian cancers in any study of statin use and ovarian cancer risk to date. CPRD data are representative of the UK population allowing generalizability to similar populations. The captured data represented over 30-years of data coverage, providing a large window to ascertain both statin use and adequate time for ovarian cancer development following the medication exposure windows. Critical covariates were created for the study from Read codes that allowed for multivariate adjustment and stratified analyses by indications and co-existing conditions. The size of the CPRD allowed for the evaluation of many independent subgroups. The CPRD utilizes the Quality and Outcomes Framework (QoF) for improving and monitoring data quality [183]. The QoF is an initiative that started in 2004 to incentivize general practitioners to capture and report high-quality data to their electronic records. The National Health Service monitors the QoF participation by general practitioners and audits the framework regularly.

Although working with data collected via EMR systems has many strengths in capturing prospective patient data over long periods of time, limitations exist in the quality of data capture that are important to address. It has been established that ovarian cancer subtypes are etiologically distinct, with different risk factor patterns and clinical characteristics [10, 184]. Unfortunately, histologic subtype information (serous, endometrioid, mucinous and clear cell) is not available in the CPRD Gold dataset. However, a large nation-wide evaluation of statin use and ovarian cancer risk in Denmark did not observe statistically significant associations between

statin use and ovarian cancer histotypes. The outcome, exposures and covariates for this analysis were created using Read codes applied to the Gold primary care database, with the assumption that if a Read code is absent, then the disease (or risk factor) is absent. While this may have led to some misclassification of statin use or other covariates as unexposed, this misclassification is expected to be nondifferential by case or control status. Laboratory data on serum cholesterol levels (a marker of CVD-related conditions) for study participants were missing in the majority of the population because laboratory test results are not comprehensively uploaded into the UK physician EMR systems after receipt. As a result, certain conditions, like dyslipidemia may not have been fully captured as cholesterol values are part of the definition. Any misclassification resulting from lack of laboratory data to define dyslipidemia or other laboratory-based variable definitions is expected to be nondifferential with respect to case status, and would bias effect estimates away from the null. Finally, as women in this analysis using statins had an indication for statin use, the results of this study are interpretable for women with underlying conditions, but not for women in the general population who are not indicated.

In summary, this national-level analysis of statin use and ovarian cancer risk was conducted in one of the largest primary care databases in the world and found no association between statin use and ovarian cancer risk overall. Indications for statin use did not affect the association, nor drug characteristics like the type of statin, number of statin prescriptions or the duration of use. Among Type 2 diabetics, hydrophilic statin use was associated with increased ovarian cancer risk, but more research is needed to better understand this finding. Results from this work have important implications for evaluating statin use for ovarian cancer prevention by demonstrating that for the general population, there may be no preventive benefit. Given the known heterogeneity of ovarian cancer, future should evaluate whether this association differs by

ovarian cancer histotypes among population-based studies whereby limitation of EMR-derived data may be minimized.

Tables for manuscript 2

Table 13. Characteristics of ovarian cancer cases and controls in CPRD

Characteristic	Cases n=9,207	Controls n=92,070	P*
Year of birth	1940 ± 13.7 (1900-	1940 ± 13.7 (1899-1979)	0.73
Age at index (year)	64.7 ± 12.0 (40-90)	64.7 ± 12.0 (40-95)	0.69
OC Use			
Ever	593 (6.4%)	7211 (7.8%)	< 0.0001
Never	8641 (93.8%)	84859 (92.1%)	
Parity			
1 pregnancy	1082 (11.7%)	11022 (11.9%)	< 0.0001
2 pregnancies	905 (9.8%)	10168 (11.0%)	
3 pregnancies	504 (5.5%)	5710 (6.2%)	
4 pregnancies	233 (2.5%)	2688 (2.9%)	
4+ pregnancies	128 (1.4%)	2126 (2.3%)	
Missing	6355 (69.0%)	60365 (65.5%)	
Family History of Ovarian			
Yes	37 (0.4%)	75 (0.08%)	< 0.0001
No	9170 (99.6%)	91995 (99.9%)	
BMI 5 years prior to			
<18.49=Low	92 (1%)	1024 (1.1%)	< 0.0001
18.5-24.99=Normal	2157 (23.4%)	23357 (25.4%)	
25-29.99=Overweight	1927 (21.0%)	18945 (20.6%)	
30+ =Obese	1487 (16.5%)	13212 (14.3%)	
Missing	3544 (38.5%)	35532 (38.6%)	
Smoking status			
Never	2169 (23.5%)	20996 (22.8%)	0.83
Former	4211 (45.7%)	41292 (44.8%)	
Current	1610 (17.4%)	15887 (17.3%)	
Missing	1217 (13.2%)	13895 (15.1%)	
Alcohol use			
Never	1244 (13.5%)	12468 (13.5%)	0.67
Former	653 (7.1%)	6296 (6.8%)	
Current	5298 (57.5%)	51755 (56.2%)	
Missing	2012 (21.9%)	21551 (23.4%)	

Abbreviations: OC = oral contraceptive; BMI = body mass index; CPRD = Clinical Practice Research Datalink

*Chi -square tests were used to determine differences between categorical variables, and t-tests were used to determine differences in mean values between cases and controls. p<0.05 was considered statistically significant

Table 14. Overall association between statin use and ovarian cancer risk in CPRD

	Cases N=9,207 n (%)	Controls N=92,070 n (%)	OR (95% CI)*
Statin use			
Ever	1671 (18.1%)	16169 (17.6%)	1.02 (0.96-1.08)
Never	7536 (81.9%)	75901 (82.4%)	Ref
Individual statins			
Simvastatin			
Ever	1160 (12.6%)	11354 (12.3%)	0.99 (0.93-1.06)
Never	8047 (87.4%)	80716 (87.7%)	Ref
Atorvastatin			
Ever	657 (7.1%)	6090 (6.6%)	1.05 (0.97-1.15)
Never	8550 (92.9%)	85980 (93.4%)	Ref
Pravastatin			
Ever	151 (1.6%)	1586 (1.7%)	0.93 (0.78-1.10)
Never	9056 (98.4%)	90484 (98.3%)	Ref
Fluvastatin			
Ever	33 (0.4%)	396 (0.4%)	0.81 (0.56-1.15)
Never	9174 (99.6%)	91674 (99.6%)	Ref
Rosuvastatin			
Ever	110 (1.2%)	956 (1.0%)	1.12 (0.91-1.37)
Never	9097 (98.9%)	91114 (99.0%)	Ref
Multiple statins			
2+ statins	552 (6.0%)	5207 (5.7%)	1.01 (0.92-1.11)
1 statin	1117 (12.1%)	10923 (11.9%)	0.99 (0.92-1.05)
Never	7538 (81.9%)	75940 (82.5%)	Ref
Statin class			
Hydrophilic only			
Ever	275 (3.0%)	2695 (2.9%)	1.00 (0.88-1.13)
Never	8932 (97.0%)	89375 (97.1%)	Ref
Lipophilic only			
Ever	1593 (17.3%)	15451 (16.8%)	1.02 (0.96-1.08)
Never	7614 (82.7%)	76619 (83.2%)	Ref
Lipo- and hydrophilic			
Ever	199 (2.2%)	2016 (2.2%)	0.94 (0.81-1.09)
Never	7538 (81.9%)	75940 (81.9%)	Ref

Abbreviations: CPRD = Clinical Practice Research Datalink; OR = odds ratio; ref = reference

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Table 15. Associations between statin use and ovarian cancer risk in CPRD by decade of birth

	Cases N=9,207	Controls N=92,070	OR (95% CI)*	OR (95% CI)†
YOB 1899-1910				
Statin				
Ever	0 (0%)	0 (0%)	-	-
Never	122 (100%)	1150 (100%)	-	-
<i>Total</i>	<i>122</i>	<i>1150</i>	-	-
YOB 1911-1920				
Statin				
Ever	21 (4.0%)	228 (4.4%)	0.91 (0.58-1.43)	-
Never	500 (96.0%)	4931 (95.6%)	Ref	-
<i>Total</i>	<i>521</i>	<i>5159</i>		-
YOB 1921-1930				
Statin				
Ever	331 (20.5%)	3101 (19.2%)	1.09 (0.96-1.24)	-
Never	1285 (79.5%)	13033 (80.8%)	Ref	-
<i>Total</i>	<i>1616</i>	<i>16134</i>		-
YOB 1931-1940				
Statin				
Ever	605 (26.7%)	6025 (26.4%)	1.01 (0.91-1.11)	-
Never	1665 (73.3%)	16703 (80.9%)	Ref	-
<i>Total</i>	<i>2270</i>	<i>22728</i>		-
YOB 1941-1950				
Statin				
Ever	480 (19.3%)	4782 (19.1%)	1.01 (0.91-1.12)	-
Never	2007 (80.7%)	20192 (80.9%)	Ref	-
<i>Total</i>	<i>2487</i>	<i>24974</i>		-
YOB 1951-1960				
Statin				
Ever	181 (12.3%)	1654 (11.3%)	1.09 (0.93-1.28)	-
Never	1291 (87.7%)	13044 (88.7%)	Ref	-
<i>Total</i>	<i>1472</i>	<i>14698</i>		-

YOB 1961-1970				
Statin				
Ever	48 (7.9%)	351 (5.7%)	1.40 (1.02-1.92)	1.40 (1.03-1.92)
Never	563 (92.1%)	5828 (94.3%)	Ref	Ref
<i>Total</i>	<i>611</i>	<i>6197</i>		
YOB 1971-1979				
Statin				
Ever	5 (4.6%)	23 (2.2%)	1.72 (0.64-4.63)	1.73 (0.64-4.63)
Never	103 (95.4%)	1020 (97.3%)	Ref	Ref
<i>Total</i>	<i>108</i>	<i>1048</i>		

Abbreviations: YOB = year of birth; CPRD = Clinical Practice Research Database; OR = odds ratio; ref = reference

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

†Adjusted for: OC use, parity (linear term, with missing/nulliparous = 0)

Table 16. Association between statin dose and duration with ovarian cancer risk in CPRD

	Cases n (%)	Controls n (%)	OR (95% CI)*
Number of statin prescriptions			
Range	1-410	1-658	
1 - 10	437 (4.7%)	4165 (4.5%)	1.02 (0.92-1.13)
11 - 28	424 (4.6%)	3974 (4.3%)	1.03 (0.93-1.14)
29 - 59	405 (4.4%)	4028 (4.4%)	0.96 (0.86-1.07)
60+	405 (4.4%)	4002 (4.3%)	0.95 (0.86-1.06)
Never	7536 (81.9%)	75901 (82.4%)	Ref
<i>Total</i>	<i>9207</i>	<i>92070</i>	
Statin duration			
≤ 1 year (low use)	338 (3.4%)	3243 (3.5%)	1.04 (0.93-1.17)
1-3 years	381 (4.1%)	3537 (3.8%)	1.07 (0.96-1.20)
3-5 years	284 (3.1%)	2774 (3.0%)	1.02 (0.90-1.15)
5-10 years	441 (4.8%)	4393 (4.8%)	1.00 (0.90-1.11)
>10 years	227 (2.5%)	2222 (2.4%)	1.02 (0.89-1.17)
Never	7536 (81.9%)	75901 (82.4%)	Ref
<i>Total</i>	<i>9207</i>	<i>92070</i>	
Defined Daily Doses			
1-308	427 (4.6%)	4068 (4.2%)	1.05 (0.94-1.16)
309-938	412 (4.5%)	4014 (4.4%)	1.02 (0.92-1.13)
939 -2203	418 (4.5%)	4062 (4.4%)	1.03 (0.93-1.14)
2203+	413 (4.5%)	4015 (4.4%)	1.03 (0.93-1.14)
Never	7537 (81.9%)	75911 (82.4%)	Ref

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Table 17. Associations between statin use and ovarian cancer risk by indication group in CPRD

	Cases N=9,207		Controls N=92,070		OR* (95% CI)*	P _{int} †
	Non-users n (%)	Statin users n (%)	Non-users n (%)	Statin users n (%)		
Hypertension						
Ever	1730 (23.0%)	1104 (66.1%)	17249 (22.7%)	10588 (65.5%)	1.03 (0.95-1.13)	0.59
Never	5806 (77.0%)	567 (33.9%)	58652 (77.3%)	5581 (34.5%)	0.98 (0.89-1.08)	
<i>Total</i>	<i>7536</i>	<i>1671</i>	<i>75901</i>	<i>16169</i>		
Heart Disease						
Ever	273 (3.6%)	392 (23.5%)	2680 (3.5%)	3832 (23.7%)	0.98 (0.82-1.18)	0.63
Never	7263 (96.0%)	1279 (76.5%)	73221 (96.5%)	12337 (76.3%)	1.01 (0.94-1.08)	
<i>Total</i>	<i>7536</i>	<i>1671</i>	<i>75901</i>	<i>16169</i>		
Dyslipidemia						
Ever	376 (5.0%)	763 (45.7%)	3869 (5.1%)	7401 (45.8%)	1.09 (0.95-1.25)	0.52
Never	7160 (95.0%)	908 (54.3%)	72032 (94.9%)	8768 (54.2%)	0.99 (0.92-1.07)	
<i>Total</i>	<i>7536</i>	<i>1671</i>	<i>75901</i>	<i>16169</i>		
BMI (kg/m²)						
< 18.5	206 (2.7%)	49 (2.9%)	2645 (3.5%)	428 (2.6%)	1.51 (1.06-2.18)	0.69
18.5 – 24.9	1644 (21.8%)	322 (19.3%)	17525 (23.1%)	3575 (22.1%)	0.95 (0.83-1.08)	
25 – 29.9	1428 (18.9%)	526 (31.5%)	13852 (18.3%)	5076 (31.4%)	0.98 (0.88-1.10)	
≥30	956 (12.7%)	531 (31.8%)	8696 (11.5%)	4671 (28.9%)	1.08 (0.97- 1.22)	
Missing	3302 (43.8%)	243 (14.5%)	33183 (43.7%)	2419 (15.0%)	0.95 (0.82-1.10)	
<i>Total</i>	<i>7536</i>	<i>1671</i>	<i>75901</i>	<i>16169</i>		
Type 2 Diabetes						
Ever	202 (2.7%)	408 (24.4%)	2013 (2.7%)	3894 (24.1%)	1.04 (0.86-1.26)	0.62
Never	7334 (97.3%)	1263 (75.6%)	73888 (97.3%)	12275 (75.9%)	1.00 (0.94-1.07)	
<i>Total</i>	<i>7536</i>	<i>1671</i>	<i>75901</i>	<i>16169</i>		

Smoking						
Current	1311 (17.4%)	299 (17.9%)	12930 (17.0%)	2957 (18.3%)	1.04 (0.91-1.20)	0.55
Former	3251 (43.1%)	960 (57.5%)	32140 (42.3%)	9152 (56.6%)	1.01 (0.93-1.09)	
Never	1772 (23.5%)	397 (23.8%)	17102 (22.5%)	3894 (24.1%)	0.96 (0.85-1.09)	
Missing	1202 (16.0%)	15 (0.1%)	13729 (18.1%)	166 (1.0%)	1.08 (0.63-1.85)	
<i>Total</i>	<i>7536</i>	<i>1671</i>	<i>75901</i>	<i>16169</i>		

Abbreviations: OR = odds ratio; BMI = body mass index; CPRD = clinical practice research datalink

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

† Interaction was determined by fitting a model between statin use and the drug indication or risk factor

Table 18. Joint effects of statin use with other CVD and diabetes medications in CPRD

	Cases N=9,207	Controls N=92,070	OR (95% CI)*
Beta-blocker			
--St / --B	6972 (75.7%)	70571 (76.6%)	Ref
+St / --B	1472 (16.0%)	14313 (15.5%)	0.99 (0.94-1.06)
--St / +B	564 (6.2%)	5330 (5.8%)	1.09 (1.00-1.20)
+St / +B	199 (2.2%)	1856 (2.0%)	1.05 (0.90-1.22)
Cardiac glycosides			
--St / --C	7395 (80.3%)	74673 (81.1%)	Ref
+St / --C	1602 (17.4%)	15476 (16.8%)	0.98 (0.94-1.06)
--St / +C	141 (1.5%)	1228 (1.3%)	1.11 (0.93-1.32)
+St / +C	69 (0.7%)	693 (0.8%)	0.94 (0.73-1.21)
metformin			
--St / --M	7454 (81.0%)	75127 (81.6%)	Ref
+St / --M	1374 (14.9%)	13453 (14.6%)	0.99 (0.93-1.05)
--St / +M	82 (0.9%)	774 (0.8%)	1.00 (0.79-1.26)
+St / +M	297 (3.2%)	2716 (2.9%)	1.01 (0.89-1.15)
Sulfonylureas			
--St / --Su	7445 (80.9%)	75007 (81.5%)	Ref
+St / --Su	1479 (16.1%)	14397 (15.6%)	0.99 (0.93-1.05)
--St / +Su	91 (1.0%)	894 (1.0%)	0.96 (0.76-1.20)
+St / +Su	192 (2.1%)	1772 (1.9%)	1.01 (0.87-1.18)
Insulin			
--St / --I	7489 (81.3%)	75540 (82.0%)	Ref
+St / --I	1581 (17.8%)	15323 (16.6%)	0.99 (0.94-1.06)
--St / +I	47 (0.5%)	361 (0.4%)	1.26 (0.93-1.72)
+St / +I	90 (1.0%)	846 (0.9%)	1.00 (0.80-1.24)

Abbreviations: OR = odds ratio; CPRD = clinical practice research datalink; CVD = cardiovascular disease

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Chapter 6: Manuscript 3

Sphingosine-1-phosphate (S1P) pathway protein expression in ovarian cancer tissues from the Polish Ovarian Cancer Study (POCS)

Abstract

Introduction Disrupted lipid signaling has been observed in many cancer types, including ovarian. Circulating levels of sphingomyelins, ceramides and cholesterol may be implicated with increased risk of ovarian cancer, but studies on lipid-related cellular pathway activity in ovarian tissue are scarce. This study evaluated expression of two primary receptors of the pro-tumorigenic lipid sphingosine-1-phosphate (S1P) to determine if expression was associated with tumor characteristics or ovarian cancer risk factors.

Methods Analyses were conducted among cases in the Polish Ovarian Cancer Study (POCS) (n=166). Expression of S1P receptors 1 and 3 (S1PRS 1,3) were examined using immunohistochemical (IHC) staining on tissue micro arrays (TMAs). Overall marker scores were determined by the product of area and intensity. Analysis of variance (ANOVA) was used to evaluate expression differences across ovarian cancer histotypes. Distributions of the staining score were used to determine dichotomized marker expression based on quartiles (S1PR1: 1 = higher quartile expression (Qs 2-4) versus 0 = lower quartile expression (Q1); S1PR3: 1= above median expression versus 0 = below median expression). Multivariable logistic regression was used to determine odds ratios (ORs) and 95% confidence intervals (CIs) for the association between tumor characteristics and ovarian cancer risk factors with marker expression status. Kaplan-Meier plots were used to examine differences in survival time according to marker expression status.

Results Expression of S1PR1 and S1PR3 differed across histotype of ovarian cancer ($p < 0.01$).

Most women with high expression of S1PR1 or S1PR3 had either moderately or poorly differentiated disease (72%). S1PR1 expression was elevated with higher disease grade, although not statistically significant (S1PR1: moderately differentiated OR 1.33, 95% CI 0.51-3.50; poorly differentiated OR 2.49, 95% CI 0.73-8.47, $p_{\text{trend}} = 0.13$). Similarly, higher disease grade was associated with higher marker expression for S1PR3 compared to well-differentiated disease (moderately differentiated: OR 1.12, 95% CI 0.27-4.59; poorly differentiated OR: 1.56, 95% CI 0.32-7.60, $p_{\text{trend}} = 0.04$). No ovarian cancer risk factors were associated with expression of either marker. Survival differences were observed by marker status, but neither S1PR1 nor S1PR3 were independent predictors of survival after accounting for cancer histotype.

Conclusions S1P receptor expression differs by ovarian cancer histotype. Higher expression of both S1PR1 and S1PR3 may be non-significantly associated with higher disease grade and stage, but not with ovarian cancer risk factors.

Impact The strong expression of lipid markers in ovarian cancer tissue suggests these pathways are involved in carcinogenesis; additional molecular studies are needed to elucidate their role further.

Introduction

Ovarian cancer is the most fatal gynecologic malignancy, with over 300,000 deaths worldwide in 2020 [2]. Overall, ovarian cancer five-year survival is 47%, but only 30% survive to five years after diagnosis with late-stage disease [184]. Survival probabilities [190], along with risk factor associations [10] and molecular underpinnings [191] differ according to etiologically distinct histotypes (serous, endometrioid, mucinous and clear cell) [8]. The complexity of ovarian cancers has contributed to the deadly nature of the disease, and biochemical pathways influencing pathogenesis and prognosis of each histotype remain poorly understood. The potential influence of lipid metabolism in ovarian cancer is one area of increasing interest [154].

Epidemiologic evaluations have identified several metabolic and hormonal conditions linked to increased ovarian cancer risk. Obesity [15] and obesity-related conditions such as diabetes [71] and metabolic syndrome (MetS) [17] are associated with lipid abnormalities and are risk factors for some histotypes. Medications used to treat metabolic conditions, such as metformin and statin drugs, have been hypothesized to reduce ovarian cancer risk [40, 47], and associations may be heterogenous by histotype [45, 47]. Prospective cohort studies also suggest that circulating lipids, such as cholesterol, phospholipids, and sphingomyelins are associated with increased histotype-specific ovarian cancer risk [111, 154, 192].

Sphingosine-1-phosphate (S1P) is a bioactive lipid that promotes tumorigenic effects such as angiogenesis, cell survival, and cellular proliferation [155]. As such, increased circulating levels of S1P are also suggested to promote tumor progression, for which angiogenesis and increased cellular proliferation are required [151, 193]. In ovarian cancer, early S1P precursors, like sphingomyelins and ceramides, are elevated in the serum of women

with ovarian cancer compared to women without ovarian cancers [150, 153]. Elevated levels of S1P have also been linked with ovarian cancer risk factors such as obesity [194]. How S1P acts at the cellular level, however, is not well understood.

It has been demonstrated that S1P exerts its effects through three primary g-protein coupled plasma membrane receptors (S1PRs1-3) [155]. Two of these receptors, S1PR1 and S1PR3 may be associated with angiogenesis activity in ovarian cancer cell lines through VEGF, IL-8 and IL-6 signaling [151], which correlates with tumor progression [195]. S1PRs 1 and 3 have been shown to promote cell migration, also in ovarian cancer cell lines [196] by modulation of PI3K and p38 kinases. S1PR1 has been linked to the accumulation of Tregs through the STAT3 signaling pathway [197], contributing to tumor evasion of a host immune response. Finally, in addition to angiogenesis, S1PR3 may promote cell migration via phospholipase C, as well as activating ERK and AKT pathway signaling that may promote cell survival [198].

Many gaps remain in understanding the activity of S1PR1 and S1PR3 and their relationships with cancer characteristics, particularly at the tissue level. Tissue-specific expression of pathway markers allows for pathway activity to be determined directly, as opposed to the evaluation of circulating markers in serum that are of uncertain origin and may not directly relate to a tumor. To date, only one study evaluated tissue expression of S1PR1 and S1PR3 in a small sample, comparing expression to normal ovarian cancer samples, but did not consider histotypes, risk factors or tumor characteristics such as stage or grade [151]. This is important, as S1PR1 and S1PR3 are hypothesized to be associated with factors that influence tumor progression [151]. To address these gaps, an exploratory study was conducted to examine tissue expression of two S1P pathway markers (S1PR1 and 3) in ovarian cancer cases from the Polish

Ovarian Cancer Study (POCS), and marker relationships were examined in relation to cancer histotype, risk factors and cancer survival.

Methods

Study population

Data for the current analysis were available from POCS, a population-based case-control study of women from Warsaw and Lodz, Poland diagnosed with incident ovarian cancer between June 2001 December 2003. Women aged 20-74 at the time of ovarian cancer diagnosis were assembled via a rapid identification system from five Polish hospitals (the Cancer Center, M. Curie Skłodowska Institute of Oncology, Polish Oncological Foundation, the Dr. Madurowicz Memorial Hospital and the Polish Mother's Health Memorial Hospital) and the Cancer Registries in both Warsaw and Lodz between the study start and end dates (n=317). Each woman had histologically-confirmed borderline or invasive ovarian cancer, and diagnoses were pathologically confirmed by hematoxylin & eosin (H&E) slide review. Reproductive and anthropometric covariate information for each woman was collected by interviewer-administered questionnaires, with a median time to interview after diagnosis of 2.4 months [167, 168]. Additional pathology and survival information was obtained from medical records. Tissue specimens were available for a subset (n=175, 55%) of Polish women enrolled in POCS. Six tumor cores were not able to be scored due to lack of sufficient cells remaining on the TMA, and three women did not have covariate information for oral contraceptive use or MHT leaving n=166 women in the analytic population. The study was originally approved by IRBs of the National Cancer Institute (NCI) as well as independent IRBs from the M. Skłodowska Curie Institute of Oncology and Cancer Center in Warsaw and the Institute of Occupational Medicine

in Lodz, Poland. Each woman provided written informed consent for both the use of interview collected data and archival ovarian cancer tissue [167]. This research study was reviewed and approved by the Institutional Review Board (IRB) of the University of Maryland (IRB).

Tissue specimen collection and creation of tissue micro arrays (TMAs)

Specimens were collected from routine medical procedures and stored in formalin fixed paraffin embedded (FFPE) tissue blocks at the respective hospitals where surgeries were performed, or from the cancer registries of Warsaw and Lodz. Slides cut from each block and stained with H&E were evaluated by a pathologist to identify representative regions of each tumor specimen. Between two to four tissue cores were punched from representative tumor regions in the FFPE whole tumor blocks and transferred to recipient blocks for creation of TMAs. Two TMA blocks containing duplicate cores for each woman were created and sections were cut from each block for IHC staining.

