Polycystic ovary syndrome – burden and adverse outcomes

by

Anuradhaa Subramanian

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

Institute of Applied Health

College of Medical and Dental Sciences

University of Birmingham

August 2022

Abstract

Background

Polycystic ovary syndrome (PCOS) is the most common, yet underdiagnosed, endocrine disorder in women of reproductive age, posing lifelong threats to their health with limited therapeutic options. Within UK primary care, there is a lack of evidence exploring the extent of missed PCOS diagnosis and possible ethnic variation in the incidence of PCOS. Furthermore, there is lack of (conclusive) evidence on the burden of PCOS including trends in the incidence of impaired glucose regulation outcomes, risk of adverse obstetric outcomes and susceptibility to COVID-19. Finally, the effectiveness of combined oral contraceptive pills (COCP), commonly prescribed for the management of PCOS symptoms, for longer term benefits such as prevention of impaired glucose regulation has not been explored. In this doctoral thesis, I undertook a series of retrospective studies using UK primary and secondary care data to address these research gaps.

Methods

A retrospective cohort of all eligible patients registered within the Clinical Practice Research Datalink (CPRD) Gold database between 1995 and 2019 was used to explore the incidence and prevalence trends of confirmed and possible PCOS diagnosis (based on relevant symptom codes) to estimate the extent of missed diagnosis, and the incidence estimates were stratified by ethnic subgroups. In subgroups of women with confirmed and possible PCOS diagnosis, incidence trends of type 2 diabetes, impaired glucose regulation (IGR) and gestational diabetes mellitus (GDM) were estimated. Using CPRD Gold Pregnancy Register and linked Hospital Episode Statistics data, pregnancies of women with PCOS between 1997 and 2020 were age matched to pregnancies of women without PCOS to observe the odds of four primary outcomes among the two groups of pregnancies: (1) preterm delivery, (2) mode of delivery, (3) high and low birthweight, and (4) stillbirth. Using contemporaneous data from The Health Improvement

Network (THIN) provided by Cegedim (January-July 2020), the hazard of confirmed or suspected COVID-19 was estimated among women with PCOS compared to an age matched cohort of women without PCOS. Finally, using a nested case-control design of a base cohort of women with PCOS identified from THIN, women with and without incident development of impaired glucose regulation were matched, and the odds of COCP prescription within a predefined exposure window was estimated.

Results

The incidence and prevalence of PCOS rose sharply in the year 2004, followed by stabilisation of the incidence rate. In addition to confirmed diagnosis, inclusion of symptom codes representing Rotterdam criteria resulted in 299 (95% CI, 198–299) missed PCOS diagnoses per 100,000 person-years. The prevalence of PCOS was highest among South Asians followed by Afro-Caribbeans. The incidence of type 2 diabetes, IGR and GDM has been rising in women with PCOS. Pregnancies of women with PCOS were at an increased risk of preterm and operative (emergency caesarean, elective caesarean and instrumental vaginal) delivery [adjusted OR: 1.11 (95% CI, 1.06-1.17), 1.10 (1.05-1.15), 1.07 (1.03-1.12), 1.04 (1.00-1.09), respectively]. Women with PCOS had a 28% increased risk of COVID-19 [adjusted HR: 1.28 (95% CI, 1.05-1.561)]. Women with PCOS and COCP use had a reduced risk of impaired glucose regulation [adjusted OR: 0.72, (95% CI, 0.59–0.87)].

Conclusion

There is a high level of missed PCOS diagnosis in primary care and an increasing ethnic disparity in the incidence of PCOS. The increased risk of adverse obstetric outcomes and susceptibility to COVID-19 should be clearly conveyed to women with PCOS. Future investigations should explore barriers to care and management of women with PCOS, especially from ethnic minority communities, and should examine the efficacy of COCP in prevention of impaired glucose regulation in a randomized controlled trial setting.

Acknowledgement

I would like to extend my deepest gratitude to my supervisors Prof. Krishnarajah Nirantharakumar, Prof. Wiebke Arlt and Associate Prof. Nicola Adderley who were inspiring, supportive, and encouraging throughout my postgraduate student life. Their passion towards their work is truly unmatched and they set great examples as academic leaders who pursued progress through collaborative science. They created an enriching environment for me that was academically, professionally, and emotionally fulfilling.

I would like to express my sincere thanks to my friend and colleague, Krishna M Gokhale, who developed the DExTER software for automated data extraction which efficiently expedited several aspects of my doctoral research.

I would also like to thank all the co-authors of the manuscripts that were published as part of this doctoral thesis. Their invaluable alliance, insights and feedback was critical for successful publication of these manuscripts.

I would like to acknowledge here the support, kindness and insights offered by all of my colleagues in the Institute of Applied Health Research and Institute of Metabolism and Systems Research.

I would specifically like to thank Dr. Siang Ing Lee who shared this doctoral journey with me and was my biggest cheerleader for all my triumphs and failures.

My parents Subramanian Venkatraman and Kala Subramanian instilled in me a mind with reason, passion, and ethics since early childhood, and for that I'm truly grateful to their way of

upbringing, without which I will not be the woman I am today. I would like to thank them for showing immense support and helping me cross countries to pursue higher education and a career of my choice.

Most importantly, I would like to thank my fiancé Bikash Deb who helped me through many of my emotional turmoil. Thank you for being an inspiration and showering me with unconditional love.

List of manuscripts and abstracts

This thesis incorporates the five following manuscripts, which corresponds to five observational studies conducted as part of Chapter 2, Chapter 3, Chapter 4, Chapter 5 and Chapter 6. Three out of the five manuscripts have been published and two manuscripts are now ready for submission.