Pilot evaluation for S1P pathway marker antibodies

Two S1P pathway markers were chosen for evaluation in the present study based on proven staining potential in ovarian cancer tissue (S1PRs 1 and 3) [151]. To optimize these stains for ovarian cancer tissue, this research project conducted a comprehensive pilot evaluation prior to staining the POCS study samples to determine the optimal staining protocol for each marker. Antibodies were identified for each marker, with a priority for monoclonal antibodies (Appendix G, Supplemental Table 1). Each antibody was first applied to multiple human controls using the recommended dilutions from the manufacturer. Following a validation of the antibodies on control tissues to determine that each antibody stained as described by the manufacturer, each antibody was applied to a test TMA (OriGene, CT565862) containing ovarian tumor samples

representing the four major histotypes. Antibody dilutions established on the initial control tissues were applied to the test TMA and optimized to ensure the most specificity to target proteins while minimizing background staining. Final antibodies used to stain POCS samples were rabbit monoclonal anti-S1P1 (Thermo Scientific, Pittsburgh, PA, USA, clone JM10-66, at 1:1000), and rabbit monoclonal anti-S1PR3 (abcam, Waltham, MA, USA, clone EPR4541, at 1:75). Tumor cores on each test TMA were evaluated by a pathologist to determine the percentage of stained cells, intensity of stained cells and an overall score.

Processing of TMAs and scoring

POCS study samples were stained for each of the S1P markers at optimized dilutions. Following staining, TMA images were de-arrayed to prepare individual cores for scoring by the pathologist (PL). TMA cores were scored by determining the percentage of stained cells (0-100%), intensity (0=negative, 1=weak, 2=moderate and 3=strong) and the product of area (percent of tumor cells stained positive for the antibody) and intensity (0-300) for each marker. Because cut points for marker expression have not been established for S1PR1 and S1PR3, the percentage of stained cells and the overall staining score were evaluated as continuous and also dichotomized according to the distributions (0 = Q1 lowest quartile expression 1 = quartiles 2-4 highest quartiles expression) for S1P1, and median split (0= below median expression, 1 = above median expression) for S1PR3. Because multiple cores per woman were included on each TMA, Spearman's rank correlation statistic was used to determine the correlations between cores on the same TMA ($\rho > 0.80$). The core with the highest percentage of stained cells was used in all analyses based on a previous analysis [52].

Covariates

The following tumor characteristics were utilized in analyses for both markers: histology (1= serous, 2= endometrioid, 3= mucinous, 4= clear cell, 5 = mixed epithelial and 6 = other), grade (1 = well differentiated, 2= moderate or poor differentiation), and stage (1 = Figo stage 1, 2 = Figo stage 2+). Risk factor covariates were also assembled for age at diagnosis (1 = <50 years, 2 = 50-54 years, 3 = 55-64 years, 4 = 65+ years), menopause status (0 = pre-menopausal, 1 = post-menopausal and 2 = missing), menopausal hormone therapy (MHT) (0= never, 1 = ever), oral contraceptive use (0 = no, 1 = yes), parity (0 = nulliparous, 1 = parous), smoking status (0 = never, 1 = former, 2 = current) and BMI kg/m² measured in the current decade (1 = <25 kg/m² , 2 = 25-30 kg/m², 3 = 30+ kg/m²). Because only two women did not have information on oral contraceptive use, and one woman was missing MHT, no missing categories were included for these two variables.

Statistical analysis

Distributions for each categorical variable were reviewed and the frequency of each covariate level determined using the total study population as the denominator. Correlations between markers were assessed using Spearman's rank statistic using the total percentage of positively stained cells, overall score for each marker, and intensity of staining. Non-parametric Kruskal-Wallis tests were used to assess differences in continuous marker expression, as percentage of stained cells, intensity and overall scores, across histotype groups. Associations between tumor characteristics and ovarian cancer risk factors were assessed for the dichotomous marker expression based on the overall score distribution for each marker. The range of

positively stained cells for S1PR3 was low, and so while a quartile analysis of expression score was possible for S1PR1, a median split of the expression score was used for S1PR3 to evaluate high or low expression status. Final outcome dichotomizations for analyses with tumor characteristics and risk factors were: S1PR1 Qs 2-4 high expression vs Q1 low expression, and S1PR3 > 200 high expression and < 200 low expression. Non-serous tumors (endometrioid, mucinous and clear cell) were grouped for analyses with marker expression status as less than five women were observed in dichotomized score categories. Multivariable logistic regression was used to determine odds ratios (ORs) and 95% confidence intervals (CIs) for the association between ovarian cancer risk factors and dichotomized expression status. Final models were adjusted for all risk factors (described in the covariate section) based on a previously published analysis of cell cycle markers in POCS [52].

Survival information was available from medical records for all cases in the present analysis. Time from interview to date of death or last follow-up was used to evaluate differences in survival probabilities between higher and lower marker expression groups. For analyses assessing the association between disease grade and marker expression, the trend across increasing categories of higher grade was assessed using a linear term for increasing category of grade in the model. The times between interview after ovarian cancer surgery and death from any cause, or last date of follow-up, were compared by marker expression status using Kaplan-Meier curves and Log-rank test for each marker overall. All analyses were conducted using SAS Software package version 9.4 (SAS Institute, Cary, NC USA) and GraphPad Prism.

Results

Characteristics of the study population

Characteristics of the study participants are shown in Table 19. The mean age at diagnosis of ovarian cancer was 56 ± 10.6 years, and the median survival time was 6.9 ± 3.5 years. More than half of cases were postmenopausal $n=94$ (57%). OC use, an established preventive factor for ovarian cancer, was low in the population $n=11$ (7%), while nulliparity, a risk-increasing factor, was moderate $n=21$ (13%). There were 78 women with BMI measures in the last decade who were in the overweight or obese categories ($BMI > 25 \text{ kg/m}^2$), and most women were either former or current smokers $n=100$ (60%). There were $n=70$ (42%) serous and $n=33$ (20%) non-serous ($n=19$ endometrioid, $n=7$ mucinous and $n=7$ clear cell) tumors on study TMAs. Mixed epithelial tumors and other histotypes $n=63$ (38%) comprised the remaining case set.

Descriptive characteristics of S1PR1 and S1PR3 marker expression overall and by histotype

The overall distributions of IHC measures for area, intensity and marker score are presented in Appendix G, Supplemental Table 2. S1PR1 showed a range for area and intensity, which were reflected in the overall score (median 180, IQR 90-200). Most (>90%) of tumor samples were determined to have 100% of cells stained positive for S1PR3, and the area range for this marker was low (median 100, IQR 100-100). However, intensity varied across samples (median 2, IQR 1-3) that was also reflected in overall score (median 200, IQR 100-200).

Figure 18 presents the distributions of each marker across the four primary ovarian cancer histotypes according to the percent of cells stained positive, intensity of staining and the staining score. S1PR1 marker expression was heterogenous across histotypes, with median

marker expression for S1PR1 for all three IHC measures differing significantly ($p < 0.001$). Overall S1PR1 expression scores for endometrioid and clear cell tumors were similar with respect to medians and interquartile ranges (endometrioid: median 90, IQR 30-190; clear cell median 80, IQR 50-190). The area measure for S1PR3 ($p=0.83$) did not differ by histotype. S1PR3 intensity measures, however, were different across histotypes ($p = 0.004$), as well as the overall score ($p < 0.006$). As with S1PR1, endometrioid and clear cell tumors appeared similar with regard to score (endometrioid: median 100, IQR 100-200; clear cell: median 100, IQR 100-200).

Correlations were assessed between S1PR1 and S1PR3 expression (Appendix G, Supplemental Table 4). Overall, receptors were significantly correlated with respect to both intensity of staining and score ($p < 0.01$). Within histotype, intensity of staining and score were correlated for both serous and mucinous tumors ($p < 0.05$).

Associations between S1PR1 and S1PR3 with ovarian cancer tumor characteristics

The ovarian cancer tumor characteristics stage, grade and histotype were examined for relationships according to higher or lower expression of S1PR1 and S1PR3 (Table 20). There were $n=46$ (28%) women who fell into the lowest quartile expression category for S1PR1 determined by the score distribution (quartile 1). The remaining $n=120$ (72%) women were classified as having higher quartile expression (quartiles 2-4). Of the 70 women with serous tumors in this analysis, $n=60$ (86%) had expression in quartiles 2-4 for S1PR1. With regard to tumor histology, women with non-serous tumors were less likely to have higher quartile expression of S1PR1 compared to women with serous tumors (OR 0.16, 95% CI 0.06-0.48). Similarly, mixed adenocarcinomas (OR 0.23, 95% CI 0.08-0.66) and a group of women with

other histotypes (OR 0.46, 95% CI 0.12-1.73) also were less likely to have higher quartile expression of S1PR1, but confidence intervals were wide and indicated uncertainty in the estimate. A comparison of non-serous to serous tumors with regard to higher quartile expression of S1PR3 was not possible given that only n=2 (6.7%) of the 30 women with non-serous types had higher quartile expression of S1PR3. However, mixed adenocarcinomas (OR 0.56, 95% CI 0.14-2.26) and other histotypes (OR 0.65, 95% CI 0.17-2.54) were less likely to have high expression of S1PR3, albeit not statistically-significant (Table 20).

Generally, higher stage or grade of disease was not significantly associated with higher quartile S1PR1 expression. For women with stage II or III disease compared to women with stage 1 disease, the higher quartile S1PR1 expression was non-significantly elevated (OR 1.55, 95% CI 0.71-3.36). Most women with higher quartile expression of S1PR1 had either moderately or poorly differentiated disease (72%). Compared to women with well-differentiated disease, the odds of higher quartile expression of S1PR1 for women with moderately differentiated disease were non-significantly elevated (OR 1.33, 95% CI 0.51-3.50) and further increased for women with poorly differentiated disease (OR 2.49, 95% CI 0.73-8.47). However, the trend across increasing differentiation status was not significant ($p_{\text{trend}}=0.13$).

Higher stage or grade of disease was also associated with expression above the median for S1PR3. There were n=25 (15%) women above the median S1PR3 expression score, and n=41 (85%) women below the median. For women with stage II or III disease compared to women with stage I, the odds of S1PR3 expression above the median was non-significantly elevated (OR 1.51, 95% CI 0.51-4.42). While estimates for moderately (OR 1.12, 95% CI 0.27-4.59) or poorly (OR 1.56, 95% CI 0.32-7.60) differentiated disease suggested elevated expression and were not statistically significant, the trend across increasing quartiles was significant ($p_{\text{trend}}=0.04$). Neither

increasing intensity category nor increasing score quartile were predictive of grade or stage in sensitivity analyses (Appendix G, Supplemental Table 5).

Associations between ovarian cancer risk factors and expression of S1PR1 and S1PR3

Table 21 shows the associations between ovarian cancer risk factors and higher quartile versus lower quartile expression of S1PR1 and above or below median expression for S1PR3. Overall, no ovarian cancer risk factors were significantly associated with high expression of either marker. Higher age at diagnosis (50-64 years, 55-64 years and 65+ years) was not associated with lower quartile expression of S1PR1 compared to women under 50 years (S1PR1: ORs 0.27-0.79; S1PR3: ORs 0.56-0.79). Similarly, menopause status at the time of diagnosis, which is positively correlated with age, was not associated with higher quartile S1PR1 expression or S1PR3 expression above the median. Neither overweight status (OR 0.66, 95% CI 0.28-1.49) or obese status (OR 0.78, 95% CI 0.31-1.93) was associated with high marker expression. Although most women with higher expression of either marker were either current or former smokers (S1PR1 58%, S1PR3 56%), smoking status was not associated with S1PR1 higher quartile expression or S1PR3 expression above the median. Finally, no association was observed between parity and expression status of either marker, and OC use was not evaluable due to low counts.

Evaluation of survival by S1PR1 and S1PR3 expression status

The median follow-up time from date of interview to the date of last recorded mortality status or death was 6.9 years (range 0-11.8 years) with n=89 (53.6%) observed deaths during the follow-up time. Kaplan-Meier plots of the overall survival by marker expression status are shown in Figure 19. Overall, no statistically significant difference in survival was observed for

S1PR1 (log-rank $p = 0.16$). A borderline significant difference in overall survival was observed for S1P3 (log-rank $p=0.06$). However, when restricting the study population to high grade serous tumors ($n=61$), the most aggressive histotype, no difference in survival was observed for either marker (S1PR1 log-rank $p = 0.42$ and S1PR3 log-rank $p=0.84$).

Discussion

This research project is the first to examine the association between expression of two primary S1P receptors and ovarian cancer histotype, tumor stage and grade, and according to ovarian cancer risk factors. While expression of both S1PR1 and S1PR3 was heterogenous across the four primary ovarian cancer histotypes, endometrioid and clear cell tumors had similar distributions for overall marker scores. However, the range of expression for S1PR3 was limited. Higher grade or stage of ovarian cancer was non-significantly associated with expression of S1PR1 in quartiles 2-4, and with S1PR3 expression above the median in ovarian tumor tissue. Non-serous tumors were less likely to express both markers. No ovarian cancer risk factors were significantly associated with expression of either S1PR1 or S1PR3. Although survival according to expression status was lower for women with higher expression of either marker, neither S1PR1 nor S1PR3 were independent predictors of survival after accounting for ovarian cancer histology.

A substantial body of literature has suggested a role for S1P signaling in cancer [55, 150, 155, 193, 199, 200], leading to intense interest in the activity of its associated cellular pathway [54, 149, 151, 201-203]. A depiction of the pathway is presented in Figure 20. Two kinases catalyze the formation of S1P, two enzymes degrade it, and five membrane receptors bind S1P and stimulate downstream angiogenesis, cell migration, cell growth and anti-apoptotic effects [201]. Intracellularly, sphingosine-1-kinase (SPHK1) and sphingosine-2-kinase (SPHK2) generate S1P from the lipid ceramide, while S1P-lyase and S1P phosphatase degrade it. Inside cells, S1P may stimulate inflammation-related pathways via nucleotide-binding oligomerization domain-containing protein 1/2 (NOD1/2), which are intracellular receptors that can activate an innate immune response when stimulated [204]. S1P may also be transported outside cells to

bind one of five G-coupled sphingosine-1-phosphate receptors (S1PR1-5). S1PRs are located in the cell membrane, and after binding S1P, activate expression various pro-tumorigenic pathways [150]. S1P receptor binding may occur on the same cell in which it was generated, or neighboring cells. This unique property of S1P to act intracellularly after generation or extracellularly in both an autocrine and paracrine manner (“inside-out signaling”) allows S1P to affect the diverse array of cellular processes [155, 203].

In the present analysis of S1PR1 and S1PR3, IHC staining was strong overall, and was significantly heterogenous by histotype for both markers. Results indicated that increased tumor stage and grade may be associated with increased expression of both S1PR1 and S1PR3. A previous study [186] evaluating the activity of S1P receptors in ovarian cancer demonstrated IHC staining on ovarian cancer tissues, finding higher expression of all three S1P receptors (S1PRs 1-3) compared to normal ovarian tissue, giving evidence that receptor expression is enhanced in ovarian cancer. This study also found that, *in vitro*, S1PR1 and S1PR3 receptors stimulated angiogenic factor secretion [151], a necessary component of ovarian tumor growth and metastasis [205, 206]. Interestingly, no associations between any ovarian cancer risk factors and differences in marker expression were observed. Circulating S1P has been linked to obesity [194], but no associations were observed between higher BMI and expression of either S1PR1 or S1PR3 in this analysis. Thus, expression of these specific receptors did not provide insight as to how S1P may be linked to weight status.

In other cancer types, links between S1PR1 and S1PR3 activity with tumor progression have been hypothesized [150]. S1PR1 activates pro-tumorigenic pathways, such as Ras/ERK, PI3K/AKT, STAT3 and PLC [150]. Activation of STAT3 via S1PR1 may promote tumor growth and progression in multiple ways, including angiogenesis [207], promotion of regulatory T cell

accumulation to the tumor microenvironment [208], or inducing tumor growth and metastasis via Il6 [209]. This has been observed with breast cancer, where S1PR1 activation of STAT3 was associated with tumor growth and progression [210], and S1PR3 activation has been linked with tumor progression and lower overall survival [150]. S1PR1 has also been implicated in the development of inflammation-associated colorectal cancers in individuals with inflammatory bowel disease [211], and the S1p/S1PR1 signaling axis may link obesity and breast cancer metastasis by activation of pro-inflammatory cytokines [212]. Although the evidence for a role of both S1P receptors in cancer is accumulating, their various tumorigenic activities across cancers, and specifically ovarian cancer, warrant further investigation.

This research work has several strengths including the large sample of cases on the POCS TMAs available for analysis. Each participant in the analysis had two cores for their ovarian cancer on the study TMAs, and the intra-individual correlations showed high agreement with respect to marker expression. The quality of the tumor characteristics evaluated in this study was high, as POCS data contains excellent pathology information that was reviewed by multiple experts who determined diagnosis upon consensus after tissue evaluation. Covariate information for all women in POCS was collected via trained interview with limited missing information for most covariates. Follow-up information for each woman was similarly well-documented allowing survival by marker status to be determined with high accuracy.

POCS TMAs contained 55% of the total POCS study participants, limiting the sample of tissue available for study. However, tumors from the women included on the POCS TMAs versus those who were not included were similar with respect to pathology, and missing stage and grade information was lower in the group included on the TMA. Risk factor characteristics were similar between cores that were included and those not included, and so selection bias is

not expected to have occurred when assembling the TMA. Limitations in this work include low case numbers in non-serous histology groups, particularly for mucinous and clear cell tumors, which is a common challenge of ovarian cancer studies such as the present analysis. However, expression across individual histotypes was examined in univariate analyses, and non-serous tumors were evaluated as a category in models for tumor characteristics and risk factors to capture the effect of this group on marker expression.

In summary, this is the first study to directly evaluate S1P receptor expression in ovarian cancer tissue according to important tumor characteristics and risk factors. Results from this work suggest that expression of either S1PR1 or S1PR3 may be heterogeneous across ovarian cancer histotypes, and related to increased stage or grade of disease. This may suggest an association with disease progression. To expand upon this work, additional molecular targets in the S1P pathway should be evaluated for expression in ovarian cancer tissue, and to further characterize expression in stratified non-serous histotypes. Understanding expression of S1P pathway makers may also be valuable in cancer precursor tissues, such as fallopian tube and endometrial tissue, where ovarian cancer histotype-specific cells of origin arise. Pathway activity could be further evaluated in other molecular datasets containing S1P pathway marker mRNA or other molecular information to further characterize downstream activity of S1P receptors.

Tables and figures for Manuscript 3

Table 19. Characteristics of ovarian cancer cases in POCS

Cases n=166	
Characteristic	
Age at diagnosis (years)	56 ± 10.6
Median survival time	6.9 ± 3.5 years
Age (years)	
<50	58 (34.9%)
50-54	18 (10.8%)
55-64	43 (24.0%)
65+	47 (25.9%)
Menopausal status	
Premenopausal	50 (30.1%)
Postmenopausal	94 (56.6%)
Missing	22 (13.3%)
Menopausal hormone use	
No	126 (75.9%)
Yes	40 (24.1%)
OC Use	
No	155 (93.3%)
Yes	11 (6.6%)
Parity	
Nulliparous	21 (12.7%)
Ever-pregnant	145 (87.3%)
Smoking	
Non-smoker	66 (39.8%)
Former	34 (20.5%)
Current	66 (39.8%)
BMI kg/m ²	
<25	89 (53.6%)
25-30	41 (24.7%)
>30	29 (17.5%)
Missing	7 (4.2%)
Grade	
Well-differentiated	33 (19.9%)
Moderately differentiated	82 (49.4%)
Poorly differentiated	37 (22.3%)
Missing	14 (8.4%)
Stage	
I	60 (36.1%)
≥II	85 (51.2%)
Missing	21 (12.7%)
Histology	
Serous	70 (42.2%)
Endometrioid	19 (11.4%)
Mucinous	7 (4.2%)
Clear Cell	7 (4.2%)
Mixed epithelial	38 (24.1%)
Other	25 (15.1%)

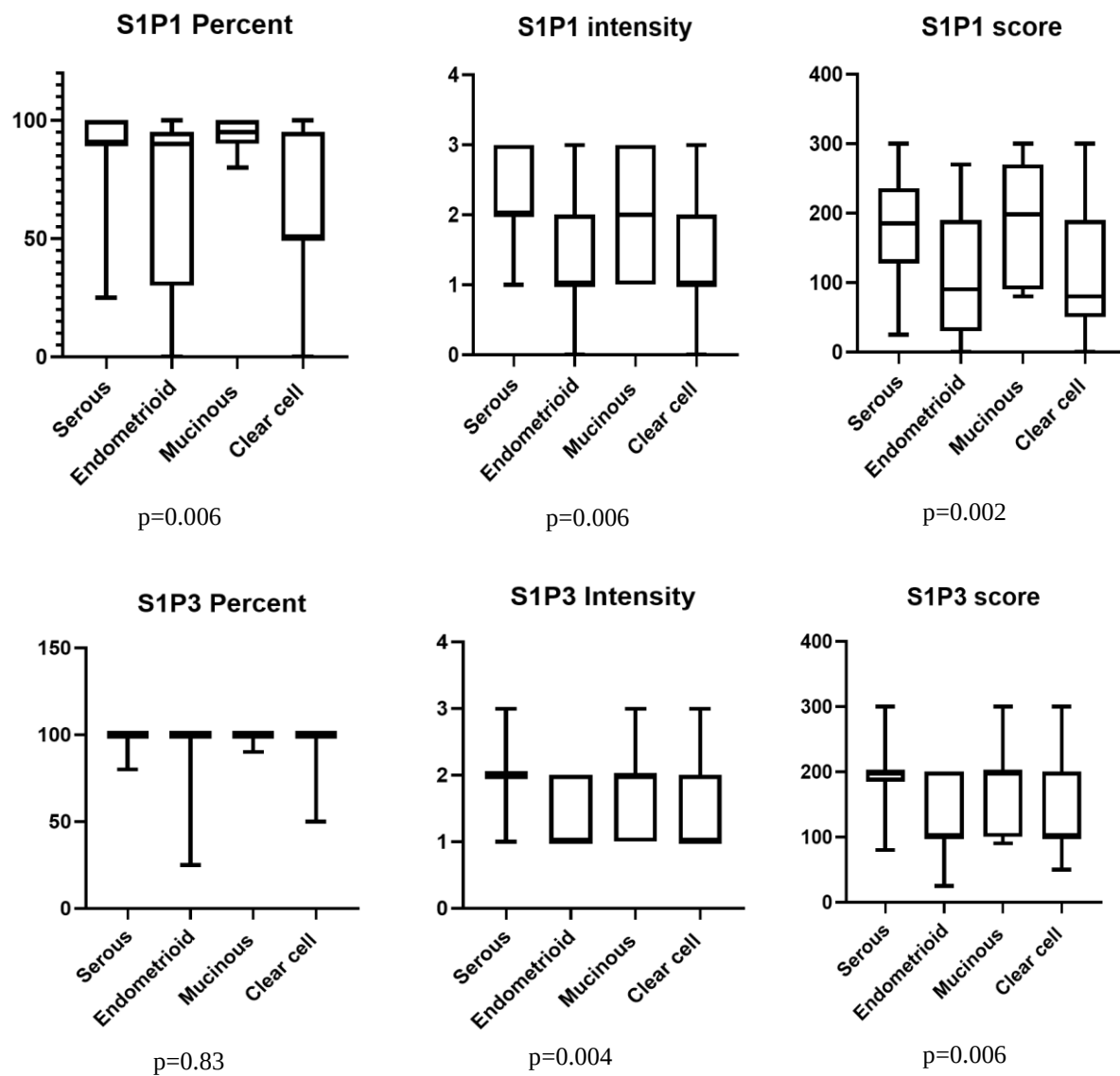


Figure 18. Distributions of area, intensity and staining score measures for S1PR1 and S1PR3

Table 20. Associations between tumor characteristics and S1PR1 and S1PR3 expression (n=166)

	S1PR1			S1PR3		
	High expression (Qs 2-4) n= 120	Low expression (Q1) n= 46	OR (95% CI)*	High expression above median n=25	Low expression below median n=141	OR (95% CI)*
Tumor characteristics						
Stage						
I	39 (32.5%)	21 (45.7%)	(ref)	8 (32.0%)	52 (36.9%)	(ref)
II-III	64 (53.3%)	21 (45.7%)	1.55 (0.71-3.36)	13 (52.0%)	72 (51.1%)	1.51 (0.51-4.42)
Missing	17 (14.2%)	4 (8.7%)	-	4 (16.0%)	17 (12.1%)	-
Grade						
Well-differentiated	21 (17.5%)	12 (26.1%)	(ref)	4 (16.0%)	29 (20.6%)	(ref)
Moderately differentiated	56 (46.7%)	26 (56.5%)	1.33 (0.51-3.50)	10 (40.0%)	72 (51.1%)	1.12 (0.27-4.59)
Poorly differentiated	30 (25.0%)	7 (15.2%)	2.49 (0.73-8.47) [†]	7 (28.0%)	30 (21.3%)	1.56 (0.32-7.60) [‡]
Missing	13 (10.8%)	1 (2.2%)	-	4 (16.0%)	10 (7.1%)	-
Histology						
Serous	60 (50.0%)	10 (21.7%)	(ref)	15 (60.0%)	55 (39.0%)	(ref)
Non-serous	22 (18.3%)	13 (28.3%)	0.16 (0.06-0.48)	2 (8.0%)	28 (19.9%)	-
Mixed adenocarcinoma	23 (19.2%)	15 (32.6%)	0.23 (0.08-0.66)	4 (16.0%)	34 (24.1%)	0.56 (0.14-2.26)
Other	20 (16.7%)	5 (10.8%)	0.46 (0.12-1.73)	4 (16.0%)	21 (14.9%)	0.65 (0.17-2.54)

Abbreviations: S1PR1 = Sphingosine phosphate receptor 1, S1PR3 = sphingosine phosphate receptor 3, OR = odds ratio, CI = confidence interval, ref = reference

*Adjusted for age in 5-year categories, menopause status (pre-menopausal, post-menopausal, missing), use of menopausal hormone therapy (ever versus never), smoking history (current, former, never), parity (ever versus never), OC use (ever versus never) and BMI (<25, 25-29.9, 30+)

[‡]ptrend = 0.13

[†]ptrend = 0.04

Table 21. Associations between S1P receptor staining and ovarian cancer risk factors

	S1PR1 Score			S1PR3 Score		
	High expression Qs (2-4) n=120	Low expression Q1 n=46	OR (95% CI)*	High Expression above median n= 25	Low expression below median n= 141	OR (95% CI)*
Age (years)						
<50	47 (39.2%)	11 (23.9%)	(ref)	6 (24.0%)	52 (36.9%)	(ref)
50-54	9 (7.5%)	9 (19.6%)	0.27 (0.08-0.92)	2 (8.0%)	16 (11.3%)	0.56 (0.07-4.64)
55-64	30 (25.0%)	13 (28.3%)	0.78 (0.28-2.16)	9 (36.0%)	34 (24.1%)	0.76 (0.13-4.54)
65+	34 (28.3%)	13 (28.3%)	0.79 (0.28-2.19)	8 (32.0%)	39 (27.2%)	0.53 (0.08-3.47)
Menopausal status						
Premenopausal	41 (34.2%)	9 (19.6%)	(ref)	3 (12.0%)	47 (33.3%)	(ref)
Postmenopausal	63 (52.5%)	31 (67.4%)	0.59 (0.24-1.47)	18 (72.0%)	76 (53.9%)	5.92 (0.8-41.54)
Missing	16 (13.3%)	6 (13.0%)	-	4 (16.0%)	18 (12.8%)	-
BMI kg/m²						
< 25	68 (56.7%)	21 (46.7%)	(ref)	15 (60.0%)	74 (52.5%)	(ref)
25 - 30	28 (23.3%)	13 (28.3%)	0.66 (0.28-1.49)	4 (16.0%)	37 (26.2%)	0.33 (0.09-1.23)
30+	20 (16.7%)	9 (19.6%)	0.78 (0.31-1.93)	3 (12.0%)	26 (18.4%)	0.38 (0.08-1.73)
missing	4 (3.3%)	3 (6.5%)	-	3 (21.0%)	4 (2.8%)	-
MHT use						
No	91 (75.8%)	35 (76.1%)	(ref)	16 (64.0%)	110 (78.0%)	(ref)
Yes	29 (24.2%)	11 (23.9%)	1.09 (0.46-2.61)	9 (36.0%)	31 (22.0%)	2.75 (0.91-8.29)
Smoking status						
Non-smoker	51 (42.5%)	15 (32.6%)	(ref)	11 (44.0%)	55 (39.0%)	(ref)
Former	21 (17.5%)	13 (28.3%)	0.52 (0.20-1.42)	6 (24.0%)	28 (19.9%)	0.76 (0.20-2.90)
Current	48 (40.0%)	18 (39.1%)	0.77 (0.31-1.90)	8 (32.0%)	58 (41.1%)	0.79 (0.25-2.52)
Parity						
No	14 (11.7%)	7 (15.2%)	(ref)	5 (20.0%)	16 (11.3%)	(ref)
Yes	106 (88.3%)	39 (84.7%)	1.22 (0.42-3.59)	20 (80.0%)	125 (88.7%)	0.56 (0.15-2.01)
OC use						
Never	110 (91.7%)	45 (97.8%)	(ref)	25 (100.0%)	130 (92.2%)	(ref)

Ever	10 (8.3%)	1 (2.2%)	-	0 (0.0%)	11 (7.8%)	-
------	-----------	----------	---	----------	-----------	---

Abbreviations: S1PR1 = sphingosine phosphate receptor 1; S1PR3 = sphingosine phosphate receptor 3; Q = quartile; OR = odds ratio; CI = confidence interval; MHT = menopausal hormone therapy; OC = oral contraceptive; BMI = body mass index

* Adjusted for age in 5-year categories, menopause status (pre-menopausal, post-menopausal, missing), use of menopausal hormone therapy (ever versus never), smoking history (current, former, never), parity (ever versus never), OC use (ever versus never) and BMI (<25, 25-29.9, 30+)

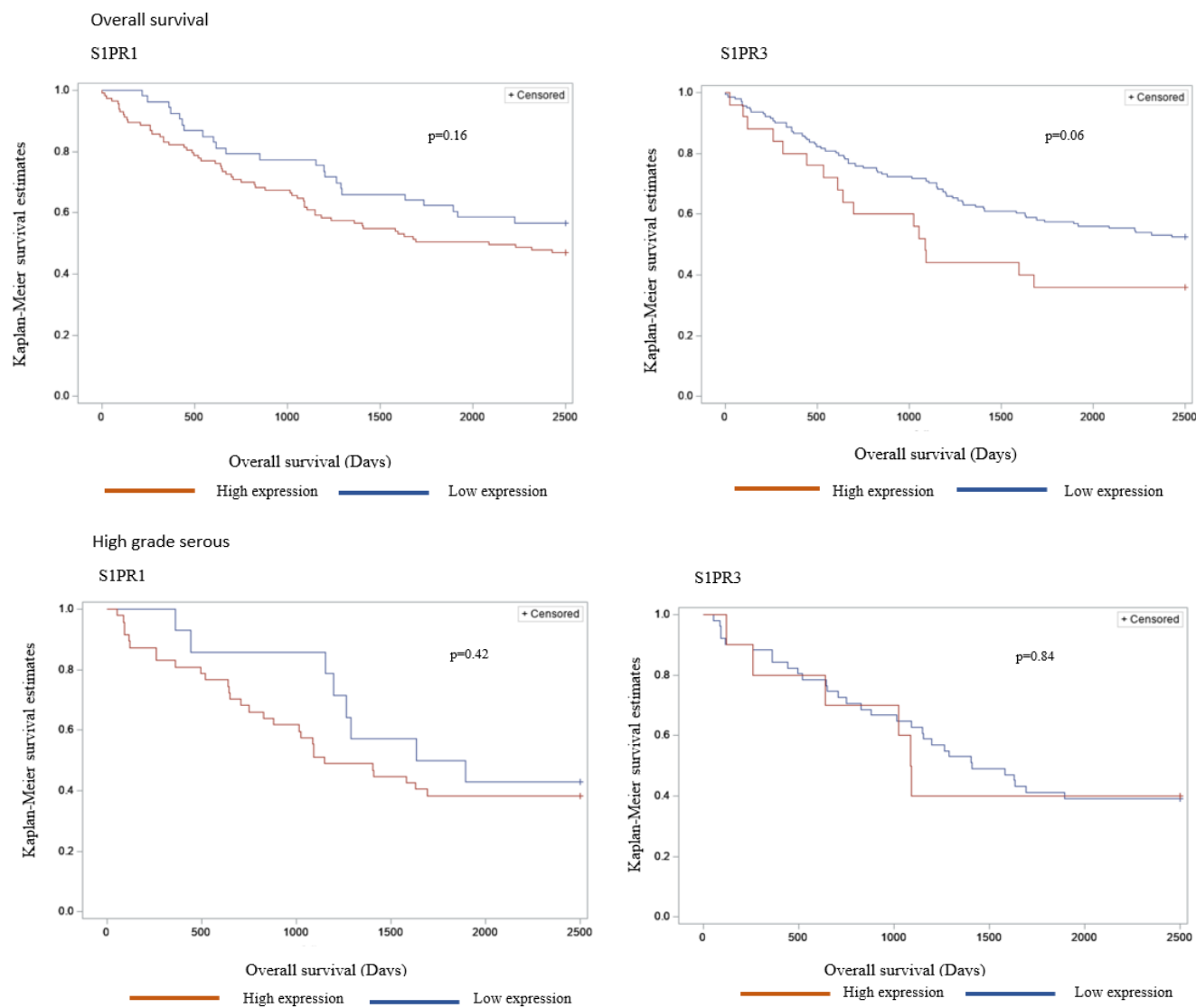


Figure 19. Kaplan-Meier overall survival curves for S1PR1 and S1PR3 expression in ovarian cancer tissue

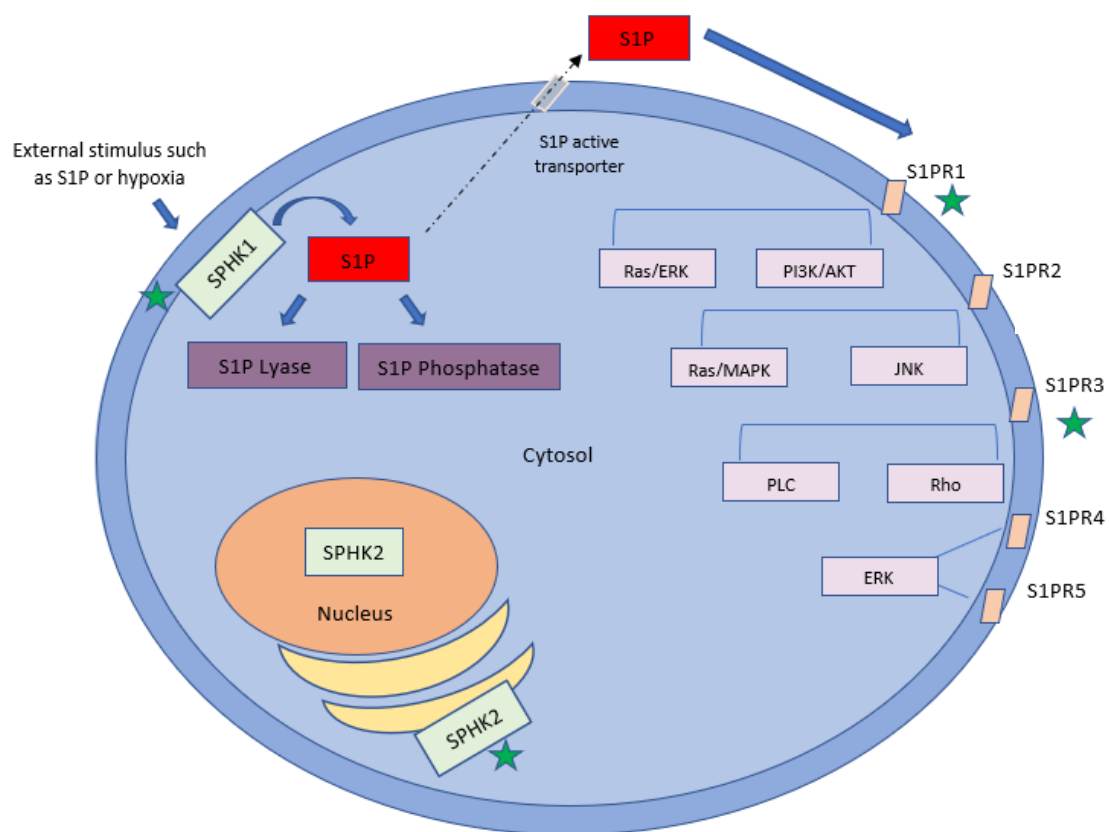


Figure 20. S1P depiction of S1P pathway components

Chapter 7: Public Health Significance

Conclusions

The findings of this dissertation suggest that although several metabolic factors are associated with increased ovarian cancer risk, many questions remain as to whether metabolic changes are targetable for chemoprevention or if they can provide insight to carcinogenic pathways. Two widely-used medications used to treat metabolic conditions, metformin and statins, were not associated with reduced risk of ovarian cancer either overall or in possible benefit groups, such as women taking each drug for specific indications. Thus, findings from this dissertation did not support the use of either drug in chemoprevention for ovarian cancer overall. However, Aims 1 and 2 showed that risk factors associated with metabolic disturbances, such as PCOS and higher BMI, were associated with increased ovarian cancer risk. In addition to the epidemiological investigation of medications used to treat metabolic and lipid disturbances, the third aim of the dissertation evaluated two cellular receptors that are components of the S1P cellular pathway, a bioactive lipid associated with angiogenesis, cellular proliferation and cell survival. Overall, expression of both S1P receptors, S1PR1 and S1PR3 were heterogenous across ovarian cancer histotype. S1PR1 and 3, which are related to angiogenesis, may also be elevated in higher grade and stage of ovarian cancer, but sample sizes were limited. Taken together, metabolic factors play a role in ovarian cancer, however, understanding the complexity in metabolic changes and avenues for possible targeting of metabolic pathways requires further investigation.

Public Health Implications

Ovarian cancer is the most fatal gynecologic malignancy [2]. Although rare, ovarian cancer is a considerable public health challenge, as no screening options are available and many women do not know they have the disease until it has reached an advanced stage [6]. The rarity of ovarian cancer as well as its heterogeneity make studying this disease challenging, as risk factor profiles and molecular characteristics differ according to histology [10]. However, studies that could identify opportunities for risk-reducing prevention strategies are critical. Aims 1 and 2 contribute two large, population-based research studies aimed at identifying possible lipid-related chemoprevention options for ovarian cancer that uncovered no associations overall or within risk groups. This research contributes to existing knowledge and ongoing investigations aimed at identifying drugs for “repurposing” in ovarian cancer as it provides clues that two prominent medications may not be beneficial for ovarian cancer overall. The exploratory analysis in Aim 3 of this dissertation provided valuable information that lipid signaling might be different by ovarian cancer histotype and may be enhanced in later stage or grade disease, even though sample sizes were limited in subgroups defined by histology and tumor characteristics. These novel findings support further investigation of the SIP pathway in ovarian cancers to determine if it is a possible targetable pathway for prevention options, as seen in other cancer types.

Limitations

Limitations to the studies presented in this dissertation were described in detail in Chapter 3, as well as in the discussion of each dissertation manuscript. The primary limitations in this research include the possible misclassification of covariates and common inconsistencies in the collection of EMR data, similar to that used for the analyses in Manuscripts 1 and 2.

Although the CPRD population for Manuscripts 1 and 2 was sampled from the general population of the UK, medication users were also indicated for drug use, meaning that the effect of chemopreventive agents was not possible to assess outside this context. This dissertation was not able to assess medication associations with ovarian cancer risk by histotype, and as obesity is associated with some non-serous types, it stands to reason that additional information could be gleaned from histotype-specific effects. Lack of laboratory data prevented the ability to separate medication users into additional risk groups based on severity. Limitations of Aim 3 of this dissertation include small case numbers for non-serous tumors, that may have been beneficial in elucidating more refined effects by histology rather than a general classification as non-serous.

Future directions

This dissertation evaluated only two medications used to treat metabolic disturbances. Future studies should assess other medications used to treat lipid disorders, considering varied mechanisms of action and additional indications for each drug. Associations according to ovarian cancer histotype also should be considered, when possible, as associations may differ by histotype and potential benefit groups are yet to be uncovered. Understanding a role for metabolic disturbances in any cancer is complex, as metabolic changes could affect disease development through numerous mechanisms, many of which are not understood. Using multiple types of epidemiologic and molecular analyses to address aspects of lipid metabolism, including observational and molecular studies, could continue to paint a continuously evolving picture of metabolic changes and associations with ovarian cancer.

Appendices

Appendix A. Read code lists for Aims 1 and 2

Codes for identifying ovarian cancers in CPRD primary care data

Med Code	Read Code	Read Term
19141	B44.00	Malignant neoplasm of ovary and other uterine adnexa
7805	B440.00	Malignant neoplasm of ovary
1986	B440.11	Cancer of ovary
49828	B441.00	Malignant neoplasm of fallopian tube
101778	B442.00	Malignant neoplasm of broad ligament
46153	B443.00	Malignant neoplasm of parametrium
97996	B44y.00	Malignant neoplasm of other site of uterine adnexa
65106	B44z.00	Malignant neoplasm of uterine adnexa NOS
30541	BB81.14	Ovarian serous tumor
18638	BB81.12	Ovarian mucinous tumor
2979	BB81.00	Ovarian cystic, mucinous and serous neoplasms

Abbreviations: CPRD = Clinical Practice Research Datalink

Codes for identifying metformin use in CPRD primary care data

Product Code	Product Name
60497	Alogliptin 12.5mg / metformin 1g tablets
17580	Avandamet 1mg/500mg tablets (GlaxoSmithKline UK Ltd)
14164	Avandamet 2mg/1000mg tablets (GlaxoSmithKline UK Ltd)

6855	Avandamet 2mg/500mg tablets (GlaxoSmithKline UK Ltd)
7325	Avandamet 4mg/1000mg tablets (GlaxoSmithKline UK Ltd)
57147	Bolamyn SR 1000mg tablets (Teva UK Ltd)
39560	Bolamyn SR 500mg tablets (Teva UK Ltd)
63929	Canagliflozin 50mg / metformin 1g tablets
31077	Competact 15mg/850mg tablets (Takeda UK Ltd)
60012	Dapagliflozin 5mg / metformin 1g tablets
63031	Dapagliflozin 5mg / metformin 850mg tablets
52221	Diagemet XL 500mg tablets (Genus Pharmaceuticals Ltd)
55270	Duformin 500mg Tablet (Dumex Ltd)
38551	Eucreas 50mg/1000mg tablets (Novartis Pharmaceuticals UK Ltd)
39203	Eucreas 50mg/850mg tablets (Novartis Pharmaceuticals UK Ltd)
25678	Glucamet 500mg Tablet (Opus Pharmaceuticals Ltd)
26258	Glucamet 850mg Tablet (Opus Pharmaceuticals Ltd)
40007	Glucophage 1000mg oral powder sachets (Merck Serono Ltd)
40110	Glucophage 500mg oral powder sachets (Merck Serono Ltd)
7166	Glucophage 500mg tablets (Merck Serono Ltd)
7610	Glucophage 850mg tablets (Merck Serono Ltd)
39729	Glucophage SR 1000mg tablets (Merck Serono Ltd)
52634	Glucophage SR 500mg tablets (DE Pharmaceuticals)
50570	Glucophage SR 500mg tablets (Lexon (UK) Ltd)
49502	Glucophage SR 500mg tablets (Mawdsley-Brooks & Company Ltd)

16044	Glucophage SR 500mg tablets (Merck Serono Ltd)
59620	Glucophage SR 500mg tablets (Waymade Healthcare Plc)
38400	Glucophage SR 750mg tablets (Merck Serono Ltd)
43684	Janumet 50mg/1000mg tablets (Merck Sharp & Dohme Ltd)
50682	Jentadueto 2.5mg/1000mg tablets (Boehringer Ingelheim Ltd)
54150	Jentadueto 2.5mg/850mg tablets (Boehringer Ingelheim Ltd)
56965	Komboglyze 2.5mg/1000mg tablets (AstraZeneca UK Ltd)
58865	Komboglyze 2.5mg/850mg tablets (AstraZeneca UK Ltd)
52445	Linagliptin 2.5mg / metformin 1g tablets
52449	Linagliptin 2.5mg / metformin 850mg tablets
51080	Metabet SR 1000mg tablets (Actavis UK Ltd)
46989	Metabet SR 1000mg tablets (Morningside Healthcare Ltd)
53774	Metabet SR 500mg tablets (Actavis UK Ltd)
45581	Metabet SR 500mg tablets (Morningside Healthcare Ltd)
20810	METFORMIN
54442	metformin (roi) 1000mg Tablet
735	metformin 100mg/ml Oral solution
43619	metformin 1g / Sitagliptin 50mg tablets
39598	metformin 1g modified-release tablets
49738	metformin 1g modified-release tablets (A H Pharmaceuticals Ltd)
62824	metformin 1g modified-release tablets (Actavis UK Ltd)
60074	metformin 1g modified-release tablets (Waymade Healthcare Plc)

40233	metformin 1g oral powder sachets sugar free
63307	metformin 1g/5ml oral solution
16213	METFORMIN 250 MG TAB
7048	metformin 500mg modified-release tablets
51135	metformin 500mg modified-release tablets (A H Pharmaceuticals Ltd)
60968	metformin 500mg modified-release tablets (Actavis UK Ltd)
62144	metformin 500mg modified-release tablets (DE Pharmaceuticals)
53478	metformin 500mg modified-release tablets (Kent Pharmaceuticals Ltd)
62265	metformin 500mg modified-release tablets (Mawdsley-Brooks & Company Ltd)
39988	metformin 500mg oral powder sachets sugar free
34135	metformin 500mg Tablet (M & A Pharmachem Ltd)
23	metformin 500mg tablets
34323	metformin 500mg tablets (A H Pharmaceuticals Ltd)
33087	metformin 500mg tablets (Actavis UK Ltd)
48149	metformin 500mg tablets (Almus Pharmaceuticals Ltd)
57457	metformin 500mg tablets (Aurobindo Pharma Ltd)
34004	metformin 500mg tablets (IVAX Pharmaceuticals UK Ltd)
34598	metformin 500mg tablets (Mylan Ltd)
34917	metformin 500mg tablets (Teva UK Ltd)
55739	metformin 500mg tablets (Tillomed Laboratories Ltd)
34504	metformin 500mg tablets (Wockhardt UK Ltd)

53867	metformin 500mg tablets (Zentiva)
58051	metformin 500mg/5ml oral solution
44250	metformin 500mg/5ml Oral solution (Hillcross Pharmaceuticals Ltd)
11990	metformin 500mg/5ml oral solution sugar free
43270	metformin 500mg/5ml oral solution sugar free (Rosemont Pharmaceuticals Ltd)
58607	metformin 500mg/5ml oral solution sugar free (Zentiva)
60286	metformin 500mg/5ml oral suspension
38355	metformin 750mg modified-release tablets
7815	METFORMIN 800 MG TAB
93	metformin 850mg tablets
33674	metformin 850mg tablets (A H Pharmaceuticals Ltd)
34836	metformin 850mg tablets (Actavis UK Ltd)
54898	metformin 850mg tablets (Almus Pharmaceuticals Ltd)
34020	metformin 850mg tablets (IVAX Pharmaceuticals UK Ltd)
62605	metformin 850mg tablets (Kent Pharmaceuticals Ltd)
34742	metformin 850mg tablets (Teva UK Ltd)
34697	metformin 850mg tablets (Wockhardt UK Ltd)
3252	METFORMIN HCl 500 MG TAB
2928	METFORMIN HCl 850 MG TAB
30316	metformin with pioglitazone 850mg + 15mg Tablet
11760	metformin with rosiglitazone 1000mg + 2mg Tablet

11737	metformin with rosiglitazone 1000mg + 4mg Tablet
11609	metformin with rosiglitazone 500mg + 1mg Tablet
11610	metformin with rosiglitazone 500mg + 2mg Tablet
31146	Metsol 500mg/5ml oral solution (Kappin Ltd)
27501	Orabet 500mg Tablet (Lagap)
42161	Orabet 500mg Tablet (Sandoz Ltd)
18220	Pioglitazone 15mg / metformin 850mg tablets
11604	Rosiglitazone 1mg / metformin 500mg tablets
11717	Rosiglitazone 2mg / metformin 1g tablets
11601	Rosiglitazone 2mg / metformin 500mg tablets
7375	Rosiglitazone 4mg / metformin 1g tablets
54891	Saxagliptin 2.5mg / metformin 1g tablets
54973	Saxagliptin 2.5mg / metformin 850mg tablets
61043	Sukkarto SR 1000mg tablets (Morningside Healthcare Ltd)
61559	Sukkarto SR 500mg tablets (Morningside Healthcare Ltd)
37902	Vildagliptin 50mg / metformin 1g tablets
37874	Vildagliptin 50mg / metformin 850mg tablets
59385	Vipdomet 12.5mg/1000mg tablets (Takeda UK Ltd)
60643	Xigduo 5mg/1000mg tablets (AstraZeneca UK Ltd)

Abbreviations: CPRD = Clinical Practice Research Datalink; mg = milligrams; UK = United Kingdom

Codes for identifying statin use in CPRD

Product Code	Product Name
56097	Atorvastatin 10mg chewable tablets sugar free
28	Atorvastatin 10mg tablets
51134	Atorvastatin 10mg tablets (A H Pharmaceuticals Ltd)
55727	Atorvastatin 10mg tablets (Actavis UK Ltd)
63140	Atorvastatin 10mg tablets (Alliance Healthcare (Distribution) Ltd)
57348	Atorvastatin 10mg tablets (Consilient Health Ltd)
55032	Atorvastatin 10mg tablets (Dexcel-Pharma Ltd)
54535	Atorvastatin 10mg tablets (Pfizer Ltd)
59357	Atorvastatin 10mg tablets (Ranbaxy (UK) Ltd)
59859	Atorvastatin 10mg tablets (Teva UK Ltd)
61149	Atorvastatin 10mg tablets (Waymade Healthcare Plc)
50236	Atorvastatin 10mg tablets (Zentiva)
48518	Atorvastatin 10mg/5ml oral solution
54992	Atorvastatin 10mg/5ml oral suspension
56165	Atorvastatin 20mg chewable tablets sugar free
75	Atorvastatin 20mg tablets
49558	Atorvastatin 20mg tablets (A H Pharmaceuticals Ltd)
52211	Atorvastatin 20mg tablets (Actavis UK Ltd)
58394	Atorvastatin 20mg tablets (Alliance Healthcare (Distribution) Ltd)
56564	Atorvastatin 20mg tablets (Almus Pharmaceuticals Ltd)

51359	Atorvastatin 20mg tablets (Arrow Generics Ltd)
52168	Atorvastatin 20mg tablets (Aspire Pharma Ltd)
51622	Atorvastatin 20mg tablets (Consilient Health Ltd)
62219	Atorvastatin 20mg tablets (DE Pharmaceuticals)
50790	Atorvastatin 20mg tablets (Dexcel-Pharma Ltd)
50788	Atorvastatin 20mg tablets (Pfizer Ltd)
58041	Atorvastatin 20mg tablets (Teva UK Ltd)
59272	Atorvastatin 20mg tablets (Waymade Healthcare Plc)
58110	Atorvastatin 20mg tablets (Zentiva)
64067	Atorvastatin 20mg/5ml oral solution
60464	Atorvastatin 20mg/5ml oral suspension
48973	Atorvastatin 30mg tablets
63469	Atorvastatin 30mg tablets (Consilient Health Ltd)
745	Atorvastatin 40mg tablets
52398	Atorvastatin 40mg tablets (A H Pharmaceuticals Ltd)
53887	Atorvastatin 40mg tablets (Actavis UK Ltd)
49751	Atorvastatin 40mg tablets (Alliance Healthcare (Distribution) Ltd)
59446	Atorvastatin 40mg tablets (Almus Pharmaceuticals Ltd)
51200	Atorvastatin 40mg tablets (Arrow Generics Ltd)
52460	Atorvastatin 40mg tablets (Aspire Pharma Ltd)
51876	Atorvastatin 40mg tablets (Consilient Health Ltd)
57834	Atorvastatin 40mg tablets (DE Pharmaceuticals)

56841	Atorvastatin 40mg tablets (Dexcel-Pharma Ltd)
52397	Atorvastatin 40mg tablets (Dr Reddy's Laboratories (UK) Ltd)
50272	Atorvastatin 40mg tablets (Pfizer Ltd)
60511	Atorvastatin 40mg tablets (Ranbaxy (UK) Ltd)
50963	Atorvastatin 40mg tablets (Teva UK Ltd)
52097	Atorvastatin 40mg tablets (Wockhardt UK Ltd)
55444	Atorvastatin 40mg tablets (Zentiva)
55034	Atorvastatin 40mg/5ml oral suspension
48346	Atorvastatin 60mg tablets
5775	Atorvastatin 80mg tablets
58418	Atorvastatin 80mg tablets (A H Pharmaceuticals Ltd)
52459	Atorvastatin 80mg tablets (Actavis UK Ltd)
53772	Atorvastatin 80mg tablets (Alliance Healthcare (Distribution) Ltd)
62476	Atorvastatin 80mg tablets (Almus Pharmaceuticals Ltd)
58742	Atorvastatin 80mg tablets (Arrow Generics Ltd)
59776	Atorvastatin 80mg tablets (Aspire Pharma Ltd)
63249	Atorvastatin 80mg tablets (Consilient Health Ltd)
60607	Atorvastatin 80mg tablets (DE Pharmaceuticals)
52821	Atorvastatin 80mg tablets (Dr Reddy's Laboratories (UK) Ltd)
53890	Atorvastatin 80mg tablets (Pfizer Ltd)
60989	Atorvastatin 80mg tablets (Phoenix Healthcare Distribution Ltd)
57836	Atorvastatin 80mg tablets (Teva UK Ltd)