(1) **Subramanian, A.,** Adderley, N.J., Toulis, K.A., Thangaratinam, S., Arlt, W., Nirantharakumar, K. Incidence and prevalence trends of polycystic ovary syndrome by ethnicity within UK primary care: 1995 – 2019 **(In preparation)**

(2) **Subramanian, A.,** Adderley, N.J., Thangaratinam, S., Arlt, W., Nirantharakumar, K. Impaired glucose regulation among women with PCOS: Incidence and prevalence trends recorded within UK primary care: 1995 – 2019 **(In preparation)**

(3) **Subramanian, A.,** Lee, S.I., Phillips, K., Toulis, K.A., Kempegowda, P., O'Reilly, M.W., Adderley, N.J., Thangaratinam, S., Arlt, W., & Nirantharakumar, K. (2022). Polycystic ovary syndrome and risk of adverse obstetric outcomes: a retrospective population-based matched cohort study in England. BMC Med 20, 298. https://doi.org/10.1186/s12916-022-02473-3 (**Published**)

(4) **Subramanian, A**., Anand, A., Adderley, N. J., Okoth, K., Toulis, K. A., Gokhale, K., Sainsbury, C., O'Reilly, M. W., Arlt, W., & Nirantharakumar, K. (2021). Increased COVID-19 infections in women with polycystic ovary syndrome: a population-based study. European Journal of Endocrinology, 184(5), 637–645. https://doi.org/10.1530/EJE-20-1163 **(Published)**

(5) Kumarendran, B., O'reilly, M. W., **Subramanian, A.,** Šsumilo, D., Toulis, K., Gokhale, K. M., Wijeratne, C. N., Coomarasamy, A., Tahrani, A. A., Azoulay, L., Arlt, W., & Nirantharakumar, K. (2021). Polycystic Ovary Syndrome, Combined Oral Contraceptives, and the Risk of Dysglycemia: A Population-Based Cohort Study With a Nested

Pharmacoepidemiological Case-Control Study. Diabetes Care, 44(12), 2758–2766. https://doi.org/10.2337/DC21-0437 **(Published)**

In addition, the following abstract arose from presentation of material in this thesis in the RCOG world congress 2022 and it was the top scoring oral presentation in the High-Risk Pregnancy Category:

(1) Category - High Risk Pregnancy. (2022). BJOG: An International Journal of Obstetrics & Gynaecology, 129(S1), 47–62. https://doi.org/10.1111/1471-0528.8_17178

Also, during the period of postgraduate study, several papers were published, which is presented in **Supplementary 1**.

Table of contents

List of tables

List of figures

List of abbreviations

AES	Androgen Excess Society
ARDS	Acute Respiratory Distress Syndrome
ASD	Autism Spectrum Disorder
BMI	Body Mass Index
COCP	Combined Oral Contraceptive Pills
COVID-19	COronaVIrus Disease - 2019
CPRD	Clinical Practice Research Datalink
CRP	C Reactive Protein
DHT	DiHydroTestosterone
FBG	Fasting Blood Glucose
GDM	Gestational Diabetes Mellitus
GLP-1	Glucagon Like Peptide-1
GP	General Practitioners
HA	HyperAndrogenism
HbA1c	Haemoglobin A1c
HES	Hospital Episode Statistics
HOMA-IR	HOmeostasis Model Assessment estimated-Insulin Resistance
HOMA-ISI	HOmeostasis Model Assessment estimated-Insulin Sensitivity Index
HPA	Hypothalamic-Pituitary-Adrenal
HR	Hazard Ratio
IGR	Impaired Glucose Regulation
IGT	Impaired Glucose Tolerance
IL	InterLeukin
IMD	Index of Multiple Deprivation
IQR	Interquartile Range
LGA	Large for Gestational Age
LNG-IUS	LevoNorGestrel-releasing IntraUterine System
MD	Mean Difference
NAFLD	Non-Alcoholic Fatty Liver Disease
NHANES	National Health and Nutrition Examination Survey
NHS	National Health Service
NICE	National Institute for health and Care and Excellence
NICHD	National Institute of Child health and Human Development
OCP	Oral Contraceptive Pill
OD	Ovulatory Dysfunction
OGTT	Oral Glucose Tolerance Test
OPCS	Operating Procedure Codes Supplement
OR	Odds Ratio

OSA	Obstructive Sleep Apnoea
PCO	PolyCystic Ovaries
PCOM	PolyCystic Ovarian Morphology
PCOS	PolyCystic Ovary Syndrome
POP	Progestogen Only Pill
QOF	Quality and Outcomes Framework
RCT	Randomized Controlled Trial
RECORD	REporting of studies Conducted using Observational Routinely-collected health Data
SARS-CoV-2	Severe Acute Respiratory Syndrome - CoronaVirus-2
SD	Standard Deviation
SGA	Small for Gestational Age
SHBG	Sex Hormone Binding Globulin
T2DM	Type 2 Diabetes Mellitus
THIN	The Health Improvement Network
TMPRSS2	TransMembrane PRroteaSe , Serine 2
TNF	Tumour Necrosis Factor
UK	United Kingdom
WHO	World Health Organization

Chapter 1 - Introduction

1.1 Prevalence of Polycystic ovary syndrome

Polycystic ovary syndrome (PCOS) is the most common endocrine disorder, that affects 4-20% of reproductive aged women.[1,2] PCOS is a lifelong metabolic condition, typically associated in most cases with androgen excess, anovulatory infertility and polycystic ovarian morphology on pelvic ultrasound.[3–5] PCOS has an estimated population prevalence between 8-16% of all women, depending on the population studied.[6,7] However, it is estimated that half of women with PCOS are undiagnosed, and the true prevalence is often difficult to estimate owing to delayed diagnosis due to the complex nature of the condition.[8] In the UK, real world evidence confirms that PCOS is a severely under-recognized condition.[9] Due to the degree of missed PCOS diagnosis by general practitioners (GPs) within primary care, a wide gap is expected between the true and diagnosed prevalence of PCOS within primary care in the UK. Notably, a diagnosis for PCOS is often made only when a woman presents to a reproductive endocrinologist with difficulty to conceive.[8] However, PCOS presentation is heterogeneous and can occur early in life even before puberty, posing silent metabolic threats to women's health.