57117	Atorvastatin 80mg tablets (Waymade Healthcare Plc)
56182	Atorvastatin 80mg tablets (Zentiva)
420	Cerivastatin 100microgram tablets
5009	Cerivastatin 200microgram tablets
5251	Cerivastatin 300microgram tablets
5278	Cerivastatin 400microgram tablets
31658	Cerivastatin 800microgram tablets
7347	Crestor 10mg tablets (AstraZeneca UK Ltd)
53460	Crestor 10mg tablets (DE Pharmaceuticals)
15252	Crestor 20mg tablets (AstraZeneca UK Ltd)
59447	Crestor 20mg tablets (Waymade Healthcare Plc)
9930	Crestor 40mg tablets (AstraZeneca UK Ltd)
57999	Crestor 40mg tablets (Lexon (UK) Ltd)
17688	Crestor 5mg tablets (AstraZeneca UK Ltd)
379	Fluvastatin 20mg capsules
62148	Fluvastatin 20mg capsules (Actavis UK Ltd)
59278	Fluvastatin 20mg capsules (Zentiva)
2137	Fluvastatin 40mg capsules
53770	Fluvastatin 40mg capsules (A H Pharmaceuticals Ltd)
11627	Fluvastatin 80mg modified-release tablets
21020	Inegy 10mg/20mg tablets (Merck Sharp & Dohme Ltd)
17059	Inegy 10mg/40mg tablets (Merck Sharp & Dohme Ltd)

16186	Inegy 10mg/80mg tablets (Merck Sharp & Dohme Ltd)
8380	Lescol 20mg capsules (Novartis Pharmaceuticals UK Ltd)
9153	Lescol 40mg capsules (Novartis Pharmaceuticals UK Ltd)
5985	Lescol XL 80mg tablets (Novartis Pharmaceuticals UK Ltd)
3411	Lipitor 10mg tablets (Pfizer Ltd)
56016	Lipitor 20mg chewable tablets (Pfizer Ltd)
7374	Lipitor 20mg tablets (Pfizer Ltd)
2955	Lipitor 40mg tablets (Pfizer Ltd)
17683	Lipitor 80mg tablets (Pfizer Ltd)
9315	Lipobay 100microgram Tablet (Bayer Plc)
9316	Lipobay 200microgram Tablet (Bayer Plc)
4961	Lipobay 300microgram Tablet (Bayer Plc)
18442	Lipobay 400microgram Tablet (Bayer Plc)
1221	Lipostat 10mg tablets (Bristol-Myers Squibb Pharmaceuticals Ltd)
3690	Lipostat 20mg tablets (Bristol-Myers Squibb Pharmaceuticals Ltd)
1223	Lipostat 40mg tablets (Bristol-Myers Squibb Pharmaceuticals Ltd)
32921	Pravastatin 10mg Tablet (Dr Reddy's Laboratories (UK) Ltd)
490	Pravastatin 10mg tablets
57137	Pravastatin 10mg tablets (Almus Pharmaceuticals Ltd)
60251	Pravastatin 10mg tablets (Sandoz Ltd)
43218	Pravastatin 10mg tablets (Teva UK Ltd)
56146	Pravastatin 10mg tablets (Waymade Healthcare Plc)

730	Pravastatin 20mg tablets
40382	Pravastatin 20mg tablets (A H Pharmaceuticals Ltd)
54607	Pravastatin 20mg tablets (Almus Pharmaceuticals Ltd)
51890	Pravastatin 20mg tablets (Medreich Plc)
56735	Pravastatin 20mg tablets (Mylan Ltd)
63074	Pravastatin 20mg tablets (PLIVA Pharma Ltd)
36377	Pravastatin 20mg tablets (Teva UK Ltd)
56607	Pravastatin 20mg tablets (Waymade Healthcare Plc)
1219	Pravastatin 40mg tablets
34820	Pravastatin 40mg tablets (A H Pharmaceuticals Ltd)
54435	Pravastatin 40mg tablets (Almus Pharmaceuticals Ltd)
51676	Pravastatin 40mg tablets (Medreich Plc)
47988	Pravastatin 40mg tablets (Mylan Ltd)
56916	Pravastatin 40mg tablets (PLIVA Pharma Ltd)
48097	Pravastatin 40mg tablets (Teva UK Ltd)
57108	Pravastatin 40mg tablets (Waymade Healthcare Plc)
44878	Ranzolont 10mg tablets (Ranbaxy (UK) Ltd)
713	Rosuvastatin 10mg tablets
57763	Rosuvastatin 10mg tablets (Waymade Healthcare Plc)
6213	Rosuvastatin 20mg tablets
58617	Rosuvastatin 20mg/5ml oral suspension
9897	Rosuvastatin 40mg tablets

7554	Rosuvastatin 5mg tablets
60160	Rosuvastatin 5mg tablets (Mawdsley-Brooks & Company Ltd)
59452	Rosuvastatin 5mg tablets (Waymade Healthcare Plc)
13041	Simvador 10mg tablets (Discovery Pharmaceuticals Ltd)
9920	Simvador 20mg tablets (Discovery Pharmaceuticals Ltd)
802	Simvador 40mg tablets (Discovery Pharmaceuticals Ltd)
39870	Simvador 80mg tablets (Discovery Pharmaceuticals Ltd)
24509	SIMVASTATIN
29438	SIMVASTATIN
34560	Simvastatin 10mg Tablet (Ratiopharm UK Ltd)
42	Simvastatin 10mg tablets
34955	Simvastatin 10mg tablets (A H Pharmaceuticals Ltd)
48078	Simvastatin 10mg tablets (Actavis UK Ltd)
61360	Simvastatin 10mg tablets (Almus Pharmaceuticals Ltd)
47774	Simvastatin 10mg tablets (Arrow Generics Ltd)
64180	Simvastatin 10mg tablets (Crescent Pharma Ltd)
53908	Simvastatin 10mg tablets (Dexcel-Pharma Ltd)
34481	Simvastatin 10mg tablets (IVAX Pharmaceuticals UK Ltd)
48051	Simvastatin 10mg tablets (Kent Pharmaceuticals Ltd)
51085	Simvastatin 10mg tablets (Medreich Plc)
34535	Simvastatin 10mg tablets (Mylan Ltd)
48058	Simvastatin 10mg tablets (Ranbaxy (UK) Ltd)

61321	Simvastatin 10mg tablets (Sandoz Ltd)
40340	Simvastatin 10mg tablets (Teva UK Ltd)
47948	Simvastatin 10mg tablets (Tillomed Laboratories Ltd)
61665	Simvastatin 10mg tablets (Waymade Healthcare Plc)
52625	Simvastatin 10mg tablets (Wockhardt UK Ltd)
52676	Simvastatin 10mg/5ml oral suspension
7552	Simvastatin 20mg / Ezetimibe 10mg tablets
34746	Simvastatin 20mg Tablet (Niche Generics Ltd)
34476	Simvastatin 20mg Tablet (Ratiopharm UK Ltd)
25	Simvastatin 20mg tablets
33082	Simvastatin 20mg tablets (A H Pharmaceuticals Ltd)
45245	Simvastatin 20mg tablets (Actavis UK Ltd)
54947	Simvastatin 20mg tablets (Almus Pharmaceuticals Ltd)
48018	Simvastatin 20mg tablets (Arrow Generics Ltd)
64104	Simvastatin 20mg tablets (Crescent Pharma Ltd)
39060	Simvastatin 20mg tablets (Dexcel-Pharma Ltd)
34366	Simvastatin 20mg tablets (IVAX Pharmaceuticals UK Ltd)
34891	Simvastatin 20mg tablets (Kent Pharmaceuticals Ltd)
50754	Simvastatin 20mg tablets (Medreich Plc)
34312	Simvastatin 20mg tablets (Mylan Ltd)
40601	Simvastatin 20mg tablets (Ranbaxy (UK) Ltd)
45235	Simvastatin 20mg tablets (Sandoz Ltd)

34316	Simvastatin 20mg tablets (Teva UK Ltd)
53676	Simvastatin 20mg tablets (Tillomed Laboratories Ltd)
58315	Simvastatin 20mg tablets (Waymade Healthcare Plc)
34814	Simvastatin 20mg tablets (Wockhardt UK Ltd)
11815	Simvastatin 20mg with ezetimibe 10mg tablet
818	Simvastatin 20mg/5ml oral solution sugar free
54266	Simvastatin 20mg/5ml oral suspension
39675	Simvastatin 20mg/5ml Oral suspension (Martindale Pharmaceuticals Ltd)
48221	Simvastatin 20mg/5ml oral suspension sugar free
54606	Simvastatin 20mg/5ml oral suspension sugar free (A H Pharmaceuticals Ltd)
44528	Simvastatin 20mg/5ml oral suspension sugar free (Rosemont Pharmaceuticals Ltd)
56065	Simvastatin 20mg/5ml oral suspension sugar free (Waymade Healthcare Plc)
57329	Simvastatin 25mg/5ml oral suspension
10172	Simvastatin 40mg / Ezetimibe 10mg tablets
34879	Simvastatin 40mg Tablet (Niche Generics Ltd)
34545	Simvastatin 40mg Tablet (Ratiopharm UK Ltd)
51	Simvastatin 40mg tablets
34502	Simvastatin 40mg tablets (A H Pharmaceuticals Ltd)
34969	Simvastatin 40mg tablets (Actavis UK Ltd)
46878	Simvastatin 40mg tablets (Almus Pharmaceuticals Ltd)
45346	Simvastatin 40mg tablets (Arrow Generics Ltd)

64307	Simvastatin 40mg tablets (Crescent Pharma Ltd)
44650	Simvastatin 40mg tablets (Dexcel-Pharma Ltd)
34381	Simvastatin 40mg tablets (IVAX Pharmaceuticals UK Ltd)
45219	Simvastatin 40mg tablets (Kent Pharmaceuticals Ltd)
51166	Simvastatin 40mg tablets (Medreich Plc)
34353	Simvastatin 40mg tablets (Mylan Ltd)
52098	Simvastatin 40mg tablets (Ranbaxy (UK) Ltd)
37434	Simvastatin 40mg tablets (Sandoz Ltd)
34376	Simvastatin 40mg tablets (Teva UK Ltd)
62137	Simvastatin 40mg tablets (Waymade Healthcare Plc)
34907	Simvastatin 40mg tablets (Wockhardt UK Ltd)
10183	Simvastatin 40mg with ezetimibe 10mg tablet
39652	Simvastatin 40mg/5ml oral solution sugar free
54985	Simvastatin 40mg/5ml oral suspension
48431	Simvastatin 40mg/5ml oral suspension sugar free
61155	Simvastatin 40mg/5ml oral suspension sugar free (A H Pharmaceuticals Ltd)
54819	Simvastatin 40mg/5ml oral suspension sugar free (Rosemont Pharmaceuticals Ltd)
14219	Simvastatin 80mg / Ezetimibe 10mg tablets
5148	Simvastatin 80mg tablets
32909	Simvastatin 80mg tablets (A H Pharmaceuticals Ltd)
49587	Simvastatin 80mg tablets (Almus Pharmaceuticals Ltd)

46956	Simvastatin 80mg tablets (Arrow Generics Ltd)
52962	Simvastatin 80mg tablets (Medreich Plc)
41657	Simvastatin 80mg tablets (Teva UK Ltd)
10206	Simvastatin 80mg with ezetimibe 10mg tablet
2718	Zocor 10mg tablets (Merck Sharp & Dohme Ltd)
7196	Zocor 20mg tablets (Merck Sharp & Dohme Ltd)
6168	Zocor 40mg tablets (Merck Sharp & Dohme Ltd)
22579	Zocor 80mg tablets (Merck Sharp & Dohme Ltd)
31930	Zocor heart-pro 10mg Tablet (McNeil Products Ltd)

Abbreviations: CPRD = Clinical Practice Research Datalink; mg = milligrams; UK = United Kingdom

Appendix B Control frequency matching

Match Type	Definition	Frequency	Percent
1	• equal YOB • control registered at practice at least 2 years prior to case index date and does not transfer out, die, develop cancer or have bilateral oophorectomy until after case index date • equal practice • equal year of registration • control is 40+ years old at case index date	29,579	32.1%
2	• equal YOB • control registered at practice at least 2 years prior to case index date and does not transfer out, die, develop cancer or have bilateral oophorectomy until after case index date • equal practice • year of registration is within 1 year • control is 40+ years old at case index date	14,784	16.1%
3	• YOB is within 1 year • control registered at practice at least 2 years prior to case index date and does not transfer out, die, develop cancer or have bilateral oophorectomy until after case index date • equal practice • year of registration is within 1 year • control is 40+ years old at case index date	24,804	26.9%
4	• YOB is within 1 year • control registered at practice at least 2 years prior to case index date and does not transfer out, die, develop cancer or have bilateral oophorectomy until after case index date • equal practice • year of registration is within 2 years • control is 40+ years old at case index date	8,581	9.3%
5	• YOB is within 2 years • control registered at practice at least 2 years prior to case index date and does not transfer out, die, develop cancer or have bilateral oophorectomy until after case index date • equal practice • year of registration is within 2 years • control is 40+ years old at case index date	6,687	7.3%
6	• YOB is within 2 years • control registered at practice at least 2 years	2,697	2.9%

	prior to case index date and does not transfer out, die, develop cancer or have bilateral oophorectomy until after case index date • equal practice • year of registration is within 3 years • control is 40+ years old at case index date		
7	• YOB is within 3 years • control registered at practice at least 2 years prior to case index date and does not transfer out, die, develop cancer or have bilateral oophorectomy until after case index date • equal practice • year of registration is within 3 years • control is 40+ years old at case index date	1,879	2.0%
8	• YOB is within 3 years • control registered at practice at least 2 years prior to case index date and does not transfer out, die, develop cancer or have bilateral oophorectomy until after case index date • equal practice • year of registration is within 4 years • control is 40+ years old at case index date	883	1.0%
9	• YOB is within 4 years • control registered at practice at least 2 years prior to case index date and does not transfer out, die, develop cancer or have bilateral oophorectomy until after case index date • equal practice • year of registration is within 4 years • control is 40+ years old at case index date	727	0.8%
10	• YOB is within 4 years • control registered at practice at least 2 years prior to case index date and does not transfer out, die, develop cancer or have bilateral oophorectomy until after case index date • equal practice • year of registration is within 5 years • control is 40+ years old at case index date	442	0.5%
11	• YOB is within 5 years • control registered at practice at least 2 years prior to case index date and does not transfer out, die, develop cancer or have bilateral oophorectomy until after case index date • equal practice • year of registration is within 5 years • control is 40+ years old at case index date	256	0.3%

12	• YOB is within 5 years • control registered at practice at least 2 years prior to case index date and does not transfer out, die, develop cancer or have bilateral oophorectomy until after case index date • equal region • year of registration is within 5 years • control is 40+ years old at case index date	751	0.8%
Total		92,070	100.0%

Abbreviations: YOB = year of birth

Appendix C Dataset specification for CPRD analyses for Aims 1 and 2

Variable	CPRD file	Description	Format	Notes
Case_status	Clinical	CASE: Any woman aged ≥ 40 years with *ovarian cancer in her patient history who has no record of an additional cancer prior to the index date of the ovarian cancer and who has ≥ 2 years of registration information in CPRD prior to index date of the ovarian cancer CONTROL: Women without ovarian cancer *frequency matched to cases up to 10:1 on year of birth (± 5 years), number of years in the CPRD (± 5 years) and the general practice of the case. Each control will be ≥ 40 years at the index date of case and have no prior history of other cancer, other than nonmelanoma skin cancer at the index date of the case, will have no prior history of both ovaries removed (oophorectomy) prior to index date, and have ≥ 2 years of registration information in CPRD prior to the index date	1= case/ 0= control	Read code list ovarian cancer - see Appendix 1 for ovarian cancer Read codes/medcodes Read code list history of other cancers-see Appendix 2 for history of other cancers Read codes/medcodes Read code list for bilateral oophorectomy

		of the case.		
		*From clinical file (independent women) *incidence-density sampling without replacement Date of first ovarian cancer Read code in patient file for cases, and corresponding match date for controls unique identifier for each patient in CPRD GOLD		
index_date_ov ary	-		date (dd/mm/yyyy)	-
patid	Patient		Integer	Available CPRD
yob	Patient	year of birth	Integer	Available CPRD
frd	Patient	first registration date; date patient first registered with the practice	date (dd/mm/yyyy)	Available CPRD
crd	Patient	date the patient's current period of registration with the practice began (date of first permanent record after latest transferred out period)	date (dd/mm/yyyy)	Available CPRD
regstat	Patient	status of registration detailing gaps and temporary patients	Integer	Available CPRD
reggap	Patient	number of days missing in the registration file	Integer	Available CPRD
deathdate	Patient	date of death of patient (CPRD algorithm)	Integer	Available CPRD
pracid	Practice	unique identifier for each specific practice in CPRD	Integer	Available CPRD
uts	Practice	date at which the practice data is deemed to be of research quality	date (dd/mm/yyyy)	Available CPRD
region	Practice	where the UK practice is based	Integer	Available CPRD

bmi_1yr-5yr	Clinical, referral or test	BMI most closely preceding 1,2,3,4, 5 years prior to index date	Categorical: 1 = low, 2 = normal, 3 = overweight, 4 = obese	Categorize using: 10-18.49=Low 18.5-24.99=Normal 25-29.99=Overweight 30+ =Obese
smoke_stat	Clinical, referral, test or entity	smoking status prior to and up to index date	Categorical: 1 = current, 2 = never, 3 = former, . = missing/unknown	available per IMS (DCEG algorithm)
alcohol_status_mr	Clinical, and entity	alcohol status prior to diagnosis	Categorical: 0 = never, 1 = former, 2 = current, . = unknown/missing	available per IMS (DCEG algorithm)
race_black	unk	ever had a medcode for black race prior to index date	0 = no, 1 = yes, . = unk/missing	available from IMS/ used in previous analyses
race_asian	unk	ever had a medcode for asian race prior to index date	0 = no, 1 = yes, . = unk/missing	available from IMS/ used in previous analyses
race_white	unk	ever had a medcode for white race prior to index date	0 = no, 1 = yes, . = unk/missing	available from IMS/ used in previous analyses
race_other	unk	ever had a medcode for other race prior to index date	0 = no, 1 = yes, . = unk/missing	available from IMS/ used in previous analyses
Aspirin_OTC_ever	unk	any OTC aspirin use prior to index date	0 = no, 1 = yes, . = unk/missing	available from IMS/ used in previous analyses
Aspirin_proph_ever	unk	any prophylactic aspirin use prior to index date	0 = no, 1 = yes, . = unk/missing	available from IMS/ used in previous analyses
oc_use_ever_BNF	Therapy	ever use of an OC prior to index date	0 = no, 1 = yes, 9 = unk/missing	-

oc_use_ever_Read	Clinical, referral	ever use of an OC prior to index date	0 = no, 1 = yes, 9 = unk/missing	-
family_history_ovca	unk	relative with ovarian cancer	0 = no, 1 = yes, 9 = unk/missing	-
parity	Pregnancy register	number of births	Integer	linkage
menstat	-	menopausal status at index date	0 = pre-menopause, 1 = post-menopause	≤ 49 years = premenopausal, > 49 years = postmenopausal at index date
diabetes_t1	Clinical, referral	diabetes type 1 prior to index date	0 = no, 1 = yes, . = unk/missing	available from IMS/ used in previous analyses;
diabetes_t2	Clinical, referral	diabetes type 2 prior to index date	0 = no, 1 = yes, . = unk/missing	available from IMS/ used in previous analyses
diabetes_unspec	Clinical, referral	diabetes unspecified prior to index date	0 = no, 1 = yes, . = unk/missing	available from IMS/ used in previous analyses
dyslipidemia	Clinical, referral	dyslipidemia prior to index date	0 = no, 1 = yes, . = unk/missing	available from IMS/ used in previous analyses
pcos_ever	Clinical, referral	polycystic ovarian syndrome prior to index date	0 = no, 1 = yes, . = unk/missing	available from IMS/ used in previous analyses
cong_heart_failure	Clinical, referral	congestive heart failure prior to index date	0 = no, 1 = yes, . = unk/missing	available from IMS/ used in previous analyses
heart_disease	Clinical, referral	ischemic heart disease prior to index date	0 = no, 1 = yes, . = unk/missing	available from IMS/ used in previous analyses
hypertension	Clinical, referral	hypertension prior to index date	0 = no, 1 = yes, . = unk/missing	available from IMS/ used in previous analyses
betablock_ever	Therapy	use of a beta blocker prior to index date	0 = no, 1 = yes, 9 = unk/missing	Read code list beta blockers

Insulin_ever	Therapy	use of insulin prior to index date	0 = no, 1 = yes, 9 = unk/missing	Read code list insulin
sulfonylureas_ever	Therapy	Use of sulfonylureas prior to index date	0 = no, 1 = yes, 9 = unk/missing	Read code list sulfonylureas
cardiacgly_ever	Therapy	use of a cardiac glycoside prior to index date	0 = no, 1 = yes, 9 = unk/missing	Read code list cardiac glycosides
statin_ever_1	Therapy	1+ prescriptions for statin drugs up to 1 year preceding index date	0 = no, 1 = yes, 9 = unk/missing	Product code list
statin_ever_2	Therapy	2+ prescriptions for statin drugs up to 1 year preceding index date	0 = no, 1 = yes, 9 = unk/missing	Product code list
hydrophilic_ever_1	Therapy	1+ prescriptions for either pravastatin or rosuvastatin up to 1 year preceding index date	0 = no, 1 = yes, 9 = unk/missing	Product code list
lipophilic_ever_1	Therapy	1+ prescriptions for either atorvastatin or simvastatin or fluvastatin up to 1 year preceding index date	0 = no, 1 = yes, 9 = unk/missing	Product code list
hydrophilic_ever_2	Therapy	2+ prescriptions for either pravastatin or rosuvastatin up to 1 year preceding index date	0 = no, 1 = yes, 9 = unk/missing	Product code list
lipophilic_ever_2	Therapy	2+ prescriptions for either atorvastatin or simvastatin or fluvastatin up to 1 year preceding index date	0 = no, 1 = yes, 9 = unk/missing	Product code list
atorvastatin_ever	Therapy	2+ prescriptions of atorvastatin up to 1 year preceding index date	0 = no, 1 = yes, 9 = unk/missing	Product code list
simvastatin_ever	Therapy	2+ prescriptions of simvastatin up to 1 year preceding index date	0 = no, 1 = yes, 9 = unk/missing	Product code list
pravastatin_ever	Therapy	2+ prescriptions of pravastatin up to 1	0 = no, 1 = yes, 9 = unk/missing	Product code list

fluvastatin_ever	Therapy	year preceding index date 2+ prescriptions of fluvastatin up to 1 year preceding index date	0 = no, 1 = yes, 9 = unk/missing	Product code list
rosuvastatin_ever	Therapy	2+ prescriptions of rosuvastatin up to 1 year preceding index date	0 = no, 1 = yes, 9 = unk/missing	Product code list
statin_individual_cat	Therapy	Categorical variable for statin use distribution	0 = none, 1 = atorvastatin, 2 = simvastatin, 3 = pravastatin, 4 = fluvastatin, 5 = rosuvastatin, 6 = multiple, 7 = unknown/missing	-
statin_duration	Therapy	Total time on statin medications (date of last statin prescription - date of first statin prescription) - months	Integer	-
atorvastatin_duration	Therapy	Total time on atorvastatin (date of last statin prescription - date of first statin prescription) - months	Integer	-
simvastatin_duration	Therapy	Total time on simvastatin (date of last statin prescription - date of first statin prescription) - months	Integer	-
pravastatin_duration	Therapy	Total time on pravastatin (date of last statin prescription - date of first statin prescription) - months	Integer	-
fluvastatin_duration	Therapy	Total time on fluvastatin (date of last statin prescription - date of first statin prescription) - months	Integer	-
rosuvastatin_duration	Therapy	Total time on rosuvastatin (date of last statin prescription	Integer	-

statin_ddd_cumul	Therapy	- date of first statin prescription) - months Cumulative DDD for all statin use across all statins (based on DDD assignments for each prescription. DDD prescription 1 + DDD prescription 2 + DDD prescription n). Up to 1-year preceding index date	Integer	See DDD/PDD definition sheet
atorvastatin_ddd_cumul	Therapy	Cumulative DDD for atorvastatin use across all statins (based on DDD assignments for each prescription. DDD prescription 1 + DDD prescription 2 + DDD prescription n). Up to 1-year preceding index date	Integer	See DDD/PDD definition sheet
simvastatin_ddd_cumul	Therapy	Cumulative DDD for simvastatin use across all statins (based on DDD assignments for each prescription. DDD prescription 1 + DDD prescription 2 + DDD prescription n). Up to 1-year preceding index date	Integer	See DDD/PDD definition sheet
pravastatin_ddd_cumul	Therapy	Cumulative DDD for pravastatin use across all statins (based on DDD assignments for each prescription. DDD prescription 1 + DDD prescription 2 + DDD prescription n). Up to 1-year preceding index date	Integer	See DDD/PDD definition sheet
fluvastatin_ddd_cumul	Therapy	Cumulative DDD for fluvastatin use across all statins (based on DDD assignments for	Integer	See DDD/PDD definition sheet

		each prescription. DDD prescription 1 + DDD prescription 2 + DDD prescription n). Up to 1-year preceding index date Cumulative DDD for rosuvastatin use across all statins (based on DDD assignments for each prescription. DDD prescription 1 + DDD prescription 2 + DDD prescription n). Up to 1-year preceding index date		
rosuvastatin_d dd_cumul	Therapy		Integer	See DDD/PDD definition sheet
statin_pdd_mis s_prescripts	Therapy	# of prescriptions in patient file missing the PDD prior to index date	Integer	See DDD/PDD definition sheet
atorvastatin_pd d_miss_prescri pts	Therapy	# of prescriptions in patient file missing the PDD up to 1 year prior to index date	Integer	See DDD/PDD definition sheet
simvastatin_pd d_miss_prescri pts	Therapy	# of prescriptions in patient file missing the PDD up to 1 year prior to index date	Integer	See DDD/PDD definition sheet
pravastatin_pd d_miss_prescri pts	Therapy	# of prescriptions in patient file missing the PDD up to 1 year prior to index date	Integer	See DDD/PDD definition sheet
fluvastatin_pdd _miss_prescrip ts	Therapy	# of prescriptions in patient file missing the PDD up to 1 year prior to index date	Integer	See DDD/PDD definition sheet
rosuvastatin_p dd_miss_presc ripts	Therapy	# of prescriptions in patient file missing the PDD up to 1 year prior to index date	Integer	See DDD/PDD definition sheet
statin_prescript s	Therapy	Number of statin prescriptions in patient record up to 1-year preceding index date	Integer	
atorvastatin_pr escripts	Therapy	Number of atorvastatin	Integer	

simvastatin_prescripts	Therapy	prescriptions in patient record up to 1-year preceding index date Number of simvastatin prescriptions in patient record up to 1-year preceding index date	Integer	
pravastatin_prescripts	Therapy	Number of pravastatin prescriptions in patient record up to 1-year preceding index date	Integer	
fluvastatin_prescripts	Therapy	Number of fluvastatin prescriptions in patient record up to 1-year preceding index date	Integer	
rosuvastatin_prescripts	Therapy	Number of rosuvastatin prescriptions in patient record up to 1-year preceding index date	Integer	
metformin_event_1	Therapy	1+ prescriptions for metformin up to 1-year preceding index date	0 = no, 1 = yes, 9 = unk/missing	Product code list - see Appendix 5 for metformin product codes Product code list - see Appendix 5 for metformin product codes
metformin_event_2	Therapy	2+ prescriptions for metformin up to 1 year preceding- index date	0 = no, 1 = yes, 9 = unk/missing	
metformin_duration	Therapy	Total time on metformin (date of last metformin prescription - date of first metformin prescription) - months Cumulative DDD for metformin use across all statins (based on DDD assignments for each prescription.	Integer	-
metformin_ddd_cumul	Therapy	DDD prescription 1 + DDD prescription 2 + DDD prescription n). Up to 1-year preceding index date	Integer	-

metformin_pdd _miss_prescrip ts	Therapy	# of prescriptions in patient file missing the PDD up to 1 year prior to index date	Integer	-
metformin_pre scripts	Therapy	Number of metformin prescriptions in patient record up to 1-year preceding index date	Integer	See DDD/PDD definition sheet