1.2 Diagnostic criteria for PCOS

The diagnostic criteria for PCOS have evolved over the last three decades.[10] The National Institute of Child Health and Human Development (NICHD) first suggested in 1990 that the patient must demonstrate both oligo- or chronic anovulation and signs of androgen excess to be diagnosed with PCOS.[11] Androgen excess is defined by biochemical evidence of elevated circulating concentrations of androgenic steroids or by clinical signs of hyperandrogenism.[12] The latter usually presents symptomatically in the form of hirsutism, hair loss and acne, while biochemical androgen excess is evidenced by serum testosterone concentrations above the healthy female reference range, dependent on assay 2.0-2.5 nmol/L.[13]

The NICHD criteria was then revised by the Rotterdam criteria in 2003, that suggested patients must display two out of three symptoms to be diagnosed with PCOS, with polycystic ovaries (PCOs) as one of the symptoms considered in addition to anovulation and hyperandrogenism that were originally considered by the NICHD criteria for PCOS diagnosis.[14] While the prevalence of PCOS was noted to be 6-8% traditionally,[4,15,16] implementation of the Rotterdam criteria resulted in an increased prevalence estimate of up to 17% in the general population (including both diagnosed and undiagnosed PCOS),[17] with up to 69% of women in the community being undiagnosed.[18] This led to controversies where some perceived the expansion of the diagnostic criteria without any real improvement in care and outcomes to be inappropriate.[19–21] The latest revision of the PCOS diagnostic criteria in 2006 by the Androgen Excess Society (AES) then placed androgen excess as the defining feature of PCOS, with a concurrent presentation of either anovulation or PCO.[22] **Table 1.1** summarises the diagnostic criteria for PCOS, as they have evolved over time.

Table 1.1: Evolution of PCOS diagnostic guidance

Criteria	Definition
National Institute of Child Health and Human Development (NICHD) criteria (1990)	(1) Symptomatic presentation of androgen excess (hair loss, hirsutism, acne) **OR** Biochemical presentation of androgen excess (Serum testosterone ≥ 2.5 nmol/l) **AND** (2) anovulation
Rotterdam criteria (2003)	Two out of three of the following: (1) Symptomatic presentation of androgen excess (hair loss, hirsutism, acne) **OR** Biochemical presentation of androgen excess (Serum testosterone ≥ 2.5 nmol/l) (2) Anovulation (3) Polycystic ovaries
Androgen Excess Society criteria (2006)	(1) Symptomatic presentation of androgen excess (hair loss, hirsutism, acne) **OR** Biochemical presentation of androgen excess (Serum testosterone ≥ 2.5 nmol/l) **AND** **(2)** Anovulation **OR** Polycystic ovaries

1.3 Biochemical hallmarks of PCOS – androgen excess and insulin resistance

While the pathophysiology of PCOS may be a complex interplay of genetic and epigenetic abnormalities alongside ovarian, endocrine and neuroendocrine dysfunction, insulin resistance and metabolic alterations to anti-Müllerian hormone, androgen, and adiponectin levels,[23] the hallmarks of PCOS are typically hyperandrogenism and insulin resistance.

Systemic hyperandrogenism, essential for PCOS diagnosis based on the recently established AES diagnostic criteria, is chronically present in a mild, moderate or severe form among all PCOS phenotypes, and forms the best hormonal marker for the diagnosis of PCOS.[24] The source of excess androgen production has been found to be of intra- and extra-ovarian origin,

produced in both ovaries and adrenal glands respectively.[25] Invariably, hyperandrogenism has been found to be the crucial mediator between PCOS and the risk of several long-term adverse outcomes including lipotoxicity and dysglycaemia.[26,27]

In a recent publication, a team of researchers, including myself, at the University of Birmingham reported a gradient increase in the risk of type 2 diabetes with increase in serum testosterone levels in women.[27] This helped restate the well-established association between androgen excess, a biochemical hallmark of PCOS, and insulin resistance, manifested as type 2 diabetes/impaired glucose regulation (IGR). Despite the strong association between PCOS and IGR,[28–30] PCOS has not been given the attention it warrants within the diabetes community,[31] possibly due to disregard of the condition (PCOS) that is commonly diagnosed among young women who are not commonly considered to be at risk of developing type 2 diabetes. This gives rise to a healthcare related economic burden of PCOS-associated diabetes estimated at £237 million and $1.77 billion in the UK and USA respectively.[32,33] In addition to type 2 diabetes, several studies have also reported an increased risk of gestational diabetes mellitus (GDM) during pregnancy among women with PCOS.[34]

With the stable increase in the prevalence of PCOS over years between 2004 and 2014,[9] it is vital to understand the changing demands of PCOS on health services to manage its related complications, especially the development of IGR. However, there has been no assessment of time trends in the incidence and prevalence of IGR, type 2 diabetes, and gestational diabetes among the high-risk group of women with PCOS. Therefore, within Chapter 3 of this doctoral thesis, I aim to estimate annual incidence rates and prevalence of IGR (a composite of prediabetes and type 2 diabetes) and type 2 diabetes among a representative cohort of women with PCOS within UK primary care between 1995 and 2019. Furthermore, I aim to estimate annual incidence of GDM among pregnant women with PCOS.

1.4 Known risk factors for PCOS

A multitude of socio-demographic factors such as high BMI, deprivation and minority ethnic background, lifestyle factors such as sleep quality, gut microbiome, and environmental exposure to endocrine disruptors, as well as metabolic disturbances such as insulin resistance, hypertension, vitamin D deficiency and thyroid disorders, have been associated with the risk of being diagnosed with PCOS and exacerbation of the severity of the syndrome.[36–38] Within the multi-ethnic population of the United Kingdom (UK), there has been no study to date specifically assessing the link between ethnicity and the incidence and prevalence of PCOS over time.

Within the dynamic UK population, it is vital to understand the changing burden of PCOS over time and to identify distinct at-risk groups for PCOS, so that interventions can be designed to target specific population subgroups for prevention and care. Therefore, part of the next chapter of this doctoral thesis (Chapter 2) aims to explore the incidence and prevalence trends of PCOS between 1995 and 2019 within UK primary care, and to further identify ethnicity based higher-risk groups by assessing the ethnic variation in the incidence and prevalence trends of PCOS.

1.5 Long-term outcomes among women with PCOS

PCOS has long been considered as a reproductive disorder. In the wake of new findings regarding the long-term consequences of PCOS, the syndrome is now identified as a metabolic disorder that affects women throughout their lifetime, and even posing intergenerational risks to their children.[15,39,40] PCOS has long been known to co-occur with infertility and pregnancy related complications, leading to women with PCOS requiring assistance for conception, and special care during pregnancy and delivery. This may be attributable to their increased risk of inflammatory and metabolic disorders prior to pregnancy and increased obstetric risk during pregnancy.[41,42] Recent population studies have reported a high morbidity burden among

women with PCOS,[43] with obesity, insulin resistance and vitamin D deficiency observed among more than 50% of women with PCOS.

Previous studies have established that women with PCOS are at a higher risk of a spectrum of long-term health outcomes and may warrant close surveillance and monitoring for preventative interventions. These health outcomes include, but are not limited to IGR, type 2 diabetes,[30] non-alcoholic fatty liver disease (NAFLD),[26] obstructive sleep apnoea (OSA),[44] asthma,[45] migraine,[46] thyroid disorder,[47] cancer,[48] mental illnesses,[49] and sexual dysfunction.[50] Mendelian randomization studies have suggested that the association between PCOS and risk of cardiometabolic outcomes may not be causal, and attributable to common features of PCOS such as obesity, high testosterone, and low levels of sex-hormone binding globulin (SHBG).[51–53] Similar Mendelian randomization studies also suggest that genetically predicted PCOS is causally linked to specific types of breast cancers such as Endocrine Receptor (ER)-positive breast cancer.[54]

Furthermore, several systematic reviews have examined the association between maternal PCOS and the risk of a range of obstetric outcomes.[55–57] However, these reviews suggest varying results across the primary studies that they included owing to methodological heterogeneity, which included differences in terms of source population, criteria employed for PCOS ascertainment, and the potential confounders matched and adjusted for in their design and analysis respectively. Several of these primary studies are further limited in terms of outdated data, small sample size,[58,59] and restrictive selection of pregnant women who have undergone assisted reproduction[60,61] within their studies. In order to obtain conclusive evidence on the association between maternal PCOS and obstetric outcomes, within Chapter 4 of this doctoral thesis, I aimed to conduct a methodologically rigorous epidemiological study using contemporary, large, and representative data.

Notably, the long-term consequences of PCOS are not limited to the women affected by PCOS, but may also affect their offspring, creating an intergenerational continuum of endocrine disturbance. There is increasing evidence that exposure to prenatal androgen excess due to maternal PCOS is associated with increased risk of adverse outcomes among their offspring including developmental delays,[62] autism spectrum disorder (ASD), neuropsychiatric disorder, learning disability, behavioural and emotional disorder, anxiety disorder,[63] obesity, diabetes,[64] and PCOS.[65]

1.6 PCOS, COVID-19 pandemic and infection susceptibility

The novel severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) reached pandemic status in March 2020 (during my second academic year as a doctoral candidate registered at the University of Birmingham), with a consequent severe impact on international healthcare systems and the global economy.[66] The resulting coronavirus disease 2019 (COVID-19) caused, and continues to cause, mild symptoms in most cases, but the incidence of severe illness, respiratory failure and mortality in high-risk groups led to mandated quarantine measures and economic shutdown across the globe in order to protect capacity within health systems and intensive care units.[67,68] Multiple large observational studies were swiftly conducted to elucidate the risk factors for severe COVID-19.[69] Prior to the possibility of vaccination as a strategy, shielding strategies were recommended for older patients and for those with significant comorbidities that placed them in a high-risk bracket for severe COVID-19 infection, such as diabetes, heart, liver and lung disease, and being immunocompromised or pregnant.

Women with polycystic ovary syndrome (PCOS) were highlighted as an overlooked and potentially high risk population for COVID-19 complications.[70] Whilst younger age and female sex are typically associated with a lower overall risk of severe COVID-19 infection and mortality, it was believed that women with PCOS may represent a distinct subgroup of patients

at higher than average risk of adverse COVID-19-related outcomes due to other shared risk factors with COVID-19 such as ethnic pre-disposition, vitamin D deficiency, hyperinflammation, obesity and cardiometabolic disturbance.

PCOS has been recognised as a pro-inflammatory condition with chronic low grade inflammation evidenced by elevated levels of C-reactive protein among women with PCOS compared to women without PCOS.[71,72] Previous studies, although preliminary and whose findings cannot be construed as causal, have suggested a possible association between PCOS and susceptibility to infections with Chlamydia pneumonia, Chlamydia trachomatis and Helicobacter pylori.[73] In the wake of the COVID-19 pandemic, it was hypothesised that women with PCOS may be at a higher risk of contracting COVID-19, and/or developing a more severe form of COVID-19. Therefore, within Chapter 5 of this doctoral thesis, and within the remit of the timely available data, I aimed to determine whether PCOS was linked to COVID-19 susceptibility and published the findings of the study in European Journal of Endocrinology in May 2021.

1.7 Therapeutic management of PCOS

Despite the high prevalence of PCOS, the therapeutic options for women with PCOS are limited to the management of their symptoms. NICE guideline recommends weight loss and healthy lifestyle behaviour including healthy eating and increased physical activity as a first line treatment for women with PCOS to prevent longer term adverse health outcomes.[35] Pharmacological agents recommended for symptom management include prescription of cyclical progestogen for prolonged amenorrhea among women with PCOS, low-dose combined oral contraceptive pills (COCP) or levonorgestrel-releasing intrauterine system (LNG-IUS) for prevention of endometrial hyperplasia, and topical retinoids and hair reduction/removal methods for acne and hirsutism. In a qualitative study exploring the experiences of care received by women with PCOS from primary care, one of the main themes

which emerged was treatment for PCOS,[74] around which women with PCOS expressed concerns on the lack of treatment of the 'root-cause' of the condition. This leads to them relying upon multifactorial symptom management as their only option, thereby exposing them to polypharmacy.