Abbreviations: YOB = year of birth; CPRD = clinical practice research datalink; BMI = body mass index; UK = United Kingdom; unk = unknown; OC = oral contraceptive; DDD = defined daily dose; PDD = prescribed daily dose

Appendix D Defined Daily Doses Algorithm

WHO 2020 Defined Daily Dose (DDD) assignments for statins and Metformin					
ATC Code	Drug Name	DDD	mg	route	
C10AA05	Atorvastatin	20	mg	oral	
C10AA01	Simvastatin	30	mg	oral	
C10AA04	Fluvastatin	60	mg	oral	
C10AA03	Pravastatin	30	mg	oral	
C10AA07	Rosuvastatin	10	mg	oral	
C10AA02	Lovastatin	45	mg	oral	
C10AA06	Cerivastatin	0.2	mg	oral	
A10BA02	Metformin	2	g	oral	

DDD Calculation:

**** Prior to DDD determinations, it may be necessary to standardize units for the medication strengths*******

Step 1: Determine DDD conversion factor

- WHO-defined DDD / strength in patient record (assuming units match) = DDD conversion factor

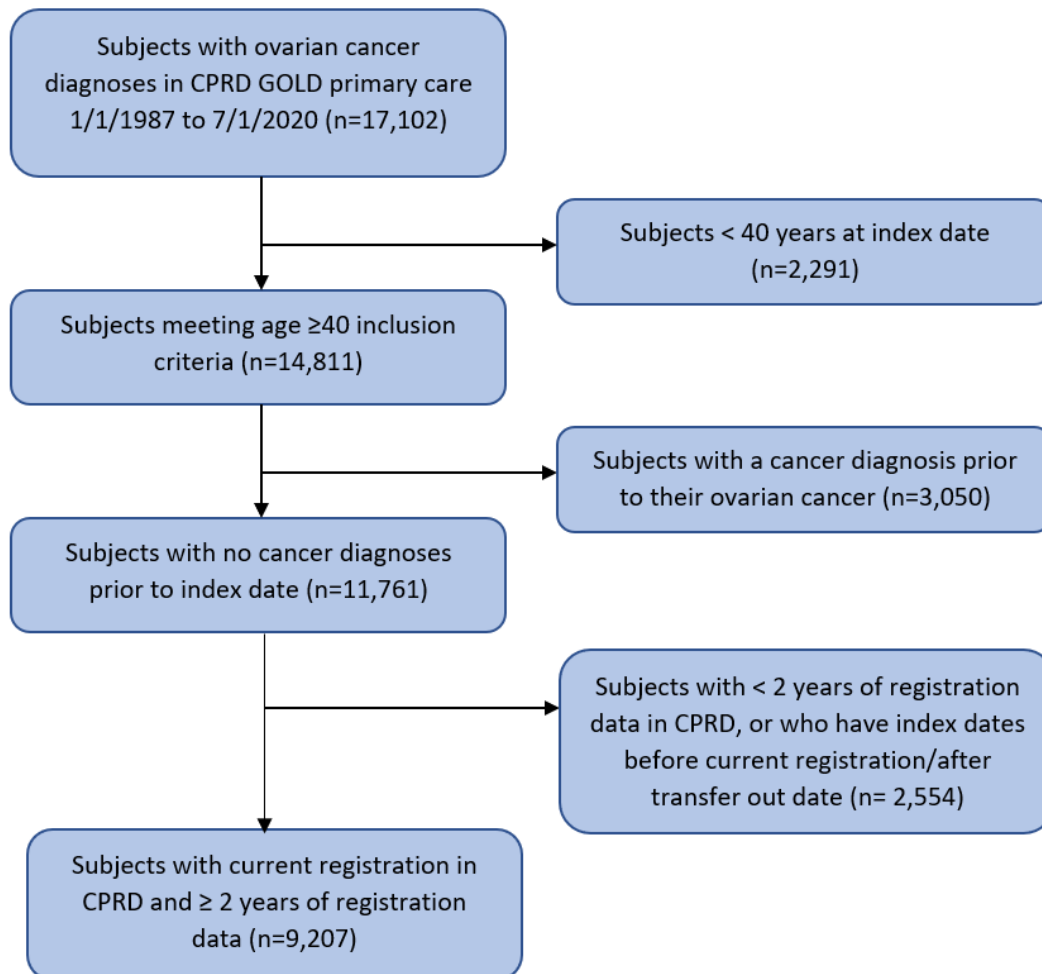
Step 2: Multiply the quantity of each prescription by this conversion factor

- DDD conversion factor (Step 1) * quantity per prescription = DDD for *that* prescription

Step 3: Cumulative DDDs

- Sum DDD's across prescriptions for each medication per patient
- For statin-specific cumulative DDDs, sum across specific statins (e.g., all atorvastatin prescriptions will be the cumulative atorvastatin value for that patient)

Supplemental Figure 1. Summary of case selection in CPRD



Supplemental Table 1. Distribution of case diagnosis dates by birth cohort

Diagnosis year	Birth Cohort								Total
	1899-1910	1911-1920	1921-1930	1931-1940	1941-1950	1951-1960	1961-1970	1971-1979	
1987	5	2	13	10	9	0	0	0	39
	4.1	0.38	0.8	0.44	0.36	0	0	0	0.4%
1988	4	14	22	18	6	0	0	0	64
	3.28	2.69	1.36	0.79	0.24	0	0	0	0.7%
1989	9	12	11	24	22	0	0	0	78
	7.38	2.3	0.68	1.06	0.88	0	0	0	0.8%
1990	11	32	48	39	15	0	0	0	145
	9.02	6.14	2.97	1.72	0.6	0	0	0	1.6%
1991	13	25	42	34	33	0	0	0	147
	10.66	4.8	2.6	1.5	1.33	0	0	0	1.6%
1992	20	37	34	39	26	2	0	0	158
	16.39	7.1	2.1	1.72	1.05	0.14	0	0	1.7%
1993	15	27	51	48	38	4	0	0	183
	12.3	5.18	3.16	2.11	1.53	0.27	0	0	2.0%
1994	17	23	51	32	42	10	0	0	175
	13.93	4.41	3.16	1.41	1.69	0.68	0	0	1.9%
1995	9	37	58	69	53	12	0	0	238
	7.38	7.1	3.59	3.04	2.13	0.82	0	0	2.6%
1996	7	22	54	46	50	20	0	0	199
	5.74	4.22	3.34	2.03	2.01	1.36	0	0	2.2%
1997	6	27	55	69	59	21	0	0	237
	4.92	5.18	3.4	3.04	2.37	1.43	0	0	2.6%
1998	2	32	53	66	60	36	0	0	249
	1.64	6.14	3.28	2.91	2.41	2.45	0	0	2.7%
1999	4	26	83	90	71	29	0	0	303
	3.28	4.99	5.14	3.96	2.85	1.97	0	0	3.3%
2000	0	24	67	83	98	44	0	0	316
	0	4.61	4.15	3.66	3.94	2.99	0	0	3.4%
2001	0	30	98	88	85	50	0	0	351

	0	5.76	6.06	3.88	3.42	3.4	0	0	3.8%
2002	0	33	76	100	93	48	6	0	356
	0	6.33	4.7	4.41	3.74	3.26	0.98	0	3.9%
2003	0	29	85	108	94	54	9	0	379
	0	5.57	5.26	4.76	3.78	3.67	1.47	0	4.1%
2004	0	24	91	126	105	60	10	0	416
	0	4.61	5.63	5.55	4.22	4.08	1.64	0	4.5%
2005	0	24	75	111	117	74	24	0	425
	0	4.61	4.64	4.89	4.7	5.03	3.93	0	4.6%
2006	0	17	95	98	120	80	15	0	425
	0	3.26	5.88	4.32	4.83	5.43	2.45	0	4.6%
2007	0	11	68	126	132	67	21	0	425
	0	2.11	4.21	5.55	5.31	4.55	3.44	0	4.6%
2008	0	8	76	96	127	77	42	0	426
	0	1.54	4.7	4.23	5.11	5.23	6.87	0	4.6%
2009	0	5	58	103	134	71	39	0	410
	0	0.96	3.59	4.54	5.39	4.82	6.38	0	4.5%
2010	0	0	50	90	119	70	58	0	387
	0	0	3.09	3.96	4.78	4.76	9.49	0	4.2%
2011	0	0	64	117	98	77	44	3	403
	0	0	3.96	5.15	3.94	5.23	7.2	2.78	4.4%
2012	0	0	37	86	111	78	47	4	363
	0	0	2.29	3.79	4.46	5.3	7.69	3.7	3.9%
2013	0	0	35	84	90	89	45	8	351
	0	0	2.17	3.7	3.62	6.05	7.36	7.41	3.8%
2014	0	0	28	59	109	71	42	8	317
	0	0	1.73	2.6	4.38	4.82	6.87	7.41	3.4%
2015	0	0	14	53	103	63	38	15	286
	0	0	0.87	2.33	4.14	4.28	6.22	13.89	3.1%
2016	0	0	13	48	63	74	46	10	254
	0	0	0.8	2.11	2.53	5.03	7.53	9.26	2.8%
2017	0	0	7	41	73	62	47	14	244
	0	0	0.43	1.81	2.94	4.21	7.69	12.96	2.7%

2018	0	0	2	35	56	51	39	18	201
	0	0	0.12	1.54	2.25	3.46	6.38	16.67	2.2%
2019	0	0	2	31	54	52	30	14	183
	0	0	0.12	1.37	2.17	3.53	4.91	12.96	2.0%
2020	0	0	0	3	22	26	9	14	74
	0	0	0	0.13	0.88	1.77	1.47	12.96	0.8%
Total	122	521	1616	2270	2487	1472	611	108	9207

Supplemental Table 2. Characteristics of ovarian cancer cases and controls in CPRD

	1899-1910		1911-1920		1921-1930		1931-1940		1941-1950		1951-1960		1961-1970		1971-1980		1899-1979	
	Cases n=122	Con trols n=1150	Cases n=521	Con trols n=5159	Cases n=1616	Con trols n=16134	Cases n=2270	Con trols n=22728	Cases n=2487	Con trols n=24974	Cases n=1472	Con trols n=14698	Cases n=611	Con trols n=6179	Cases n=108	Con trols n=1048	Cases n=9207	Con trols n=92070
Age at index	85.2 ± 3.1	85.4 ± 3.3	80.8 ± 5.4	80.8 ± 5.4	76.0 ± 7.2	76.0 ± 7.1	69.0 ± 7.8	69.0 ± 7.8	60.8 ± 7.8	60.8 ± 7.8	53.9 ± 6.7	53.8 ± 6.6	47.9 ± 4.5	47.9 ± 4.6	43.3 ± 2.6	43.3 ± 2.5	64.7 ± 12.0	64.7 ± 12.0
Year at diagnosis	1992 (1987-1999)	NA	1997 (1987-2009)	NA	2002 (1987-2019)	NA	2004 (1987-2020)	NA	2006 (1987-2020)	NA	2008 (1992-2020)	NA	2012 (2002-2020)	NA	2016 (2011-2020)	NA	2005 (1987-2020)	NA
Time in CPRD (years)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	22.49 ± 16.7	22.49 ± 16.8
GP (num practices)																	938	938
OC Use																		
Ever	0 (0%)	0 (0%)	0 (0%)	1 (0.01%)	1 (0.1%)	15 (0.1%)	15 (0.6%)	109 (0.5%)	81 (3.3%)	850 (3.4%)	176 (11.9%)	2320 (15.8%)	161 (26.4%)	2085 (33.7%)	36 (33.3%)	444 (42.4%)	470 (5.1%)	5824 (6.3%)
Never	0 (0%)	0 (0%)	521 (100%)	5158 (99.9%)	1615 (99.9%)	1611 (99.9%)	2255 (99.4%)	2261 (99.5%)	2406 (96.7%)	2412 (96.6%)	1296 (88.0%)	1237 (84.2%)	450 (73.6%)	4094 (66.3%)	72 (66.7%)	604 (57.6%)	8737 (94.9%)	8624 (93.7%)
Family History of Ovarian Cancer																		
Ever	0 (0%)	0 (0%)	0 (0%)	1 (0.01%)	3 (0.2%)	1 (0.01%)	10 (0.4%)	8 (0.04%)	12 (0.5%)	23 (0.1%)	6 (0.4%)	24 (0.2%)	5 (0.8%)	17 (0.3%)	1 (0.9%)	1 (0.1%)	37 (0.4%)	75 (0.1%)
Never	0 (0%)	0 (0%)	521 (100%)	5158 (99.9%)	1613 (99.8%)	1613 (99.9%)	2260 (99.6%)	2272 (99.9%)	2475 (99.5%)	2495 (99.9%)	1466 (99.6%)	1467 (99.8%)	606 (99.2%)	6162 (99.7%)	107 (99.1%)	1047 (99.9%)	9170 (99.6%)	9199 (99.9%)
Parity																		
1 pregnancy	1 (0.8%)	10 (0.9%)	14 (2.7%)	198 (3.8%)	99 (6.1%)	1287 (8.0%)	275 (12.1%)	2689 (11.8%)	318 (1.3%)	3412 (13.7%)	236 (16.6%)	2244 (15.3%)	115 (18.8%)	1023 (16.6%)	24 (22.2%)	159 (15.2%)	1082 (11.7%)	1102 (11.9%)

2 pregnancies	0 (0%)	2 (0.2%)	7 (1.3%)	45 (0.9%)	53 (3.3%)	642 (4.0%)	166 (7.3%)	1953 (8.6%)	316 (1.3%)	3450 (13.9%)	229 (15.6%)	2540 (17.3%)	114 (18.8%)	130 (2.1%)	20 (18.5%)	235 (22.4%)	905 (9.8%)	1016 8 (11.0%)
3 pregnancies	0 (0%)	0 (0%)	1 (0.2%)	30 (0.6%)	23 (1.4%)	268 (1.7%)	93 (4.1%)	992 (4.4%)	151 (6.1%)	1822 (7.2%)	145 (9.9%)	1558 (10.6%)	78 (12.8%)	864 (14.0%)	13 (12.0%)	176 (16.8%)	504 (5.5%)	5710 (6.2%)
4 pregnancies	0 (0%)	0 (0%)	1 (0.2%)	8 (0.2%)	7 (0.4%)	132 (0.8%)	37 (1.6%)	416 (1.8%)	71 (2.9%)	771 (3.1%)	67 (4.6%)	767 (5.2%)	44 (7.2%)	496 (8.0%)	6 (5.6%)	98 (9.4%)	233 (2.5%)	2688 (2.9%)
4+ pregnancies	0 (0%)	0 (0%)	0 (0%)	7 (0.1%)	3 (0.2%)	102 (0.6%)	24 (1.1%)	274 (1.2%)	36 (1.4%)	453 (1.8%)	25 (1.7%)	578 (3.9%)	31 (5.1%)	553 (8.9%)	9 (8.3%)	159 (15.2%)	128 (1.4%)	2126 (2.3%)
Missing/ nulliparous	121 (99.2%)	1138 (99.1%)	498 (95.6%)	4871 (94.4%)	1431 (88.6%)	1370 3 (84.9%)	1675 (73.7%)	1640 4 (72.2%)	1595 (64.1%)	1506 6 (60.3%)	770 (52.3%)	7011 (47.8%)	229 (37.5%)	1942 (31.4%)	36 (33.3%)	221 (21.1%)	6355 (69.0%)	6036 5 (65.5%)
BMI																		
Low	1 (0.8%)	7 (0.6%)	7 (1.3%)	108 (2.1%)	44 (2.7%)	465 (2.9%)	53 (2.3%)	620 (2.7%)	69 (2.8%)	768 (3.1%)	59 (4.0%)	680 (4.6%)	16 (2.6%)	358 (5.8%)	6 (5.6%)	67 (6.4%)	255 (2.8%)	3073 (3.3%)
Normal	4 (3.3%)	24 (2.1%)	62 (11.9%)	570 (11.0%)	290 (17.9%)	2924 (18.1%)	450 (19.8%)	4920 (21.6%)	581 (23.4%)	6121 (24.5%)	365 (24.8%)	4225 (28.7%)	181 (29.6%)	1965 (31.8%)	33 (30.6%)	351 (33.5%)	1966 (21.4%)	2110 0 (22.9%)
Overweight	2 (1.6%)	19 (1.7%)	63 (12.1%)	538 (10.4%)	326 (20.2%)	2969 (18.4%)	553 (24.4%)	5246 (23.1%)	563 (22.6%)	5611 (22.5%)	291 (19.8%)	3075 (20.9%)	134 (21.9%)	1246 (20.2%)	22 (20.3%)	224 (21.4%)	1954 (21.2%)	1892 8 (20.6%)
Obese	0 (0%)	10 (0.9%)	26 (5.0%)	262 (5.1%)	167 (10.3%)	1693 (10.5%)	392 (17.3%)	3536 (15.6%)	422 (16.9%)	3973 (15.9%)	310 (21.1%)	2532 (17.2%)	139 (22.7%)	1144 (18.5%)	31 (28.7%)	217 (20.7%)	1487 (16.2%)	1336 7 (14.5%)
Missing	115 (94.2%)	1090 (94.8%)	363 (69.7%)	3681 (71.4%)	789 (48.8%)	8083 (50.1%)	822 (36.2%)	8406 (37.0%)	852 (34.3%)	8501 (34.0%)	447 (30.4%)	4186 (28.5%)	141 (23.1%)	1466 (23.7%)	16 (14.8%)	189 (18.0%)	3545 (38.5%)	3560 2 (38.7%)
Smoking status																		
Never	2 (1.6%)	42 (3.7%)	67 (12.9%)	585 (11.3%)	321 (19.9%)	3035 (18.8%)	522 (23.0%)	5043 (22.2%)	580 (23.3%)	5728 (22.9%)	427 (29.0%)	4045 (27.5%)	206 (33.7%)	2110 (34.1%)	44 (40.7%)	408 (38.9%)	2169 (23.6%)	2099 6 (22.8%)

	Former	44 (36.1)	365 (31.7%)	216 (41.5)	2148 (41.6%)	773 (47.8)	7434 (46.1)	1119 (49.3)	1093 3 (48.1)	1177 (47.3)	1146 4 (45.9)	622 (42.3)	6199 (42.2)	175 (2.8%)	1526 (24.7%)	37 (34.3)	376 (35.9%)	4211 (45.7)	4129 2 (44.8)
	Current	4 (3.3%)	22 (1.9)	40 (7.7%)	366 (7.1)	207 (12.8)	1985 (12.3)	326 (14.4)	3402 (14.9)	480 (19.3)	4945 (19.8)	351 (23.8)	3402 (23.1)	175 (2.8%)	1526 (24.7%)	27 (25.0)	239 (22.8%)	1610 (17.5)	1588 7 (17.3)
	Missing	72 (59.0)	721 (62.7%)	216 (41.5)	2148 (41.6%)	315 (19.5)	3680 (22.8)	303 (13.3)	3350 (14.7)	250 (10.1)	2837 (11.4)	72 (4.9%)	1052 (7.2)	7 (0.1%)	170 (2.8)	0 (0%)	25 (2.4)	1217 (13.2)	1389 5 (15.1)
Alcohol use																			
	Never	12 (9.8%)	138 (12.0%)	89 (17.1)	869 (16.8%)	311 (19.2)	2861 (17.7)	345 (15.2)	3613 (15.9)	278 (11.2)	2785 (11.2)	136 (9.2%)	1459 (9.9)	58 (9.5%)	640 (10.4%)	15 (14.3)	103 (9.8)	1244 (13.5)	1246 8 (13.5)
	Former	1 (0.8%)	9 (0.8)	16 (3.1%)	192 (3.7)	121 (7.5%)	1144 (7.1)	173 (7.6%)	1860 (8.2)	188 (0.8%)	1710 (6.8)	98 (6.7%)	914 (6.2)	52 (8.5%)	398 (6.4)	4 (3.7%)	69 (6.6)	653 (7.1%)	6296 (6.8)
	Current	27 (22.0)	192 (16.7%)	173 (33.2)	1553 (30.1%)	726 (44.9)	6962 (43.2)	1278 (56.3)	1219 7 (53.7)	1571 (63.2)	1565 3 (62.7)	1007 (68.4)	1001 3 (68.1)	435 (71.2)	4431 (71.7%)	81 (75.0)	754 (71.9%)	5298 (57.5)	5175 5 (56.2)
	Missing	82 (67.2)	811 (70.5%)	243 (46.6)	2545 (49.3%)	458 (28.3)	5167 (32.0)	474 (20.9)	5058 (22.3)	450 (18.1)	4826 (19.3)	231 (15.7)	2312 (15.7)	66 (10.8)	710 (11.5%)	8 (7.4%)	122 (11.6%)	2012 (21.9)	2155 1 (23.4)

Abbreviations: OC = oral contraceptives; CPRD = Clinical Practice Research Datalink

Supplemental Table 3. Association between metformin use and ovarian cancer risk in CPRD by groups defined according to historical CPRD database changes

	Cases N=9,207	Controls N=92,070	OR (95% CI)
<i>CPRD index dates 1987-1992</i>			
metformin			
Ever	3 (0.5%)	15 (0.2%)	1.77 (0.50-6.20)
Never	628 (99.5%)	6295 (99.8%)	-
<i>Total</i>	<i>631</i>	<i>6310</i>	<i>-</i>
<i>CPRD index dates 1993-2012</i>			
metformin			
Ever	231 (3.5%)	2223 (3.3%)	0.99 (0.86-1.14)
Never	6435 (96.5%)	64437 (96.7%)	-
<i>Total</i>	<i>6666</i>	<i>66660</i>	<i>-</i>
<i>CPRD index dates 2012-2019</i>			
metformin			
Ever	138 (7.5%)	1202 (6.5%)	1.03 (0.85-1.25)
Never	1698 (92.5%)	17158 (93.4%)	-
<i>Total</i>	<i>1836</i>	<i>18360</i>	<i>-</i>

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Supplemental Table 4. Association between metformin use and ovarian cancer risk in CPRD by groups defined according to density of cases in birth cohorts

	Cases N=9,207	Controls N=92,070	OR (95% CI)
<i>CPRD index dates 1987-2000</i>			
metformin			
Ever	27 (1.1%)	213 (0.8%)	1.23 (0.82-1.84)
Never	2504 (98.9%)	25097 (99.2%)	-
Total	2531	25310	-
<i>CPRD index dates 2001-2009</i>			
metformin			
Ever	129 (3.6%)	1319 (3.7%)	0.95 (0.79-1.15)
Never	3484 (96.4%)	34811 (96.3%)	-
Total	3613	36130	-
<i>CPRD index dates 2010-2019</i>			
metformin			
Ever	216 (7.2%)	1908 (6.4%)	1.02 (0.88-1.18)
Never	2773 (92.8%)	27982 (93.6%)	-
Total	2989	29890	-

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Supplemental Table 5. Association between BMI and ovarian cancer risk in CPRD

	Cases N=9,207	Controls N=92,070	OR (95% CI)
BMI (kg/m²) at least 5 years prior to diagnosis			
Low <18.5	255 (2.8%)	3073 (3.3%)	0.89 (0.78-1.02)
Normal 18.5-24.9	1966 (21.4%)	21100 (22.9%)	(ref)
Overweight 25-29.9	1954 (21.2%)	18928 (20.6%)	1.10 (1.03-1.18)
Obese 30+	1487 (16.2%)	13367 (14.5%)	1.19 (1.11-1.28)
missing	3545 (38.5%)	35602 (38.7%)	1.05 (0.99-1.11)

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; ref = reference; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Supplemental Table 6. Association between BMI and metformin indications in CPRD