Despite common use of COCPs, there is ongoing concern regarding the risk benefit profile of long-term COCP use among women with PCOS.[75] This is largely attributable to the lack of evidence on the possible preventative implications of the drug for this specific subgroup of women who are at risk of several adverse metabolic outcomes. An international evidence-based guideline for the assessment and management of PCOS recommends consideration of COCP as a treatment for adults and adolescents respectively for the management of hyperandrogenism and menstrual irregularity.[76] Literature suggests that COCP ameliorates hyperandrogenism through two mechanistic pathways: (1) by increasing the sex-hormone-binding globulin (SHBG) synthesised in the liver, thereby reducing free circulating testosterone (an action facilitated by the oestrogen component of COCP), and (2) by suppressing a surge in luteinizing hormone, thereby reducing excess androgen production by the ovaries (an action facilitated by the progestogen component of COCP).[77] Notably, an anti-androgenic progestin component could block and lower the effect of free circulating testosterone, thereby having a beneficial effect,[78] however there is little to no evidence regarding the addition of anti-androgenic component to COCP therapy.

I have previously contributed to a study in which a dose-response relationship between serum testosterone and the risk of developing IGR was established among women.[27] Considering the effect of COCP on reducing androgen excess and the effect of androgen on increased risk of IGR, it was hypothesised that COCP could have a protective effect among women with PCOS and prevent development of IGR in this high-risk group. Therefore, within Chapter 6 of this doctoral thesis, I aimed to estimate the real-world effectiveness of COCP on preventing IGR

among women with PCOS. This was performed through a methodologically rigorous pharmacoepidemiological study with a nested case control design using real-world primary care data.

While initiation of insulin-desensitizing drugs such as metformin is not recommended in primary care within the NICE guidelines,[35] an international evidence-based guideline for the assessment and management of PCOS recommends use of metformin alongside COCP when desirable management of the metabolic features of PCOS is not achieved by COCP and lifestyle modification alone.[76]

Notably, several advancements in weight management and glucose regulation have been made in recent years, such as bariatric surgery and Glucagon Like Peptide-1 (GLP-1) agonists, which may be applicable to the management of PCOS, a condition commonly observed among women with high BMI and IGR. However, the effectiveness of these advanced treatment strategies is not in routine dialogue and there is no research to date to estimate the effectiveness of these strategies among women with PCOS.

This doctoral thesis explores the hypothesis that women with PCOS are underdiagnosed and are burdened by adverse outcomes. It aimed to assess the extent of underdiagnosis and long-term consequences among women with PCOS, such as development of impaired glucose regulation, adverse obstetric outcomes during pregnancy, and susceptibility to infection, particularly SARS-CoV-2.

1.8 Specific aims of doctoral thesis

- Chapter 2 aimed to explore the incidence and prevalence trends of PCOS between 1995 and 2019 within UK primary care based on diagnostic code records, followed by a combination of symptom records representing each of the three diagnostic criteria used to define PCOS, to assess the extent of missed PCOS diagnosis. This chapter further

aimed to assess the ethnic variation in the incidence and prevalence trends of PCOS diagnosis.

- Chapter 3 aimed to estimate annual incidence rates and prevalence of IGR (a composite of prediabetes and type 2 diabetes) and type 2 diabetes among a representative cohort of women with PCOS within UK primary care between 1995 and 2019. This chapter further aimed to estimate annual incidence of GDM among pregnant women with PCOS.
- Chapter 4 aimed to assess the risk of obstetric outcomes (primary outcomes including preterm delivery, mode of delivery, high and low birthweight, and stillbirth, and secondary outcomes including very preterm delivery, extremely preterm delivery, and small and large for gestational age) among a population-based representative cohort of women with PCOS compared to an age matched cohort of women without PCOS.
- Chapter 5 aimed to examine the incident risk of reported suspected/confirmed COVID-19 in women with PCOS in the UK utilizing a large primary care database, in comparison to age matched control women without PCOS.
- Chapter 6 aimed to assess if use of COCP reduced the risk of development of IGR and type 2 diabetes among women with PCOS.
- Chapter 7 aimed to discuss the findings and implications of the findings and provide future directions for research to reduce the burden of PCOS.

Chapter 2 - Incidence and prevalence trends of polycystic ovary syndrome by ethnicity within UK primary care: 1995 – 2019

Manuscript ready for submission

Anuradhaa Subramanian[1], Nicola J Adderley[1], Konstantinos A Toulis[1], Shakila Thangaratinam[2], Wiebke Arlt[3], Krishnarajah Nirantharakumar[1,4]

[1] Institute of Applied Health Research, University of Birmingham, Birmingham, UK

[2] Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK

[3] National Institute for Health Research (NIHR) Birmingham Biomedical Research Centre, University Hospitals Birmingham NHS Foundation Trust and University of Birmingham, Birmingham, UK

[4] Midlands Health Data Research UK, Birmingham, UK

A.S conceived the idea for the study with guidance from W.A and K.N. A.S extracted the data using DExtER software. A.S designed the study and performed data cleaning and analysis. A.S wrote the initial draft of the manuscript. All authors reviewed and revised the manuscript. All authors read and approved the final manuscript.

2.1 Abstract

2.1.1 Introduction

In the UK, real world evidence suggests that PCOS is a severely under-recognized condition. In this study, I aimed to explore the incidence and prevalence trends of PCOS within UK primary care and assess variation in the incidence and prevalence of diagnosed PCOS by ethnicity.