	Type 2 diabetes Cases N= 6517	Type 2 diabetes Controls N=94,760	OR* (95% CI) (diabetes)
BMI (kg/m²) 5 years before index			
Low <18.5	61 (0.9%)	3267 (3.4%)	0.57 (0.44-0.74)
Normal 18.5-24.9	733 (11.2%)	22333 (23.6%)	(ref)
Overweight 25-29.9	1575 (24.2%)	19307 (20.4%)	2.49 (2.28-2.72)
Obese 30+	2835 (43.5%)	12019 (12.7%)	7.18 (6.60-7.82)
<i>missing</i>	1313 (20.1%)	37834 (39.9%)	1.06 (0.97-1.16)
BMI (kg/m²) 5 years before index	PCOS Cases N= 276	PCOS Controls N= 10,100	OR* (95% CI) (PCOS)
Low <18.5	15 (5.4%)	3313 (32.8%)	1.83 (1.04-3.23)
Normal 18.5-24.9	57 (20.7%)	23009 (22.8%)	(ref)
Overweight 25-29.9	59 (21.4%)	20823 (20.6%)	1.14 (0.79-1.64)
Obese 30+	87 (31.5%)	14767 (14.6%)	2.32 (1.65-3.28)
<i>missing</i>	58 (21.0%)	39089 (38.7%)	0.60 (0.42-0.86)

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; ref = reference; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Supplemental Table 7. Association between metformin indications and ovarian cancer risk in CPRD

	Cases N=9,207	Controls N=92,070	OR* (95% CI)
Diabetes Type 2			
Ever	610 (6.6%)	5907 (6.4%)	0.98 (0.89-1.07)
Never	8597 (93.4%)	86163 (93.6%)	(ref)
Diabetes Type 2 (removing 0-3 months)			
Ever	609 (6.6%)	5901 (6.4%)	0.97 (0.89-1.06)
Never	8580 (93.2%)	86024 (93.4%)	(ref)
PCOS			
Ever	37 (0.4%)	239 (0.3%)	1.66 (1.17-2.35)
Never	9107 (96.6%)	91831 (99.7%)	(ref)

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; ref = reference; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Supplemental Table 8. Timing of metformin use and sulfonylureas use with ovarian cancer risk in CPRD 1987-2020

	Cases N=9,207	Controls N=92,070	OR (95% CI)
Timing of prescriptions			
Sulfonylureas before metformin	91	772	1.18 (0.93-1.50)
Sulfonylureas prescribed the same day as metformin	19	197	1.09 (0.88-1.36)
Sulfonylureas prescribed after metformin	116	1032	0.96 (0.59-1.55)
No use of metformin or Sulfonylureas	8981	90069	(ref)

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; ref = reference; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Appendix F Supplemental material for manuscript 2

Supplemental Table 1. Association between statin use and ovarian cancer risk in CPRD by groups defined according to historical CPRD database changes

	Cases N=9,207	Controls N=92,070	OR* (95% CI)
<i>CPRD index dates 1987-1992</i>			
Statins			
Ever	0 (0%)	2 (0.03%)	-
Never	631 (100%)	6308 (99.9%)	-
<i>Total</i>	<i>631</i>	<i>6310</i>	-
<i>CPRD index dates 1993-2012</i>			
Statins			
Ever	1046 (15.7%)	9906 (14.9%)	1.03 (0.96-1.11)
Never	5620 (84.3%)	56754 (85.1%)	Ref
<i>Total</i>	<i>6666</i>	<i>66660</i>	-
<i>CPRD index dates 2012-2019</i>			
Statins			
Ever	608 (33.1%)	6053 (33.0%)	0.92 (0.82-1.02)
Never	1228 (66.9%)	12307 (67.0%)	Ref
<i>Total</i>	<i>1836</i>	<i>18360</i>	-

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; ref = reference; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Supplemental Table 2. Association between statin use and ovarian cancer risk in CPRD by groups defined according to density of cases in birth cohorts

	Cases N=9,207	Controls N=92,070	OR* (95% CI)
<i>CPRD index dates 1987-2000</i>			
Statins			
Ever	31 (1.2%)	307 (1.2%)	0.97 (0.67-1.41)
Never	2500 (98.8%)	25003 (98.8%)	Ref
<i>Total</i>	<i>2531</i>	<i>25310</i>	-
<i>CPRD index dates 2001-2009</i>			
Statins			
Ever	646 (17.9%)	6091 (16.9%)	1.05 (0.96-1.15)
Never	2967 (82.1%)	30039 (83.1%)	Ref
<i>Total</i>	<i>3613</i>	<i>36130</i>	-
<i>CPRD index dates 2010-2019</i>			
Statins			
Ever	977 (32.7%)	9563 (32.0%)	0.95 (0.87-1.03)
Never	2012 (67.3%)	20327 (68.0%)	Ref
<i>Total</i>	<i>2989</i>	<i>29890</i>	-

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; ref = reference; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Supplemental Table 3. Association between cumulative prescription counts for individual statins and ovarian cancer risk

	Cases N=9,207	Controls N=92,070	OR* (95% CI)
Simvastatin prescriptions			
Never use	7930 (86.1%)	79629 (86.5%)	(ref)
1-8	430 (4.7%)	4016 (4.4%)	1.04 (0.93-1.15)
9-22	270 (2.9%)	2808 (3.0%)	0.93 (0.82-1.06)
23-48	300 (3.3%)	2842 (3.1%)	1.01 (0.89-1.14)
49+	277 (3.0%)	2775 (3.0%)	0.94 (0.83-1.07)
Atorvastatin prescriptions			
Never use	8481 (92.1%)	85390 (92.7%)	(ref)
1-9	258 (2.8%)	2236 (2.4%)	1.13 (0.99-1.28)
10-21	134 (1.5%)	1403 (1.5%)	0.93 (0.78-1.11)
22-44	163 (1.8%)	1544 (1.7%)	1.02 (0.86-1.20)
44+	171 (1.9%)	1497 (1.6%)	1.08 (0.92-1.27)
Pravastatin prescriptions			
Never use	9029 (98.1%)	90242 (98.0%)	(ref)
1-6	61 (0.7%)	681 (0.7%)	0.87 (0.67-1.13)
7-15	37 (0.4%)	366 (0.4%)	0.97 (0.69-1.36)
16-36	44 (0.5%)	394 (0.4%)	1.07 (0.78-1.46)
36+	36 (0.4%)	387 (0.4%)	0.88 (0.62-1.24)
Fluvastatin prescriptions			
Never use	9163 (99.5%)	91586 (99.5%)	(ref)
1-5	24 (0.3%)	190 (0.2%)	1.22 (0.80-1.86)
6-14	10 (0.1%)	101 (0.1%)	0.93 (0.49-1.79)
15-30	1 (<0.1%)	95 (0.1%)	-
30+	9 (<0.1%)	98 (0.1%)	0.87 (0.44-1.72)
Rosuvastatin prescriptions			
Never use	9083 (98.6%)	90975 (99.0%)	(ref)
1-5	43 (0.5%)	379 (0.4%)	1.10 (0.80-1.51)
6-18	28 (0.4%)	252 (0.3%)	1.08 (0.73-1.59)
19-39	29 (0.4%)	226 (0.3%)	1.23 (0.83-1.81)
40+	24 (0.3%)	238 (0.3%)	0.95 (0.63-1.45)

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; ref = reference; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Supplemental Table 4. Association between individual statin durations and ovarian cancer risk

	Cases n=9207	Controls N=92070	OR* (95% CI)
Simvastatin			
<1 year (low use)	163 (11.8%)	1605 (1.7%)	0.98 (0.84-1.16)
1-3 years	254 (2.8%)	2426 (2.6%)	1.02 (0.89-1.16)
3-5 years	213 (2.3%)	2085 (2.3%)	0.98 (0.85-1.13)
5-10 years	341 (3.7%)	3459 (3.8%)	0.94 (0.83-1.05)
10+ years	189 (2.1%)	1779 (1.9%)	1.00 (0.86-1.17)
Never	7536 (81.9%)	75901 (82.4%)	(ref)
Atorvastatin			
<1 year (low use)	78 (<0.1%)	605 (<0.1%)	1.24 (0.98-1.58)
1-3 years	117 (1.3%)	1131 (1.2%)	1.00 (0.82-1.21)
3-5 years	107 (1.2%)	1044 (1.1%)	0.98 (0.80-1.20)
5-10 years	212 (2.3%)	1975 (2.1%)	1.02 (0.88-1.18)
10+ years	143 (1.6%)	1335 (1.4%)	1.01 (0.84-1.21)
Never	7536 (81.9%)	75901 (82.4%)	(ref)
Pravastatin			
<1 year (low use)	13 (0.4%)	148 (1.6%)	0.86 (0.49-1.52)
1-3 years	27 (0.3%)	279 (0.3%)	0.92 (0.62-1.37)
3-5 years	25 (0.3%)	235 (0.3%)	1.02 (0.67-1.54)
5-10 years	49 (0.5%)	530 (0.6%)	0.88 (0.65-1.17)
10+ years	37 (0.4%)	394 (0.4%)	0.89 (0.63-1.25)
Never	7536 (81.9%)	75901 (82.4%)	(ref)
Fluvastatin			
<1 year (low use)	1 (<0.1%)	30 (<0.1%)	0.32 (0.05-2.36)
1-3 years	6 (<0.1%)	55 (<0.1%)	1.04 (0.45-2.41)
3-5 years	6 (<0.1%)	68 (<0.1%)	0.84 (0.36-1.92)
5-10 years	10 (0.1%)	125 (0.1%)	0.75 (0.40-1.43)
10+ years	10 (0.1%)	118 (0.1%)	-
Never	7536 (81.9%)	75901 (82.4%)	(ref)
Rosuvastatin			
<1 year (low use)	6 (<0.1%)	51 (<0.1%)	1.16 (0.50-2.70)
1-3 years	21 (0.2%)	145 (0.2%)	1.40 (0.89-2.22)
3-5 years	18 (0.2%)	139 (0.2%)	1.24 (0.76-2.02)
5-10 years	36 (0.4%)	338 (0.4%)	1.01 (0.71-1.42)
10+ years	29 (0.3%)	283 (0.3%)	0.97 (0.66-1.42)
Never	7536 (81.9%)	75901 (82.4%)	(ref)

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; ref = reference; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Supplemental Table 5. Association between statin durations and ovarian cancer risk by statin class

	Cases N=9207	Controls N=92070	OR* (95% CI)
Hydrophilic			
<1 year (low use)	31 (0.3%)	298 (0.3%)	1.02 (0.70-1.48)
1-3 years	50 (0.4%)	463 (0.4%)	1.04 (0.77-1.29)
3-5 years	43 (0.3%)	401 (0.3%)	1.03 (0.75-1.41)
5-10 years	88 (1.0%)	891 (1.0%)	0.93 (0.75-1.17)
10+ years	63 (0.4%)	642 (0.4%)	0.93 (0.72-1.21)
Never	7536 (81.9%)	75901 (82.4%)	(ref)
Lipophilic			
<1 year (low use)	316 (3.4%)	3013 (3.3%)	1.02 (0.91-1.15)
1-3 years	357 (3.9%)	3336 (3.6%)	1.04 (0.93-1.16)
3-5 years	270 (2.9%)	2682 (2.9%)	0.96 (0.85-1.10)
5-10 years	426 (4.6%)	4237 (4.6%)	0.95 (0.86-1.06)
10+ years	224 (2.4%)	2183 (2.4%)	0.97 (0.84-1.17)
Never	7536 (81.9%)	75901 (82.4%)	(ref)

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; ref = reference; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Supplemental Table 6. Defined daily doses for statin drugs assigned by the World Health Organization (WHO)

Statin	Defined Daily Dose (mg)
Atorvastatin	20
Simvastatin	30
Fluvastatin	60
Rosuvastatin	10
Pravastatin	30

Supplemental Table 7. Association between individual statin cumulative DDDs and ovarian cancer risk

	Cases n=9207	Controls N=92070	OR* (95% CI)
Simvastatin			
1-308	220 (2.4%)	2128 (2.3%)	1.00 (0.87-1.16)
309-938	277 (3.0%)	2888 (3.1%)	0.92 (0.81-1.05)
939-2203	326 (3.5%)	3182 (3.5%)	0.98 (0.87-1.11)
2203+	337 (3.7%)	3151 (3.4%)	1.01 (0.90-1.13)
Never	7536 (81.9%)	75901 (82.4%)	(ref)
Atorvastatin			
1-308	101 (1.1%)	903 (1.0%)	1.08 (0.88-1.33)
309-938	158 (1.7%)	1375 (1.5%)	1.10 (0.93-1.30)
939-2203	169 (1.8%)	1610 (1.7%)	1.01 (0.86-1.19)
2203+	229 (2.5%)	2200 (2.4%)	0.98 (0.85-1.13)
Never	7536 (81.9%)	75901 (82.4%)	(ref)
Pravastatin			
1-308	28 (0.3%)	250 (0.3%)	1.09 (0.73-1.61)
309-938	33 (0.3%)	363 (0.4%)	0.87 (0.61-1.25)
939-2203	36 (0.3%)	412 (0.4%)	0.83 (0.59-1.18)
2203+	54 (0.6%)	561 (0.6%)	0.90 (0.68-1.20)
Never	7536 (81.9%)	75901 (82.4%)	(ref)
Fluvastatin			
1-308	5 (<0.1%)	63 (<0.1%)	0.76 (0.31-1.89)
309-938	8 (<0.1%)	90 (<0.1%)	0.86 (0.41-1.76)
939-2203	8 (<0.1%)	106 (<0.1%)	0.70 (0.34-1.45)
2203+	11 (<0.1%)	137 (<0.1%)	0.76 (0.41-1.41)
Never	7536 (81.9%)	75901 (82.4%)	(ref)
Rosuvastatin			
1-308	10 (<0.1%)	89 (<0.1%)	1.11 (0.58-2.14)
309-938	20 (0.1%)	191 (0.2%)	0.99 (0.63-1.58)
939-2203	37 (0.4%)	261 (0.3%)	1.35 (0.96-1.91)
2203+	43 (0.5%)	415 (0.5%)	0.98 (0.71-1.34)
Never	7536 (81.9%)	75901 (82.4%)	(ref)

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; ref = reference; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Supplemental Table 8. Association between cumulative statin DDDs and ovarian cancer risk by statin class

	Cases N=9207	Controls N=92070	OR* (95% CI)
Hydrophilic			
1-308	53 (0.6%)	484 (0.5%)	1.07 (0.80-1.42)
309-938	57 (0.6%)	595 (0.6%)	0.92 (0.70-1.20)
939-2203	74 (0.8%)	690 (0.7%)	1.02 (0.80-1.30)
2203+	91 (0.9%)	926 (1.0%)	0.93 (0.75-1.15)
Never	7536 (81.9%)	75901 (82.4%)	(ref)
Lipophilic			
1-308	397 (4.3%)	3807 (4.1%)	1.01 (0.91-1.13)
309-938	395 (4.3%)	3811 (4.1%)	1.00 (0.90-1.11)
939-2203	399 (4.3%)	3925 (4.3%)	0.98 (0.88-1.09)
2203+	401 (4.4%)	3899 (4.2%)	0.97 (0.87-1.08)
Never	7536 (81.9%)	75901 (82.4%)	(ref)

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; ref = reference; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Supplemental Table 9. Association between statin indications and ovarian cancer risk in CPRD

	Cases N=9,207	Controls N=92,070	OR* (95% CI)
Hypertension			
Ever	2834 (30.8%)	27837 (30.2%)	1.01 (0.96-1.05)
Never	6373 (69.2%)	64233 (69.8%)	(ref)
Heart Disease			
Ever	665 (7.2%)	6512 (7.1%)	0.99 (0.92-1.08)
Never	8542 (92.8%)	85558 (92.9%)	(ref)
Dyslipidemia			
Ever	1139 (12.4%)	11270 (12.2%)	1.01 (0.94-1.08)
Never	8068 (87.6%)	80800 (87.8%)	(ref)
Diabetes Type 2			
Ever	610 (6.6%)	5907 (6.4%)	0.98 (0.89-1.07)
Never	8597 (93.4%)	86163 (93.6%)	(ref)
Smoking			
Current	1610 (17.5%)	15887 (17.3%)	1.00 (0.93-1.07)
Former	4211 (45.7%)	41292 (44.8%)	0.99 (0.93-1.04)
Never	2169 (23.6%)	20996 (22.8%)	(ref)
Missing	1217 (13.2%)	13895 (15.1%)	0.82 (0.76-0.88)
BMI at least 5 years Prior to diagnosis (kg/m²)			
Low <18.5	255 (2.8%)	3073 (3.3%)	0.89 (0.78-1.02)
Normal 18.5-24.9	1966 (21.4%)	21100 (22.9%)	(ref)
Overweight 25-29.9	1954 (21.2%)	18928 (20.6%)	1.10 (1.03-1.18)
Obese 30+	1487 (16.2%)	13367 (14.5%)	1.19 (1.11-1.28)
Missing	3545 (38.5%)	35602 (38.7%)	1.05 (0.99-1.11)

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; ref = reference; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Supplemental Table 10. Associations between statins and ovarian cancer risk by indication group and statin class in CPRD

	Cases		Controls		OR* (95% CI)	p ^{int}
	Non-users n (%)	Statin users n (%)	Non-users n (%)	Statin users n (%)		
Hydrophilic						
Hypertension						
Ever	2642 (29.6%)	192 (69.8%)	26066 (29.2%)	1771 (65.7%)	0.88 (0.70-1.01)	0.21
Never	6290 (71.4%)	83 (30.2%)	63309 (70.8%)	924 (34.3%)	1.07 (0.91-1.24)	
Total	8932	275	89375	2695	-	
Heart Disease						
Ever	581 (6.5%)	84 (30.5%)	5695 (6.4%)	817 (30.3%)	1.01 (0.70-1.29)	0.97
Never	8351 (93.5%)	191 (69.5%)	83680 (93.6%)	1878 (69.7%)	0.99 (0.85-1.15)	
Total	8932	275	89375	2695	-	
Lipophilic						
Hypertension						
Ever	1781 (23.4%)	1053 (66.1%)	17704 (23.1%)	10133 (65.6%)	1.00 (0.91-1.10)	0.85
Never	5833 (76.6%)	540 (33.9%)	58915 (76.9%)	5318 (34.4%)	1.03 (0.65-1.12)	
Total	7614	1593	76619	15451	-	
Heart Disease						
Ever	296 (3.9%)	369 (23.2%)	2880 (3.8%)	3632 (23.5%)	0.98 (0.83-1.17)	0.63
Never	7318 (96.1%)	1224 (76.8%)	73739 (96.2%)	11819 (76.5%)	1.02 (0.95-1.09)	
Total	7614	1593	76619	15451	-	

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; ref = reference; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Supplemental Table 11. Associations between statin class and ovarian cancer risk by other factors in CPRD

	Cases		Controls		OR* (95% CI)	p ^{int}
	Non-users n (%)	Statin users n (%)	Non-users n (%)	Statin users n (%)		
Hydrophilic						
Type 2 diabetes						
Ever	528 (5.9%)	82 (29.8%)	5268 (5.9%)	639 (23.7%)	1.29 (1.01-1.66)	0.04
Never	8404 (94.1%)	193 (70.2%)	84107 (94.1%)	2056 (76.3%)	0.92 (0.79-1.07)	
Total	8932	275	89375	2695	-	
BMI kg/m ²						
< 18.5	248 (2.8%)	7 (2.5%)	2996 (3.4%)	77 (2.9%)	1.06 (0.49-2.34)	0.08
18.5-24.9	1919 (21.5%)	47 (17.1%)	20478 (22.9%)	622 (23.1%)	0.78 (0.58-1.05)	
25-29.9	1857 (20.8%)	97 (35.3%)	18025 (20.2%)	903 (33.5%)	1.01 (0.82-1.26)	
30+	1404 (15.7%)	83 (30.2%)	12612 (14.1%)	755 (28.0%)	0.97 (0.75-1.24)	
missing	3504 (39.2%)	41 (14.9%)	35264 (39.5%)	338 (12.5%)	1.21 (0.88-1.68)	
Total	8932	275	89375	2695	-	
Smoking						
Current	1555 (17.4%)	55 (20.0%)	15415 (17.2%)	472 (17.5%)	1.19 (0.89-1.59)	0.41
Former	4053 (45.4%)	158 (57.5%)	39706 (44.4%)	1586 (58.8%)	0.94 (0.80-1.12)	
Never	2110 (23.6%)	59 (21.5%)	20385 (22.8%)	611 (22.7%)	0.92 (0.70-1.21)	
Missing	1214 (13.6%)	3 (1.1%)	13869 (15.5%)	26 (1.0%)	-	
Total	8932	275	89375	2695	-	
Lipophilic						
Type 2 diabetes						
Ever	224 (2.9%)	386 (24.2%)	2162 (2.8%)	3745 (24.2%)	0.98 (0.82-1.18)	0.67
Never	7390 (97.1%)	1207 (75.8%)	74457 (97.2%)	11706 (75.8%)	1.02 (0.96-1.09)	
Total	7614	1593	76619	15451	-	
BMI kg/m ²						
< 18.5	209 (2.7%)	46 (2.9%)	2663 (3.5%)	410 (2.7%)	1.38 (0.98-1.94)	0.51

	1660	306	1768	3419	0.92	
18.5-24.9	(21.8%)	(19.2%)	(2.3%)	(22.1%)	(0.81-1.04)	
	1451	503	14063	4865	0.98	
25-29.9	(19.1%)	(31.6%)	(18.4%)	(31.5%)	(0.88-1.09)	
	974	513	8864	4503	1.03	
30+	(12.8%)	(32.2%)	(11.6%)	(29.1%)	(0.92-1.50)	
	3320	225	33348	2254	1.07	
missing	(43.6%)	(14.1%)	(43.5%)	(14.6%)	(0.98-1.17)	
Total	7614	1593	76619	15451	-	
Smoking						
	1329	281	13063	2824	1.02	
Current	(17.5%)	(17.6%)	(17.0%)	(18.3%)	(0.88-1.18)	
	3293	918	32533	8759	1.00	
Former	(43.2%)	(57.6%)	(42.5%)	(56.7%)	(0.92-1.09)	0.10
	1790	379	17275	3712	0.96	
Never	(23.5%)	(23.8%)	(22.5%)	(24.0%)	(0.85-1.09)	
	1202	15	13748	147	1.22	
Missing	(15.8%)	(1.0%)	(17.9%)	(1.0%)	(0.71-2.09)	
Total	7614	1593	76619	15451		

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; ref = reference; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Supplemental Table 12. Associations between individual hydrophilic statins and ovarian cancer risk in CPRD

Cases			Controls			p _{int}
Non-users n (%)	Statin users n (%)	Non-users n (%)	Statin users n (%)	OR* (95% CI)		
Pravastatin						
Type 2 diabetes						
Ever	567 (6.3%)	43 (28.5%)	5532 (39.4%)	375 (23.6%)	1.13 (0.82-1.58)	0.23
Never	8489 (93.7%)	108 (71.5%)	8495 (60.6%)	1211 (76.4%)	0.87 (0.71-1.06)	
Total	9056	151	14027	1586	-	
Rosuvastatin						
Type 2 diabetes						
Ever	576 (6.3%)	34 (30.9%)	5644 (6.2%)	263 (27.5%)	1.26 (0.87-1.82)	0.49
Never	8521 (93.7%)	76 (69.1%)	85470 (93.8%)	693 (72.5%)	1.07 (0.85-1.36)	
Total	9097	110	91114	956	-	

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; ref = reference; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Supplemental Table 13. Joint effects of hydrophilic statin use with concomitant medications in CPRD

	Cases N= 7811	Controls N= 78596	
			OR* (95% CI)
Beta-blocker			
--St / --B	6972 (89.3%)	70571 (89.8%)	Ref
+St / --B	240 (3.1%)	2354 (3.0%)	0.98 (0.86-1.12)
--St / +B	564 (7.2%)	5330 (6.8%)	1.10 (1.00-1.20)
+St / +B	35 (0.4%)	341 (0.4%)	0.99 (0.70-1.40)
Cardiac glycosides			
--St / --C	7395 (94.7%)	74673 (95.0%)	Ref
+St / --C	261 (3.3%)	2549 (3.2%)	0.98 (0.86-1.12)
--St / +C	141 (1.8%)	1228 (1.6%)	1.10 (0.92-1.31)
+St / +C	14 (0.2%)	146 (0.2%)	0.90 (0.52-1.56)
Metformin			
--St / --M	7454 (95.4%)	75127 (95.6%)	Ref
+St / --M	214 (0.3%)	2229 (2.8%)	0.92 (0.80-1.07)
--St / +M	82 (1.0%)	744 (0.9%)	1.00 (0.80-1.26)
+St / +M	61 (0.9%)	466 (0.6%)	1.21 (0.92-1.58)
Sulfonylureas			
--St / --Su	7445 (95.3%)	75007 (95.4%)	Ref
+St / --Su	234 (3.0%)	2375 (3.0%)	0.94 (0.82-1.08)
--St / +Su	91 (1.2%)	894 (1.1%)	0.96 (0.77-1.20)
+St / +Su	41 (0.5%)	320 (0.4%)	1.18 (0.86-1.64)
Insulin			
--St / --I	7489 (95.9%)	75540 (96.1%)	Ref
+St / --I	258 (3.3%)	2541 (3.2%)	0.97 (0.85-1.11)
--St / +I	47 (0.6%)	361 (0.5%)	1.29 (0.95-1.75)
+St / +I	17 (0.2%)	154 (0.2%)	1.02 (0.62-1.69)

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; ref = reference; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Supplemental Table 14. Joint effects of lipophilic statin use with other CVD medications in CPRD

	Cases N= 9129	Controls N= 91352	OR* (95% CI)
Beta-blocker			
--St / --B	6972 (76.3%)	70571 (77.3%)	Ref
+St / --B	1402 (15.4%)	13660 (15.0%)	0.99 (0.93-1.06)
--St / +B	564 (6.2%)	5330 (5.8%)	1.09 (1.00-1.20)
+St / +B	191 (2.1%)	1791 (2.0%)	1.04 (0.89-1.21)
Cardiac glycosides			
--St / --C	7395 (81.0%)	74673 (81.7%)	Ref
+St / --C	1532 (16.8%)	1804 (2.0%)	1.00 (0.94-1.06)
--St / +C	141 (1.5%)	1228 (1.3%)	1.06 (0.93-1.32)
+St / +C	61 (0.7%)	647 (0.7%)	0.89 (0.68-1.16)
Metformin			
--St / --M	7454 (81.7%)	75127 (82.2%)	Ref
+St / --M	1314 (14.4%)	12833 (14.0%)	0.99 (0.93-1.06)
--St / +M	82 (0.9%)	774 (0.8%)	1.00 (0.79-1.26)
+St / +M	279 (3.1%)	2618 (2.9%)	0.99 (0.87-1.12)
Sulfonylureas			
--St / --Su	7445 (81.6%)	75007 (82.1%)	Ref
+St / --Su	1415 (15.5%)	13751 (15.1%)	0.99 (0.93-1.05)
--St / +Su	91 (1.0%)	894 (1.0%)	0.96 (0.77-1.20)
+St / +Su	178 (1.9%)	1700 (1.9%)	0.97 (0.83-1.14)
Insulin			
--St / --I	7489 (82.0%)	75540 (82.7%)	Ref
+St / --I	1506 (16.5%)	14640 (16.0%)	0.99 (0.93-1.05)
--St / +I	47 (0.5%)	361 (0.4%)	1.29 (0.95-1.76)
+St / +I	87 (1.0%)	811 (1.0%)	1.01 (0.81-1.26)