2.1.2 Methods

A retrospective open cohort study of all eligible reproductive aged women was conducted between 1st of January 1995 and 31st December 2019 within the Clinical Practice Research Datalink (CPRD) Gold database. The study period was stratified into 25 yearly time periods, from 1995 to 2019. Annual incidence rates of PCOS were estimated based on diagnostic code records, followed by a combination of symptom records fulfilling each of three diagnostic criteria – National Institute of Child Health and Human Development (NICHD), Rotterdam, and Androgen Excess Society (AES) – to define PCOS. Cross-sectional analyses were performed at the start of each year to identify annual point prevalence estimates of PCOS. Finally, incidence and prevalence trend estimates were reported after stratifying by ethnicity.

2.1.3 Results and discussion

A total of 4,236,388 reproductive aged women were identified and followed up for a median of 4.81 years. There was a gradual increase in the incidence of PCOS based on diagnostic coding over the years until 2004, when the incidence leaped from 13.3 (in 2003) to 75.6 (in 2004) per 100,000 person-years. Diagnosis of PCOS based on a combination of symptom codes fulfilling each of the diagnostic criteria remained stable between the years 1995 (216, 259 and 234 per 100,000 person-years based on NICHD, Rotterdam and AES criteria, respectively) and 2014 (213, 295 and 272 per 100,000 person-years based on NICHD, Rotterdam and AES criteria, respectively), but it gradually dropped thereafter to 139, 200 and 185 per 100,000

person-years in 2019. Overall, the incidence and prevalence of PCOS throughout the study period was highest among South Asians, followed by Black Afro-Caribbeans, and ethnic difference in the prevalence of PCOS have widened over the last two and half decades.

2.1.4 Conclusion

There is a high level of missed PCOS diagnosis in primary care. Addressing this will facilitate improved risk stratification for long term health surveillance and care pathway implementation. Further investigations should be carried out to elucidate the reasons for lower reporting of PCOS related symptoms from 2014 onwards. Future research should focus on ethnic disparity in the management and access to care for women with PCOS.

2.2 Introduction

Polycystic ovary syndrome (PCOS) is the most common endocrine disorder, affecting 5-20% of the reproductive aged women. Due to the complexity in PCOS pathogenesis, the marked heterogeneity in its cardinal clinical manifestations and controversies in case definition, the confirmation of PCOS diagnosis is often delayed, thus the actual prevalence is often difficult to estimate and possibly variable in different subpopulations.

In specific, the multifactorial pathogenesis of the syndrome largely consists of ovarian and adrenal hyperandrogenism, resulting from genetic, hormonal, transgenerational and environmental components.[79] Specific gene loci (i.e. INS-VNTR, FSHR, and THADA), a blunted responsiveness to negative feedback control at the hypothalamic level and early prenatal androgen exposure have all been implicated in the pathogenesis of the syndrome.[79] This might be further complicated with insulin resistance and lipotoxicity.

The complexity and variability in aetiology is reflected in the heterogeneous set of symptoms PCOS may manifest. Chronic anovulation and oligomenorrhea, acne and hirsutism of variable intensity mark the effects of hyperandrogenism on the reproductive system, whereas obesity and long-term metabolic health risks complement the symptomatology with the metabolic aspect of the syndrome, at least in a significant subset of patients. Reproductive challenges and subfertility remain the single most important morbidity and a diagnosis for PCOS is often made only when a woman presents to a reproductive endocrinologist with difficulty in conceiving.[8]

This complexity has been further perplexed by the evolution of diagnostic guidance for PCOS.[10] The National Institute of Child Health and Human Development (NICHD)-proposed criteria,[11] requiring both oligo- or chronic anovulation and signs of hyperandrogenism for establishing the diagnosis, were replaced by the Rotterdam criteria in 2003, which maintained the minimum requirement of two symptoms, but expanded the phenotypic spectrum of the syndrome by recognising polycystic ovaries in ultrasound (PCO) as an additional phenotype.[14]

Naturally, the prevalence of PCOS, previously noted to be 6-8%,[4,15,16] rose to be up to 17% in the general population,[17] by the implementation of the Rotterdam criteria, with the majority of patients considered to be undiagnosed.[18] The latest revision of the PCOS diagnostic criteria in 2006 by the Androgen Excess Society (AES) then placed hyperandrogenism as the defining feature of PCOS, with a concurrent display of either anovulation or PCO.[22]

Whether the evolution of PCOS diagnostic criteria has been associated with actual improvement in health-care and reproductive and metabolic outcomes is still debatable.[19–21] However, the PCOS-related health care economic burden, associated with disease screening and treating its various morbidities is unequivocally significant. Therefore, any evidence that would permit a more targeted screening approach is warranted. Considering the definite genetic and transgenerational component in PCOS aetiology and the phenotypical variation in the presentation of PCOS across ethnic and racial sub-groups that have been reported before,[80–82] a formal exploration of differences in PCOS prevalence and incidence on the basis of ethnic origin would be needed.

In fact, a review study reporting pooled prevalence estimates of PCOS from various geographical locations (the United States, the United Kingdom, Spain, Greece, Australia, Asia, and Mexico) suggested the possible presence of ethnic difference in the prevalence of PCOS.[83] However, small sample size of individual studies, selection bias, and lack of comparability limit the internal and external validity of the findings. More recently, differences in PCOS prevalence on the basis of ethnic origin have also been reported in a single study involving a total of ~700 women with PCOS.[84] However, this elegant study was a secondary data analysis of a clinical trial, thus possibly not accurately reflecting community trends.

To address these shortcomings, I aimed to explore the incidence and prevalence trends of PCOS between 1995 and 2019 within UK primary care based on diagnostic code records, followed by a combination of symptom records representing each of the three diagnostic criteria used to

define PCOS, to assess the extent of missed PCOS diagnosis. I further aimed to assess, the ethnic variation in the incidence and prevalence trends of PCOS diagnosis.