Abbreviations: OR = odds ratio; CPRD = Clinical Practice Research Datalink; ref = reference; CI = confidence interval

* Adjusted for: OC use (1= ever, 0= never), BMI (1= <18.5, 2= 18.5-24.9, 3= 25-30, 4=30+, 8=missing), parity (1=1 birth, 2= 2 births, 3=3 births, 4=4 births, 5 = 4+ births, 8=missing/nulliparous)

Appendix G. Supplemental material for manuscript 3

Supplemental Table 1. Antibodies for S1P pathway markers

Receptor	Antibody	Antibody Manufacturer	Polyclonal or monoclonal	Species	Clone number	RNA or protein expression
S1PR1	MA5-32587	ThermoFisher	monoclonal	rabbit	JM10-66	Protein receptor
S1PR3	ab126622	abcam	monoclonal	rabbit	EPR454 1(2)	Protein receptor

Supplemental Table 2. Characteristics of ovarian cancer cases in POCS by TMA inclusion status

Characteristic	Cases not on TMA n=142	Cases included on TMA n=166	P*
Age (years)			
<50	39 (27.5%)	58 (34.9%)	0.47
50-54	26 (18.3%)	18 (10.8%)	
55-64	34 (23.9%)	43 (24.0%)	
65+	43 (30.3%)	47 (25.9%)	
Menopausal status			
Premenopausal	33 (23.2%)	50 (30.1%)	0.31
Postmenopausal	93 (65.5%)	94 (56.6%)	
Missing	16 (11.3%)	22 (13.3%)	
Menopausal hormone use			
No	106 (74.6%)	134 (76.7%)	0.83
Yes	36 (25.3%)	40 (22.9%)	
Missing	0 (0%)	1 (0.6%)	
OC Use			
No	128 (90.1%)	126 (75.9%)	0.91
Yes	12 (8.4%)	40 (24.1%)	
Missing	2 (1.4%)	0 (0.0%)	
Parity			
Nulliparous	30 (21.1%)	35 (20.0%)	0.74
Ever-pregnant	112 (72.5%)	145 (87.3%)	
Smoking			
Non-smoker	68 (47.9%)	66 (39.8%)	0.51
Former	27 (19.0%)	34 (20.5%)	
Current	47 (33.1%)	66 (39.8%)	
BMI kg/m ²			
<25	90 (63.4%)	96 (54.9%)	0.32
25-30	34 (23.9%)	50 (28.6%)	
>30	16 (11.3%)	28 (16.0%)	
Missing	2 (1.4%)	0 (0.0%)	
Grade			
I	9 (6.3%)	33 (19.8%)	<0.01
≥II	44 (31.0%)	119 (76.7%)	
Missing	89 (62.7%)	14 (8.4%)	
Stage			
I	23 (16.2%)	60 (36.1%)	<0.01
≥II	36 (25.3%)	85 (51.2%)	
Missing	83 (58.4%)	21 (12.7%)	
Histology			
Serous	60 (42.2%)	70 (42.2%)	<0.01
Endometrioid	29 (20.4%)	19 (11.4%)	
Mucinous	10 (7.0%)	7 (4.2%)	
Clear Cell	4 (2.8%)	7 (4.2%)	
Mixed epithelial	11 (7.7%)	40 (24.1%)	
Other	28 (19.7%)	25 (15.1%)	

*chi square tests were used for comparisons of categorical variables; p values < 0.05 were considered statistically significant

Supplemental Table 3. Overall distributions for S1PR1 and S1PR3 for all three IHC measures

	S1PR1	S1PR3
Area		
Median (IQR)	90 (80-100)	100 (100-100)
Range	0-100	25-100
Intensity		
Median	2 (1-3)	2 (1-3)
Range	0-3	1-3
Score		
Median (IQR)	180 (90-200)	200 (100-200)
Range	0-300	25-300

Abbreviations: IQR = interquartile range, S1PR1 = sphingosine-1-phosphate receptor 1; S1PR3 = sphingosine 1-phosphate receptor 3

Supplemental Table 4. Spearman correlation coefficients between S1PR1 and S1PR3 by four primary ovarian cancer histotypes

	Correlation Coefficient	<i>p</i> [*]
Ovarian Cancer, n=105[†]		
Percent	0.11	0.26
Intensity	0.37	<0.0001
Score	0.38	<0.0001
Serous histology, n=72		
Percent	-0.01	0.93
Intensity	0.24	0.04
Score	0.23	0.05
Endometrioid histology, n=19		
Percent	0.31	0.20
Intensity	0.19	0.44
Score	0.20	0.40
Mucinous histology, n= 7		
Percent	0.62	0.13
Intensity	0.81	0.03
Score	0.86	0.03
Clear cell histology, n=7		
Percent	0.63	0.13
Intensity	0.42	0.35
Score	0.57	0.18

* P <0.05 was considered statistically significant

[†] Includes only the four primary histotypes (serous, endometrioid, mucinous, and clear cell)

Supplemental Table 5. Associations between S1PR1 and ovarian cancer tumor characteristics

	Histology			Grade			Stage		
	Serous n=70	Non-serous n=35	OR* (95% CI)	well differentiated n=33	Poor or moderately differentiated n=119	OR* (95% CI)	I n=60	II-III n=85	OR* (95% CI)
S1P1									
Percent									
	16	12		15	36		26	20	
Q1 <25	(22.8%)	(34.3%)	1.00	(45.5%)	(30.3%)	1.00	(43.3%)	(23.5%)	1.00
	21	8	0.53	9	32	1.06	13	26	3.10
Q2 <90	(30.0%)	(22.8%)	(0.15-1.81)	(27.3%)	(26.9%)	(0.34-3.32)	(21.7%)	(30.6%)	(1.20-8.03)
	13	8	0.77	8	51	1.18		18	3.73
Q3 <99	(18.6%)	(22.8%)	(0.22-2.64)	(23.7%)	(42.0%)	(0.33-4.25)	7 (11.7%)	(21.2%)	(1.17-11.9)
	20	7	0.35	1	0	-	14	21	2.06
Q4 <100	(28.6%)	(20.0%)	(0.10-1.25)	(3.0%)	(0.0%)	-	(23.3%)	(24.7%)	(0.77-5.52)
Intensity									
	0	2		3	3				
0	(0.0%)	(5.7%)	-	(9.1%)	(2.5%)	-	3 (5.0%)	2 (2.4%)	-
	16	12		10	39		20	24	
1	(22.9%)	(34.3%)	1.00	(30.3%)	(32.8%)	1.00	(33.3%)	(28.2%)	1.00
	36	12	0.29	14	54	0.96	21	41	1.90
2	(51.4%)	(34.3%)	(0.10-0.94)	(42.4%)	(45.4%)	(0.33-2.78)	(35.0%)	(48.2%)	(0.79-4.56)
	18		0.62	6	23	1.13	16	18	0.93
3	(25.7%)	9 (25.7%)	(0.18-2.13)	(18.2%)	(19.3%)	(0.33-2.78)	(26.7%)	(21.2%)	(0.79-4.56)
Score									
	10	13		12	33		21	21	
Q1 <90	(14.3%)	(37.1%)	1.00	(36.4%)	(27.7%)	1.00	(35.0%)	(24.7%)	1.00
	26		0.05	9	39	1.69	15	25	1.54
Q2 <180	(37.1%)	4 (11.4%)	(0.01-0.28)	(27.3%)	(32.8%)	(0.53-5.37)	(25.0%)	(29.0%)	(0.59-4.07)
	16		0.25	6	25	1.43		22	3.02
Q3 <200	(22.9%)	9 (25.7%)	(0.06-1.04)	(18.3%)	(21.0%)	(0.39-5.26)	8 (13.3%)	(25.9%)	(1.01-9.05)
	18		0.29	6	22	1.53	16	17	0.92
Q4 <300	(25.7%)	9 (25.7%)	(0.08-1.14)	(18.2%)	(18.5%)	(0.43-5.48)	(26.7%)	(20.0%)	(0.34-2.47)

*Adjusted for age in 5-year categories, menopause status (pre-menopausal, post-menopausal, missing), use of menopausal hormone therapy (ever versus never), smoking history (current, former, never), parity (ever versus never), OC use (ever versus never) and BMI (<25, 25-29.9, 30+)

Supplemental Table 6. Associations between S1PR3 and ovarian cancer tumor characteristics

Histology			Grade			Stage		
Serous N=70	Non-serous N=35	OR* (95% CI)	Well-differentiated N=33	Poor or moderately differentiated N=119	OR* (95% CI)	I N=60	II-III N=85	OR* (95% CI)
S1P3								
Percent								
Q1 = 100	-	-	-	-	-	-	-	-
Q2 = 100	-	-	-	-	-	-	-	-
Q3 = 100	-	-	-	-	-	-	-	-
Q4 = 100	-	-	-	-	-	-	-	-
Intensity								
0	0	-	0	0	-	0	0	-
(0.0%)	(0.0%)		(0%)	(0%)		(0%)	(0%)	
15	14		15	42		25	26	
(21.4%)	(40.0%)	1.00	(45.5%)	(35.3%)	1.00	(41.7%)	(30.1%)	1.00
40	18	0.59	14	60	2.07	27	46	1.56
(57.1%)	(51.4%)	(0.21-1.60)	(42.4%)	(50.4%)	(0.81-5.28)	(45.0%)	(54.1%)	(0.72-3.33)
15	3	0.21	4	17	2.18	8	13	1.64
(21.4%)	(8.6%)	(0.04-0.99)	(12.1%)	(14.3%)	(0.56-8.50)	(13.3%)	(15.3%)	(0.55-4.92)
Score								
15	14		15	42		25	26	
(21.4%)	(40.0%)	1.00	(45.5%)	(35.3%)	1.00	(41.7%)	(30.1%)	1.00
Q2 < 200	-	-	-	-	-	-	-	-
40	18	0.59	14	60	2.07	27	46	1.56
(57.1%)	(51.4%)	(0.21-1.60)	(42.4%)	(50.4%)	(0.81-5.28)	(45.0%)	(54.1%)	(0.73-3.33)
15	3	0.21	4	17	2.18	8	13	1.64
(21.4%)	(8.6%)	(0.04-0.99)	(12.1%)	(14.3%)	(0.56-8.50)	(13.3%)	(15.3%)	(1.55-4.92)

Abbreviations: S1PR1 = Sphingosine phosphate receptor 1; S1PR3 = Sphingosine phosphate receptor 3; OR = odds ratio, CI = confidence interval, Q = quartile

*Adjusted for age in 5-year categories, menopause status (pre-menopausal, post-menopausal, missing), use of menopausal hormone therapy (ever versus never), smoking history (current, former, never), parity (ever versus never), OC use (ever versus never) and BMI (<25, 25-29.9, 30+)

Appendix H

IRB exempt status by the University of Maryland



1204 Marie Mount Hall
College Park, MD 20742-5125
TEL 301.405.4212
FAX 301.314.1475
irb@umd.edu
www.umresearch.umd.edu/IRB

DATE: August 24, 2020

TO: Sarah Irvin, MS
FROM: University of Maryland College Park (UMCP) IRB

PROJECT TITLE: [1615403-1] Lipid-modifying Medications and Ovarian Cancer: Implications for Risk and Potential Etiologic Pathways

REFERENCE #:

SUBMISSION TYPE: New Project

ACTION: DETERMINATION OF EXEMPT STATUS

DECISION DATE: August 24, 2020

REVIEW CATEGORY: Exemption category # 4

Thank you for your submission of New Project materials for this project. The University of Maryland College Park (UMCP) IRB has determined this project is **EXEMPT FROM IRB REVIEW** according to federal regulations.

We will retain a copy of this correspondence within our records.

If you have any questions, please contact the IRB Office at 301-405-4212 or irb@umd.edu. Please include your project title and reference number in all correspondence with this committee.

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within University of Maryland College Park (UMCP) IRB's records.

Bibliography

1. ACS, *Cancer Facts & Figures*. 2018, American Cancer Society.
2. Sung, H., et al., *Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries*. CA Cancer J Clin, 2021. **71**(3): p. 209-249.
3. Mallen, A.R., M.K. Townsend, and S.S. Tworoger, *Risk Factors for Ovarian Carcinoma*. Hematol Oncol Clin North Am, 2018. **32**(6): p. 891-902.
4. Bristow, R.E., I. Puri, and D.S. Chi, *Cytoreductive surgery for recurrent ovarian cancer: a meta-analysis*. Gynecol Oncol, 2009. **112**(1): p. 265-74.
5. Jemal, A., et al., *Cancer statistics, 2009*. CA Cancer J Clin, 2009. **59**(4): p. 225-49.
6. Buys, S.S., et al., *Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial*. JAMA, 2011. **305**(22): p. 2295-303.
7. Jacobs, I.J., et al., *Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial*. Lancet, 2016. **387**(10022): p. 945-956.
8. Kurman, R.J. and M. Shih Ie, *The Dualistic Model of Ovarian Carcinogenesis: Revisited, Revised, and Expanded*. Am J Pathol, 2016. **186**(4): p. 733-47.
9. Yang, H.P., et al., *Ovarian cancer risk factors by histologic subtypes in the NIH-AARP Diet and Health Study*. Int J Cancer, 2012. **131**(4): p. 938-48.
10. Wentzensen, N., et al., *Ovarian Cancer Risk Factors by Histologic Subtype: An Analysis From the Ovarian Cancer Cohort Consortium*. J Clin Oncol, 2016. **34**(24): p. 2888-98.
11. Trabert, B., et al., *Aspirin, nonaspirin nonsteroidal anti-inflammatory drug, and acetaminophen use and risk of invasive epithelial ovarian cancer: a pooled analysis in the Ovarian Cancer Association Consortium*. J Natl Cancer Inst, 2014. **106**(2): p. djt431.
12. Sieh, W., et al., *Tubal ligation and risk of ovarian cancer subtypes: a pooled analysis of case-control studies*. Int J Epidemiol, 2013. **42**(2): p. 579-89.
13. Pearce, C.L., et al., *Association between endometriosis and risk of histological subtypes of ovarian cancer: a pooled analysis of case-control studies*. Lancet Oncol, 2012. **13**(4): p. 385-94.
14. Rasmussen, C.B., et al., *Pelvic Inflammatory Disease and the Risk of Ovarian Cancer and Borderline Ovarian Tumors: A Pooled Analysis of 13 Case-Control Studies*. Am J Epidemiol, 2017. **185**(1): p. 8-20.
15. Olsen, C.M., et al., *Obesity and risk of ovarian cancer subtypes: evidence from the Ovarian Cancer Association Consortium*. Endocr Relat Cancer, 2013. **20**(2): p. 251-62.
16. Shah, M.M., et al., *Diabetes mellitus and ovarian cancer: more complex than just increasing risk*. Gynecol Oncol, 2014. **135**(2): p. 273-7.
17. Bjorge, T., et al., *Metabolic risk factors and ovarian cancer in the Metabolic Syndrome and Cancer project*. Int J Epidemiol, 2011. **40**(6): p. 1667-77.
18. Olsen, C.M., et al., *Obesity and the risk of epithelial ovarian cancer: a systematic review and meta-analysis*. Eur J Cancer, 2007. **43**(4): p. 690-709.
19. Lee, J.Y., et al., *Diabetes mellitus and ovarian cancer risk: a systematic review and meta-analysis of observational studies*. Int J Gynecol Cancer, 2013. **23**(3): p. 402-12.

20. Chen, H.F., et al., *A large scale population-based cohort study on the risk of ovarian neoplasm in patients with type 2 diabetes mellitus*. Gynecol Oncol, 2014. **134**(3): p. 576-80.
21. Esposito, K., et al., *Metabolic syndrome and risk of cancer: a systematic review and meta-analysis*. Diabetes Care, 2012. **35**(11): p. 2402-11.
22. CDC. *Adult Obesity Facts*. 2021 September 20, 2021]; Available from: <https://www.cdc.gov/obesity/data/adult.html>.
23. CDC. *Prevalence of Obesity and Severe Obesity Among Adults: United States, 2017–2018*. 2021; Available from: <https://www.cdc.gov/nchs/products/databriefs/db360.htm>.
24. CDC. *National Diabetes Statistics Report*. 2020; Available from: <https://www.cdc.gov/diabetes/pdfs/data/statistics/national-diabetes-statistics-report.pdf>.
25. Hirode, G. and R.J. Wong, *Trends in the Prevalence of Metabolic Syndrome in the United States, 2011-2016*. JAMA, 2020. **323**(24): p. 2526-2528.
26. Craig, E.R., et al., *Metabolic risk factors and mechanisms of disease in epithelial ovarian cancer: A review*. Gynecol Oncol, 2016. **143**(3): p. 674-683.
27. Doerstling, S.S., C.H. O'Flanagan, and S.D. Hursting, *Obesity and Cancer Metabolism: A Perspective on Interacting Tumor-Intrinsic and Extrinsic Factors*. Front Oncol, 2017. **7**: p. 216.
28. Hursting, S.D. and S.M. Dunlap, *Obesity, metabolic dysregulation, and cancer: a growing concern and an inflammatory (and microenvironmental) issue*. Ann N Y Acad Sci, 2012. **1271**: p. 82-7.
29. Neagu, M., et al., *Inflammation and Metabolism in Cancer Cell-Mitochondria Key Player*. Front Oncol, 2019. **9**: p. 348.
30. Wagh, A. and N.J. Stone, *Treatment of metabolic syndrome*. Expert Rev Cardiovasc Ther, 2004. **2**(2): p. 213-28.
31. Saisho, Y., *Metformin and Inflammation: Its Potential Beyond Glucose-lowering Effect*. Endocr Metab Immune Disord Drug Targets, 2015. **15**(3): p. 196-205.
32. Jain, M.K. and P.M. Ridker, *Anti-inflammatory effects of statins: clinical evidence and basic mechanisms*. Nat Rev Drug Discov, 2005. **4**(12): p. 977-87.
33. Febbraro, T., E. Lengyel, and I.L. Romero, *Old drug, new trick: repurposing metformin for gynecologic cancers?* Gynecol Oncol, 2014. **135**(3): p. 614-21.
34. Liu, Y., et al., *Effect of statin on risk of gynecologic cancers: a meta-analysis of observational studies and randomized controlled trials*. Gynecol Oncol, 2014. **133**(3): p. 647-55.
35. Home, P.D., et al., *Experience of malignancies with oral glucose-lowering drugs in the randomised controlled ADOPT (A Diabetes Outcome Progression Trial) and RECORD (Rosiglitazone Evaluated for Cardiovascular Outcomes and Regulation of Glycaemia in Diabetes) clinical trials*. Diabetologia, 2010. **53**(9): p. 1838-45.
36. Urpilainen, E., et al., *The role of metformin and statins in the incidence of epithelial ovarian cancer in type 2 diabetes: a cohort and nested case-control study*. Bjog, 2018. **125**(8): p. 1001-1008.
37. Gong, Z., et al., *Diabetes, metformin and incidence of and death from invasive cancer in postmenopausal women: Results from the women's health initiative*. Int J Cancer, 2016. **138**(8): p. 1915-27.

38. Baur, D.M., et al., *Type 2 diabetes mellitus and medications for type 2 diabetes mellitus are associated with risk for and mortality from cancer in a German primary care cohort*. *Metabolism*, 2011. **60**(10): p. 1363-71.
39. Tseng, C.H., *Metformin reduces ovarian cancer risk in Taiwanese women with type 2 diabetes mellitus*. *Diabetes Metab Res Rev*, 2015. **31**(6): p. 619-26.
40. Bodmer, M., et al., *Use of metformin and the risk of ovarian cancer: a case-control analysis*. *Gynecol Oncol*, 2011. **123**(2): p. 200-4.
41. Clearfield, M., et al., *Air Force/Texas Coronary Atherosclerosis Prevention Study (AFCAPS/TexCAPS): efficacy and tolerability of long-term treatment with lovastatin in women*. *J Womens Health Gend Based Med*, 2001. **10**(10): p. 971-81.
42. Desai, P., et al., *An analysis of the association between statin use and risk of endometrial and ovarian cancers in the Women's Health Initiative*. *Gynecol Oncol*, 2018. **148**(3): p. 540-546.
43. Yu, O., et al., *Statin use and female reproductive organ cancer risk in a large population-based setting*. *Cancer Causes Control*, 2009. **20**(5): p. 609-16.
44. Friedman, G.D., et al., *Screening statins for possible carcinogenic risk: up to 9 years of follow-up of 361,859 recipients*. *Pharmacoepidemiol Drug Saf*, 2008. **17**(1): p. 27-36.
45. Baandrup, L., et al., *Statin use and risk for ovarian cancer: a Danish nationwide case-control study*. *Br J Cancer*, 2015. **112**(1): p. 157-61.
46. Lavie, O., et al., *The effect of statins on risk and survival of gynecological malignancies*. *Gynecol Oncol*, 2013. **130**(3): p. 615-9.
47. Akinwunmi, B., et al., *Statin therapy and association with ovarian cancer risk in the New England Case Control (NEC) study*. *Int J Cancer*, 2019. **144**(5): p. 991-1000.
48. Kaye, J.A. and H. Jick, *Statin use and cancer risk in the General Practice Research Database*. *Br J Cancer*, 2004. **90**(3): p. 635-7.
49. Irvin, S., et al., *Systematic review and meta-analysis of studies assessing the relationship between statin use and risk of ovarian cancer*. *Cancer Causes Control*, 2020. **31**(10): p. 869-879.
50. Herrett, E., et al., *Data Resource Profile: Clinical Practice Research Datalink (CPRD)*. *Int J Epidemiol*, 2015. **44**(3): p. 827-36.
51. Wentzensen, N., et al., *Discovery and validation of methylation markers for endometrial cancer*. *Int J Cancer*, 2014. **135**(8): p. 1860-8.
52. Felix, A.S., et al., *Cell-cycle protein expression in a population-based study of ovarian and endometrial cancers*. *Front Oncol*, 2015. **5**: p. 25.
53. Cuvillier, O., et al., *Suppression of ceramide-mediated programmed cell death by sphingosine-1-phosphate*. *Nature*, 1996. **381**(6585): p. 800-3.
54. Shida, D., et al., *Targeting SphK1 as a new strategy against cancer*. *Curr Drug Targets*, 2008. **9**(8): p. 662-73.
55. Ogretmen, B., *Sphingolipid metabolism in cancer signalling and therapy*. *Nat Rev Cancer*, 2018. **18**(1): p. 33-50.
56. Hart, P.C., et al., *SPHK1 Is a Novel Target of Metformin in Ovarian Cancer*. *Mol Cancer Res*, 2019.
57. NCI. *Cancer Staging*. 2019; Available from: <https://www.cancer.gov/about-cancer/diagnosis-staging/staging>.

58. Prat, J., *Ovarian, fallopian tube and peritoneal cancer staging: Rationale and explanation of new FIGO staging 2013*. Best Pract Res Clin Obstet Gynaecol, 2015. **29**(6): p. 858-69.
59. Matz, M., et al., *Worldwide comparison of ovarian cancer survival: Histological group and stage at diagnosis (CONCORD-2)*. Gynecol Oncol, 2017. **144**(2): p. 396-404.
60. Jarboe, E.A., et al., *Tubal and ovarian pathways to pelvic epithelial cancer: a pathological perspective*. Histopathology, 2008. **53**(2): p. 127-38.
61. Labidi-Galy, S.I., et al., *High grade serous ovarian carcinomas originate in the fallopian tube*. Nat Commun, 2017. **8**(1): p. 1093.
62. Olivier, M., M. Hollstein, and P. Hainaut, *TP53 mutations in human cancers: origins, consequences, and clinical use*. Cold Spring Harb Perspect Biol, 2010. **2**(1): p. a001008.
63. Reid, B.M., J.B. Permuth, and T.A. Sellers, *Epidemiology of ovarian cancer: a review*. Cancer Biol Med, 2017. **14**(1): p. 9-32.
64. Coburn, S.B., et al., *International patterns and trends in ovarian cancer incidence, overall and by histologic subtype*. Int J Cancer, 2017. **140**(11): p. 2451-2460.
65. Torre, L.A., et al., *Ovarian cancer statistics, 2018*. CA Cancer J Clin, 2018. **68**(4): p. 284-296.
66. Nicholson, B.D., et al., *Weight loss as a predictor of cancer in primary care: a systematic review and meta-analysis*. Br J Gen Pract, 2018. **68**(670): p. e311-e322.
67. Doubeni, C.A., A.R. Doubeni, and A.E. Myers, *Diagnosis and Management of Ovarian Cancer*. Am Fam Physician, 2016. **93**(11): p. 937-44.
68. Dixon-Suen, S.C., et al., *Adult height is associated with increased risk of ovarian cancer: a Mendelian randomisation study*. Br J Cancer, 2018. **118**(8): p. 1123-1129.
69. Brinton, L.A., et al., *Relationships of uterine and ovarian tumors to pre-existing chronic conditions*. Gynecol Oncol, 2007. **107**(3): p. 487-94.
70. American Diabetes, A., *Diagnosis and classification of diabetes mellitus*. Diabetes Care, 2010. **33 Suppl 1**: p. S62-9.
71. Zhang, D., et al., *Diabetes mellitus and risk of ovarian cancer. A systematic review and meta-analysis of 15 cohort studies*. Diabetes Res Clin Pract, 2017. **130**: p. 43-52.
72. Gapstur, S.M., et al., *Type II diabetes mellitus and the incidence of epithelial ovarian cancer in the cancer prevention study-II nutrition cohort*. Cancer Epidemiol Biomarkers Prev, 2012. **21**(11): p. 2000-5.
73. Huang, P.L., *A comprehensive definition for metabolic syndrome*. Dis Model Mech, 2009. **2**(5-6): p. 231-7.
74. Rezaianzadeh, A., S.M. Namayandeh, and S.M. Sadr, *National Cholesterol Education Program Adult Treatment Panel III Versus International Diabetic Federation Definition of Metabolic Syndrome, Which One is Associated with Diabetes Mellitus and Coronary Artery Disease?* Int J Prev Med, 2012. **3**(8): p. 552-8.
75. Ibrahim, M.A. and I. Jialal, *Hypercholesterolemia*, in *StatPearls*. 2019: Treasure Island (FL).
76. Melvin, J.C., et al., *Lipid profiles and risk of breast and ovarian cancer in the Swedish AMORIS study*. Cancer Epidemiol Biomarkers Prev, 2012. **21**(8): p. 1381-4.
77. Hu, J., et al., *Dietary cholesterol intake and cancer*. Ann Oncol, 2012. **23**(2): p. 491-500.
78. Kabat, G.C., et al., *Serum lipids and risk of obesity-related cancers in postmenopausal women*. Cancer Causes Control, 2018. **29**(1): p. 13-24.

79. Inoue, M., et al., *Diabetes mellitus and the risk of cancer: results from a large-scale population-based cohort study in Japan*. Arch Intern Med, 2006. **166**(17): p. 1871-7.
80. Chodick, G., et al., *Diabetes and risk of incident cancer: a large population-based cohort study in Israel*. Cancer Causes Control, 2010. **21**(6): p. 879-87.
81. Harding, J.L., et al., *Cancer risk among people with type 1 and type 2 diabetes: disentangling true associations, detection bias, and reverse causation*. Diabetes Care, 2015. **38**(2): p. 264-70.
82. Faber, M.T., et al., *Cigarette smoking and risk of ovarian cancer: a pooled analysis of 21 case-control studies*. Cancer Causes Control, 2013. **24**(5): p. 989-1004.
83. Harris, H.R., et al., *Polycystic Ovary Syndrome, Oligomenorrhea, and Risk of Ovarian Cancer Histotypes: Evidence from the Ovarian Cancer Association Consortium*. Cancer Epidemiol Biomarkers Prev, 2018. **27**(2): p. 174-182.
84. Khan, M., et al., *Site-specific cancer risk due to diabetes mellitus history: evidence from the Japan Collaborative Cohort (JACC) Study*. Asian Pac J Cancer Prev, 2006. **7**(2): p. 253-9.
85. Rapp, K., et al., *Fasting blood glucose and cancer risk in a cohort of more than 140,000 adults in Austria*. Diabetologia, 2006. **49**(5): p. 945-52.
86. Stattin, P., et al., *Prospective study of hyperglycemia and cancer risk*. Diabetes Care, 2007. **30**(3): p. 561-7.
87. Johnson, J.A., et al., *Time-varying incidence of cancer after the onset of type 2 diabetes: evidence of potential detection bias*. Diabetologia, 2011. **54**(9): p. 2263-71.
88. Lambe, M., et al., *Impaired glucose metabolism and diabetes and the risk of breast, endometrial, and ovarian cancer*. Cancer Causes Control, 2011. **22**(8): p. 1163-71.
89. Wotton, C.J., D.G. Yeates, and M.J. Goldacre, *Cancer in patients admitted to hospital with diabetes mellitus aged 30 years and over: record linkage studies*. Diabetologia, 2011. **54**(3): p. 527-34.
90. Lo, S.F., et al., *Modest increase in risk of specific types of cancer types in type 2 diabetes mellitus patients*. Int J Cancer, 2013. **132**(1): p. 182-8.
91. Attner, B., et al., *Cancer among patients with diabetes, obesity and abnormal blood lipids: a population-based register study in Sweden*. Cancer Causes Control, 2012. **23**(5): p. 769-77.
92. O'Mara, B.A., T. Byers, and E. Schoenfeld, *Diabetes mellitus and cancer risk: a multisite case-control study*. J Chronic Dis, 1985. **38**(5): p. 435-41.
93. Adler, A.I., et al., *Is diabetes mellitus a risk factor for ovarian cancer? A case-control study in Utah and Washington (United States)*. Cancer Causes Control, 1996. **7**(4): p. 475-8.
94. Parazzini, F., et al., *Ovarian cancer risk and history of selected medical conditions linked with female hormones*. Eur J Cancer, 1997. **33**(10): p. 1634-7.
95. Kuriki, K., K. Hirose, and K. Tajima, *Diabetes and cancer risk for all and specific sites among Japanese men and women*. Eur J Cancer Prev, 2007. **16**(1): p. 83-9.
96. Reis, N. and N. Kizilkaya Beij, *Risk factors for ovarian cancer: results from a hospital-based case-control study*. Turkiye Klin J Medical Science, 2010. **30**: p. 79-87.
97. Bosetti, C., et al., *Diabetes mellitus and cancer risk in a network of case-control studies*. Nutr Cancer, 2012. **64**(5): p. 643-51.
98. Hemminki, K., et al., *Risk of cancer following hospitalization for type 2 diabetes*. Oncologist, 2010. **15**(6): p. 548-55.

99. Weiderpass, E., et al., *Diabetes mellitus and ovarian cancer (Sweden)*. Cancer Causes Control, 2002. **13**(8): p. 759-64.
100. Hsu, P.C., et al., *A Population-Based Cohort Study of All-Cause and Site-Specific Cancer Incidence Among Patients With Type 1 Diabetes Mellitus in Taiwan*. J Epidemiol, 2015. **25**(9): p. 567-73.
101. Shu, X., et al., *Cancer risk among patients hospitalized for Type 1 diabetes mellitus: a population-based cohort study in Sweden*. Diabet Med, 2010. **27**(7): p. 791-7.
102. Ness, R.B. and C. Cottreau, *Possible role of ovarian epithelial inflammation in ovarian cancer*. J Natl Cancer Inst, 1999. **91**(17): p. 1459-67.
103. Choi, J.H., et al., *Gonadotropins and ovarian cancer*. Endocr Rev, 2007. **28**(4): p. 440-61.
104. Cramer, D.W. and W.R. Welch, *Determinants of ovarian cancer risk. II. Inferences regarding pathogenesis*. J Natl Cancer Inst, 1983. **71**(4): p. 717-21.
105. Risch, H.A., *Hormonal etiology of epithelial ovarian cancer, with a hypothesis concerning the role of androgens and progesterone*. J Natl Cancer Inst, 1998. **90**(23): p. 1774-86.
106. Diergaarde, B. and M.L. Kurta, *Use of fertility drugs and risk of ovarian cancer*. Curr Opin Obstet Gynecol, 2014. **26**(3): p. 125-9.
107. Gronwald, J., et al., *Treatment of infertility does not increase the risk of ovarian cancer among women with a BRCA1 or BRCA2 mutation*. Fertil Steril, 2016. **105**(3): p. 781-785.
108. Kurta, M.L., et al., *Use of fertility drugs and risk of ovarian cancer: results from a U.S.-based case-control study*. Cancer Epidemiol Biomarkers Prev, 2012. **21**(8): p. 1282-92.
109. Risch, H.A., et al., *Events of reproductive life and the incidence of epithelial ovarian cancer*. Am J Epidemiol, 1983. **117**(2): p. 128-39.
110. Whittemore, A.S., R. Harris, and J. Itnyre, *Characteristics relating to ovarian cancer risk: collaborative analysis of 12 US case-control studies. IV. The pathogenesis of epithelial ovarian cancer. Collaborative Ovarian Cancer Group*. Am J Epidemiol, 1992. **136**(10): p. 1212-20.
111. Pyragius, C.E., et al., *Aberrant lipid metabolism: an emerging diagnostic and therapeutic target in ovarian cancer*. Int J Mol Sci, 2013. **14**(4): p. 7742-56.
112. Niemi, R.J., et al., *Ovarian tumours of different histologic type and clinical stage induce similar changes in lipid metabolism*. Br J Cancer, 2018. **119**(7): p. 847-854.
113. Long, J., et al., *Lipid metabolism and carcinogenesis, cancer development*. Am J Cancer Res, 2018. **8**(5): p. 778-791.
114. Munir, R., et al., *Lipid metabolism in cancer cells under metabolic stress*. Br J Cancer, 2019. **120**(12): p. 1090-1098.
115. Braicu, E.I., et al., *High-grade ovarian serous carcinoma patients exhibit profound alterations in lipid metabolism*. Oncotarget, 2017. **8**(61): p. 102912-102922.
116. Cheng, C., et al., *Lipid metabolism reprogramming and its potential targets in cancer*. Cancer Commun (Lond), 2018. **38**(1): p. 27.
117. O'Sullivan, J., et al., *Obesity and gastrointestinal cancer: the interrelationship of adipose and tumour microenvironments*. Nat Rev Gastroenterol Hepatol, 2018. **15**(11): p. 699-714.
118. O'Neill, L.A., R.J. Kishton, and J. Rathmell, *A guide to immunometabolism for immunologists*. Nat Rev Immunol, 2016. **16**(9): p. 553-65.

119. Graham, G.G., et al., *Clinical pharmacokinetics of metformin*. Clin Pharmacokinet, 2011. **50**(2): p. 81-98.
120. Pernicova, I. and M. Korbonits, *Metformin--mode of action and clinical implications for diabetes and cancer*. Nat Rev Endocrinol, 2014. **10**(3): p. 143-56.
121. Bailey, C.J., *Metformin: historical overview*. Diabetologia, 2017. **60**(9): p. 1566-1576.
122. Marshall, S.M., *60 years of metformin use: a glance at the past and a look to the future*. Diabetologia, 2017. **60**(9): p. 1561-1565.
123. Sharma, M., I. Nazareth, and I. Petersen, *Trends in incidence, prevalence and prescribing in type 2 diabetes mellitus between 2000 and 2013 in primary care: a retrospective cohort study*. BMJ Open, 2016. **6**(1): p. e010210.
124. Chow, C.K., et al., *Availability and affordability of essential medicines for diabetes across high-income, middle-income, and low-income countries: a prospective epidemiological study*. Lancet Diabetes Endocrinol, 2018. **6**(10): p. 798-808.
125. WHO. *Essential Medicines*. 2019; Available from: https://www.who.int/topics/essential_medicines/en/.
126. Johnson, N.P., *Metformin use in women with polycystic ovary syndrome*. Ann Transl Med, 2014. **2**(6): p. 56.
127. Baptiste, C.G., et al., *Insulin and hyperandrogenism in women with polycystic ovary syndrome*. J Steroid Biochem Mol Biol, 2010. **122**(1-3): p. 42-52.
128. Heard, M.J., et al., *Pregnancies following use of metformin for ovulation induction in patients with polycystic ovary syndrome*. Fertil Steril, 2002. **77**(4): p. 669-73.
129. FDA. *Glucophage, Glucophage XR*. 2019; Available from: https://www.accessdata.fda.gov/drugsatfda_docs/label/2017/020357s037s039,021202s021s023lbl.pdf.
130. Scarpello, J.H. and H.C. Howlett, *Metformin therapy and clinical uses*. Diab Vasc Dis Res, 2008. **5**(3): p. 157-67.
131. Demierre, M.F., et al., *Statins and cancer prevention*. Nat Rev Cancer, 2005. **5**(12): p. 930-42.
132. Gazerro, P., et al., *Pharmacological actions of statins: a critical appraisal in the management of cancer*. Pharmacol Rev, 2012. **64**(1): p. 102-46.
133. Gu Q, P.-R.R., Burt VL, Kit BK. , *Prescription cholesterol-lowering medication use in adults aged 40 and over: United States, 2003–2012*. NCHS data brief, no 177. . 2014, National Center for Health Statistics: Hyattsville, MD.
134. O'Keeffe, A.G., I. Nazareth, and I. Petersen, *Time trends in the prescription of statins for the primary prevention of cardiovascular disease in the United Kingdom: a cohort study using The Health Improvement Network primary care data*. Clin Epidemiol, 2016. **8**: p. 123-32.
135. Liu, Z., et al., *Statin use and reduced risk of biliary tract cancers in the UK Clinical Practice Research Datalink*. Gut, 2019. **68**(8): p. 1458-1464.
136. Schachter, M., *Chemical, pharmacokinetic and pharmacodynamic properties of statins: an update*. Fundam Clin Pharmacol, 2005. **19**(1): p. 117-25.
137. Centre", N.C.G., *Lipid Modification: Cardiovascular Risk Assessment and the Modification of Blood Lipids for the Primary and Secondary Prevention of Cardiovascular Disease*. 2014, London: National Institute for Health and Care Excellence (UK).

138. (NICE)", N.I.f.H.C.E. *Lipid-Modifying Drugs*. 2017; Available from: <https://www.nice.org.uk/advice/ktt3/chapter/evidence-context>.
139. USPSTF. *Statin Use for the Primary Prevention of Cardiovascular Disease in Adults: Preventive Medication*. 2016; Available from: Statin Use for the Primary Prevention of Cardiovascular Disease in Adults: Preventive Medication.
140. Permanente", K. *Atherosclerotic Cardiovascular Disease (ASCVD) Secondary Prevention Guideline*. 2018.
141. Duncan, R.E., A. El-Sohemy, and M.C. Archer, *Statins and cancer development*. *Cancer Epidemiol Biomarkers Prev*, 2005. **14**(8): p. 1897-8.
142. Bathaie, S.Z., et al., *Mevalonate Pathway and Human Cancers*. *Curr Mol Pharmacol*, 2017. **10**(2): p. 77-85.
143. Greenaway, J.B., et al., *Ovarian tumour growth is characterized by mevalonate pathway gene signature in an orthotopic, syngeneic model of epithelial ovarian cancer*. *Oncotarget*, 2016. **7**(30): p. 47343-47365.
144. Gong, L., et al., *The mevalonate coordinates energy input and cell proliferation*. *Cell Death & Disease*, 2019. **10**(4): p. 327.
145. Hamelin, B.A. and J. Turgeon, *Hydrophilicity/lipophilicity: relevance for the pharmacology and clinical effects of HMG-CoA reductase inhibitors*. *Trends Pharmacol Sci*, 1998. **19**(1): p. 26-37.
146. Kobayashi, Y., et al., *Drug repositioning of mevalonate pathway inhibitors as antitumor agents for ovarian cancer*. *Oncotarget*, 2017. **8**(42): p. 72147-72156.
147. Kato, S., et al., *Lipophilic but not hydrophilic statins selectively induce cell death in gynaecological cancers expressing high levels of HMGCoA reductase*. *J Cell Mol Med*, 2010. **14**(5): p. 1180-93.
148. Boudreau, D.M., O. Yu, and J. Johnson, *Statin use and cancer risk: a comprehensive review*. *Expert Opin Drug Saf*, 2010. **9**(4): p. 603-21.
149. Bao, Y., et al., *Sphingosine Kinase 1 and Sphingosine-1-Phosphate Signaling in Colorectal Cancer*. *Int J Mol Sci*, 2017. **18**(10).
150. Wang, P., et al., *Roles of sphingosine-1-phosphate signaling in cancer*. *Cancer Cell Int*, 2019. **19**: p. 295.
151. Dai, L., et al., *Sphingosine kinase 1/sphingosine-1-phosphate (S1P)/S1P receptor axis is involved in ovarian cancer angiogenesis*. *Oncotarget*, 2017. **8**(43): p. 74947-74961.
152. Zheng, X., et al., *The sphingosine kinase-1/sphingosine-1-phosphate axis in cancer: Potential target for anticancer therapy*. *Pharmacol Ther*, 2019. **195**: p. 85-99.
153. Knapp, P., et al., *Plasma and ovarian tissue sphingolipids profiling in patients with advanced ovarian cancer*. *Gynecol Oncol*, 2017. **147**(1): p. 139-144.
154. Zeleznik, O.A., et al., *Circulating Lysophosphatidylcholines, Phosphatidylcholines, Ceramides, and Sphingomyelins and Ovarian Cancer Risk: a 23-year Prospective Study*. *J Natl Cancer Inst*, 2019.
155. Pyne, N.J. and S. Pyne, *Sphingosine 1-phosphate and cancer*. *Nat Rev Cancer*, 2010. **10**(7): p. 489-503.
156. Spiegel, S. and S. Milstien, *Sphingosine-1-phosphate: an enigmatic signalling lipid*. *Nat Rev Mol Cell Biol*, 2003. **4**(5): p. 397-407.
157. Kousoulis, A.A., I. Rafi, and S. de Lusignan, *The CPRD and the RCGP: building on research success by enhancing benefits for patients and practices*. *Br J Gen Pract*, 2015. **65**(631): p. 54-5.

158. Williams, T., et al., *Recent advances in the utility and use of the General Practice Research Database as an example of a UK Primary Care Data resource*. Ther Adv Drug Saf, 2012. **3**(2): p. 89-99.
159. Roland, M., *Linking physicians' pay to the quality of care--a major experiment in the United kingdom*. N Engl J Med, 2004. **351**(14): p. 1448-54.
160. Ury, H.K., *Efficiency of case-control studies with multiple controls per case: continuous or dichotomous data*. Biometrics, 1975. **31**(3): p. 643-9.
161. Berek, J. and R. Bast, *Nonepithelial ovarian cancer*. Holland-Frei Cancer Medicine. 6th edition., ed. P.R. In: Kufe DW, Weichselbaum RR, et al., editors. 2003, Hamilton (ON).
162. Lewis, J.D., et al., *The relationship between time since registration and measured incidence rates in the General Practice Research Database*. Pharmacoepidemiol Drug Saf, 2005. **14**(7): p. 443-51.
163. Minassian, C., et al., *Methods to generate and validate a Pregnancy Register in the UK Clinical Practice Research Datalink primary care database*. Pharmacoepidemiol Drug Saf, 2019. **28**(7): p. 923-933.
164. Digital, N. *National Data Opt Out, September 2021*. 2022; Available from: <https://digital.nhs.uk/data-and-information/publications/statistical/national-data-opt-out/september-2021>.
165. Rundle, A.G., P. Vineis, and H. Ahsan, *Design options for molecular epidemiology research within cohort studies*. Cancer Epidemiol Biomarkers Prev, 2005. **14**(8): p. 1899-907.
166. Macedo, A.F., et al., *Statins and the risk of type 2 diabetes mellitus: cohort study using the UK clinical practice research datalink*. BMC Cardiovasc Disord, 2014. **14**: p. 85.
167. Brinton, L.A., et al., *Reproductive risk factors for endometrial cancer among Polish women*. Br J Cancer, 2007. **96**(9): p. 1450-6.
168. Garcia-Closas, M., et al., *Ovarian cancer risk and common variation in the sex hormone-binding globulin gene: a population-based case-control study*. BMC Cancer, 2007. **7**: p. 60.
169. Snaebjornsson, M.T., S. Janaki-Raman, and A. Schulze, *Greasing the Wheels of the Cancer Machine: The Role of Lipid Metabolism in Cancer*. Cell Metab, 2020. **31**(1): p. 62-76.
170. Viollet, B., et al., *Cellular and molecular mechanisms of metformin: an overview*. Clin Sci (Lond), 2012. **122**(6): p. 253-70.
171. Hay, N., *Reprogramming glucose metabolism in cancer: can it be exploited for cancer therapy?* Nat Rev Cancer, 2016. **16**(10): p. 635-49.
172. Higurashi, T., et al., *Metformin for chemoprevention of metachronous colorectal adenoma or polyps in post-polypectomy patients without diabetes: a multicentre double-blind, placebo-controlled, randomised phase 3 trial*. Lancet Oncol, 2016. **17**(4): p. 475-483.
173. Zi, F., et al., *Metformin and cancer: An existing drug for cancer prevention and therapy*. Oncol Lett, 2018. **15**(1): p. 683-690.
174. Tang, G.H., et al., *Association of Metformin with Breast Cancer Incidence and Mortality in Patients with Type II Diabetes: A GRADE-Assessed Systematic Review and Meta-analysis*. Cancer Epidemiol Biomarkers Prev, 2018. **27**(6): p. 627-635.
175. Evans, J.M., et al., *Metformin and reduced risk of cancer in diabetic patients*. BMJ, 2005. **330**(7503): p. 1304-5.

176. Libby, G., et al., *New users of metformin are at low risk of incident cancer: a cohort study among people with type 2 diabetes*. Diabetes Care, 2009. **32**(9): p. 1620-5.
177. Landman, G.W., et al., *Metformin associated with lower cancer mortality in type 2 diabetes: ZODIAC-16*. Diabetes Care, 2010. **33**(2): p. 322-6.
178. SEER, *urveillance, Epidemiology, and End Results (SEER) Program Populations (1969-2017)*, D. National Cancer Institute, Surveillance Research Program, Editor. 2016.
179. WHO. *Defined Daily Dose (DDD)*. 2019; Available from: https://www.who.int/medicines/regulation/medicines-safety/toolkit_ddd/en/.
180. Pottegard, A., et al., *Considerations for Pharmacoepidemiological Studies of Drug-Cancer Associations*. Basic Clin Pharmacol Toxicol, 2018. **122**(5): p. 451-459.
181. Sirmans, S.M. and K.A. Pate, *Epidemiology, diagnosis, and management of polycystic ovary syndrome*. Clin Epidemiol, 2013. **6**: p. 1-13.
182. Pollak, M., *Insulin and insulin-like growth factor signalling in neoplasia*. Nat Rev Cancer, 2008. **8**(12): p. 915-28.
183. Forbes, L.J., et al., *The role of the Quality and Outcomes Framework in the care of long-term conditions: a systematic review*. Br J Gen Pract, 2017. **67**(664): p. e775-e784.
184. Peres, L.C., et al., *Invasive Epithelial Ovarian Cancer Survival by Histotype and Disease Stage*. J Natl Cancer Inst, 2019. **111**(1): p. 60-68.
185. Adhyaru, B.B. and T.A. Jacobson, *Safety and efficacy of statin therapy*. Nat Rev Cardiol, 2018. **15**(12): p. 757-769.
186. Virani, S.S., et al., *Secondary Prevention for Atherosclerotic Cardiovascular Disease: Comparing Recent US and European Guidelines on Dyslipidemia*. Circulation, 2020. **141**(14): p. 1121-1123.
187. Shi, J., et al., *Association of metformin use with ovarian cancer incidence and prognosis: a systematic review and meta-analysis*. Int J Gynecol Cancer, 2019. **29**(1): p. 140-146.
188. Hindler, K., et al., *The role of statins in cancer therapy*. Oncologist, 2006. **11**(3): p. 306-15.
189. Mullen, P.J., et al., *The interplay between cell signalling and the mevalonate pathway in cancer*. Nat Rev Cancer, 2016. **16**(11): p. 718-731.
190. Seidman, J.D., et al., *Distribution and case-fatality ratios by cell-type for ovarian carcinomas: a 22-year series of 562 patients with uniform current histological classification*. Gynecol Oncol, 2015. **136**(2): p. 336-40.
191. Bodelon, C., et al., *Molecular Classification of Epithelial Ovarian Cancer Based on Methylation Profiling: Evidence for Survival Heterogeneity*. Clin Cancer Res, 2019. **25**(19): p. 5937-5946.
192. Sutphen, R., et al., *Lysophospholipids are potential biomarkers of ovarian cancer*. Cancer Epidemiol Biomarkers Prev, 2004. **13**(7): p. 1185-91.
193. Rodriguez, Y.I., et al., *Sphingosine-1 Phosphate: A New Modulator of Immune Plasticity in the Tumor Microenvironment*. Front Oncol, 2016. **6**: p. 218.
194. Kowalski, G.M., et al., *Plasma sphingosine-1-phosphate is elevated in obesity*. PLoS One, 2013. **8**(9): p. e72449.
195. Chaudhary, B. and E. Elkord, *Regulatory T Cells in the Tumor Microenvironment and Cancer Progression: Role and Therapeutic Targeting*. Vaccines (Basel), 2016. **4**(3).
196. Park, K.S., et al., *S1P stimulates chemotactic migration and invasion in OVCAR3 ovarian cancer cells*. Biochem Biophys Res Commun, 2007. **356**(1): p. 239-44.

197. Ou, Y., M.J. Cannon, and M. Nakagawa, *Regulatory T Cells in Gynecologic Cancer*. MOJ Immunol, 2018. **6**(2): p. 34-42.
198. Baudhuin, L.M., et al., *S1P3-mediated Akt activation and cross-talk with platelet-derived growth factor receptor (PDGFR)*. FASEB J, 2004. **18**(2): p. 341-3.
199. Fan, Q., et al., *Sphingosine-1-phosphate promotes ovarian cancer cell proliferation by disrupting Hippo signaling*. Oncotarget, 2017. **8**(16): p. 27166-27176.
200. Oskouian, B. and J. Saba, *Sphingosine-1-phosphate metabolism and intestinal tumorigenesis: lipid signaling strikes again*. Cell Cycle, 2007. **6**(5): p. 522-7.
201. Kunkel, G.T., et al., *Targeting the sphingosine-1-phosphate axis in cancer, inflammation and beyond*. Nat Rev Drug Discov, 2013. **12**(9): p. 688-702.
202. Geffken, K. and S. Spiegel, *Sphingosine kinase 1 in breast cancer*. Adv Biol Regul, 2018. **67**: p. 59-65.
203. Takabe, K., et al., *"Inside-out" signaling of sphingosine-1-phosphate: therapeutic targets*. Pharmacol Rev, 2008. **60**(2): p. 181-95.
204. Pei, G., et al., *Cellular stress promotes NOD1/2-dependent inflammation via the endogenous metabolite sphingosine-1-phosphate*. EMBO J, 2021. **40**(13): p. e106272.
205. Bamberger, E.S. and C.W. Perrett, *Angiogenesis in epithelial ovarian cancer*. Mol Pathol, 2002. **55**(6): p. 348-59.
206. Brown, M.R., J.O. Blanchette, and E.C. Kohn, *Angiogenesis in ovarian cancer*. Baillieres Best Pract Res Clin Obstet Gynaecol, 2000. **14**(6): p. 901-18.
207. Chen, Z. and Z.C. Han, *STAT3: a critical transcription activator in angiogenesis*. Med Res Rev, 2008. **28**(2): p. 185-200.
208. Priceman, S.J., et al., *S1PR1 is crucial for accumulation of regulatory T cells in tumors via STAT3*. Cell Rep, 2014. **6**(6): p. 992-999.
209. Lee, H., et al., *STAT3-induced S1PR1 expression is crucial for persistent STAT3 activation in tumors*. Nat Med, 2010. **16**(12): p. 1421-8.
210. Takabe, K. and S. Spiegel, *Export of sphingosine-1-phosphate and cancer progression*. J Lipid Res, 2014. **55**(9): p. 1839-46.
211. Liang, J., et al., *Sphingosine-1-phosphate links persistent STAT3 activation, chronic intestinal inflammation, and development of colitis-associated cancer*. Cancer Cell, 2013. **23**(1): p. 107-20.
212. Nagahashi, M., et al., *Targeting the SphK1/S1P/S1PR1 Axis That Links Obesity, Chronic Inflammation, and Breast Cancer Metastasis*. Cancer Res, 2018. **78**(7): p. 1713-1725.