2.3 Methods

2.3.1 Study design and data source

A retrospective open cohort study of all eligible reproductive aged women was conducted between 1st of January 1995 and 31st December 2019 within the Clinical Practice Research Datalink (CPRD) Gold database. CPRD Gold is a primary care database that contains pseudo-anonymised patient medical records of over 20 million patients from 973 general practices that use the Vision electronic healthcare records system, comprising around 7% of the UK population and representative in terms of age, sex and national mortality rates.[85] The patient medical records are comprised of information on patient demographics, symptoms and diagnoses, drug prescriptions, and physical and biochemical measurements from laboratories. A hierarchical clinical coding system called Read codes are used to record symptoms and diagnosis within CPRD Gold. Primary care databases from the UK are well established as reliable datasets for the study of incidence and prevalence trends of both rare and common disease conditions[86–90] and to study ethnic differences in the prevalence of conditions.[91]

2.3.2 Study population and follow-up

Practices were considered eligible to be included in the study one year after reporting Up To Standard date, a date within the CPRD Gold database from which practices are considered to consistently provide high quality data that is fit for research. Patients were eligible for inclusion one year after registration with an eligible general practice if they had an 'acceptable' patient flag, an indicator that determines the validity of the patient record.[85] Eligible women were followed up from their date of eligibility or when they turned 15 years old (whichever was the latest), and until the earliest of the following time points: date when a PCOS diagnosis was made or one of the PCOS definitions considered (given in the section below) was met, date of

de-registration from the practice, last date of data collection from the practice, study end date (31st December 2019), or women reaching 50 years of age.

2.3.3 PCOS definition

The primary PCOS exposure status, diagnosed PCOS, was ascertained by a Read code record of PCOS. Due to underdiagnosis of PCOS within primary care, I also considered additional PCOS definitions based on diagnostic criteria in guidelines. These definitions considered records of combinations of symptom codes indicating a PCOS diagnosis based on (1) NICHD criteria, (2) Rotterdam criteria, and (3) AES criteria (**Table 1.1**). The NICHD criteria was met when a woman had symptomatic or biochemical presentation of hyperandrogenism and a record of anovulation within her medical records. Read code records of hair loss, hirsutism or acne were considered to indicate a symptomatic presentation of hyperandrogenism. A record of serum testosterone greater than or equal to 2.5 nmol/l was considered as a biochemical presentation of hyperandrogenism. The Rotterdam criteria was met if a woman presented with two out of the following three symptoms: polycystic ovaries (PCO), hyperandrogenism and anovulation. The Androgen Excess Society criteria was met if a woman presented with hyperandrogenism and one of the other two symptoms – anovulation or PCO. The Read code lists for PCOS diagnosis and symptoms including hair loss, hirsutism, acne, anovulation and PCO are presented in **Supplementary 2, Supplementary 3, Supplementary 4, Supplementary 5, Supplementary 6** and **Supplementary 7** respectively.

In order to understand the true burden of PCOS, I considered an additional broader definition of PCOS, which was defined as a record of either a PCOS diagnostic code or symptoms meeting the Rotterdam criteria. Due to the possibility of misclassification between "PCOS" and "PCO" owing to their similarity, I considered another definition which additively included PCO in the previous definition.

2.3.4 Ethnicity

Ethnicity data recorded as Read codes within primary care through direct patient self-reports are considered gold standard. The wide range of available ethnicity Read codes were consolidated to represent five ethnicity subgroups as present in the 2011 UK census classification: 1) White (British, Irish, other White), 2) South Asian (Bangladeshi, Pakistani, Indian, Sri Lankan, British Asian or other South Asian), 3) Black Afro-Caribbean (Black African, Black Caribbean, Black British or other Black people), 4) mixed ethnicity, and 5)other minority ethnic groups (including Chinese, Vietnamese and other Southeast Asian). Read code lists for each ethnicity subgroup are presented in **Supplementary 8, Supplementary 9, Supplementary 10, Supplementary 11** and **Supplementary 12** in the order mentioned above. Where more than one ethnicity was recorded for a patient, the latest record was used to categorise the ethnicity variable. When there was no Read code record of ethnicity, ethnicity was categorised as missing.

2.3.5 Statistical Analysis

2.3.5.1 Incidence trend

The study period was stratified into 25 yearly time periods, from 1995 to 2019. For each year, the follow-up time started on 1st of January that year or the start of the patient's follow-up (whichever was the latest) and ended on 31st of December that year or the end of patient's follow-up (whichever was the earliest). For each year, incidence rate of PCOS was calculated by dividing the total number of patients newly diagnosed with PCOS/fulfilling the PCOS definition of interest divided by the total person-years of follow-up for that year.

2.3.5.2 Prevalence trend

A series of cross-sectional analyses were performed on 1st January of each year from 1995 to 2019 to calculate a series of annual point prevalence estimates. The total number of women

meeting the PCOS definition prior to the start of the year of interest was divided by the total number of eligible women aged 15-50 at the start of the year of interest.

2.3.5.3 Ethnic difference

The incidence and prevalence trends of PCOS were estimated stratified by ethnicity, again using definitions of PCOS based on diagnostic coding and each of the above-mentioned diagnostic criteria.

In an additional analysis, I performed a time-dependent Cox proportional hazards regression model with annual intervals to estimate unadjusted and adjusted hazard ratios of PCOS diagnosis among women from each of the ethnic minority subgroups compared to women of White ethnicity. Year of follow-up, age and BMI recorded prior to the year of follow-up were considered as covariates to adjust for in the model. BMI was categorised according to the World Health Organization (WHO) criteria as underweight (<18.5 kg/m^2), normal weight (18.5-25 kg/m^2), overweight (25-35 kg/m^2) and obese (≥35 kg/m^2). In case of missing BMI data, a separate missing category for this was created. Age at the start of follow-up year was categorised into 5-years age bands as 15-20, 20-25, 25-30, 30-35, 35-40 ,40-45, and 45-50 years. Furthermore, I ran the regression analysis with an interaction term between ethnicity and BMI categories to estimate the compounded effects of BMI and ethnicity subgroups.

Data cleaning and incidence/prevalence estimations were performed in Stata IC version 15. Selection of Read code lists was performed using an inhouse developed software platform called Code Builder.

2.3.6 Ethics

Observational research using CPRD data was approved by National Research Ethics Service Committee. This study has been approved by the Independent Scientific Advisory Committee (Reference: 21_000412). As CPRD data are anonymized, individual patients are not required to give consent for the use of their data.

2.4 Results

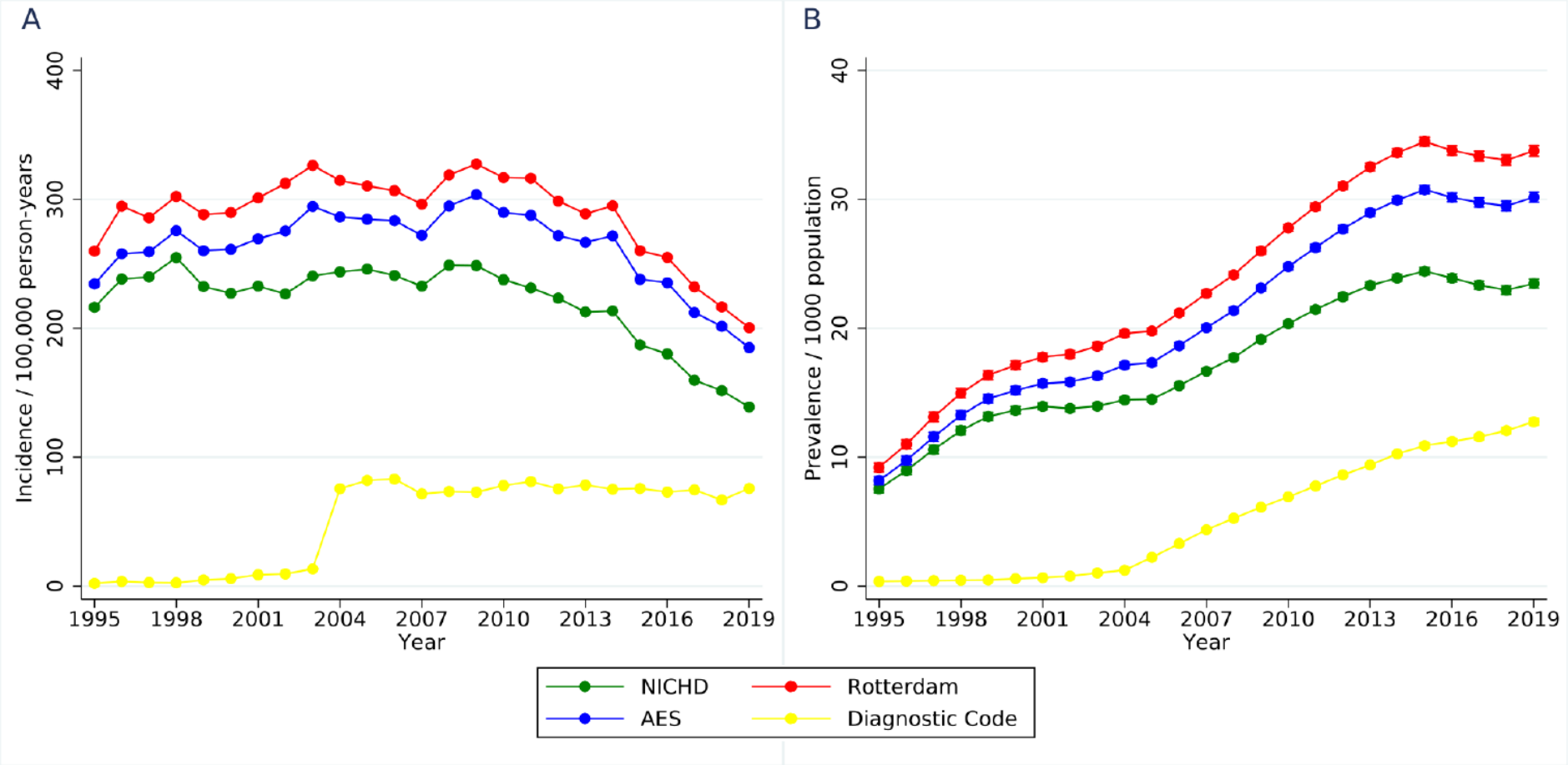
2.4.1 Characteristics of the cohort

An eligible cohort of 6,604,130 women were identified, of which 4,236,388 were eligible during their reproductive age of 15-50. The mean age of included women was 28.13 (Standard Deviation, SD 12.10) years and they were followed up for a median of 4.81 (Interquartile Range, IQR 1.81-9.88) years. Of the included women, 45.2% (n=1,914,189) had a coded record of ethnicity. Of these, 80.38% (n=1,536,710) were White Caucasian, 8.9% (n=170,443) were Black Afro-Caribbean, 4.6%, (n=88,495) were South-Asian, 1.2% (n=22,800) were mixed ethnicity, and 5.0% (n=95,741) were from other ethnic minority groups.

2.4.2 Incidence trend of PCOS

The PCOS incidence trend estimates are presented in **Supplementary 13** and **Figure 2.1**. Based on diagnostic codes alone, the incidence rate of PCOS in the year 1995 was 2.1 (95% CI, 1.9-2.3) per 100,000 person-years. There was a gradual increase in the incidence rate over the years until 2004, when the incidence leaped from 13.3 (95% CI, 13.1-13.5) in 2003 to 75.6 (95% CI, 75.5-76.1) in 2004, per 100,000 person-years. The incidence rate of PCOS remained stable between 2004 and 2019.

Figure 2.1: Incidence (A) and prevalence (B) trends of PCOS based on: (1) diagnostic code for PCOS alone, (2) combination of symptom codes fulfilling National Institute of Child Health and Human Development criteria for PCOS diagnosis (NICHD), (3) combination of symptom codes fulfilling Rotterdam criteria for PCOS diagnosis, and (4) combination of symptom codes fulfilling Androgen Excess Society (AES) criteria for PCOS diagnosis



Diagnosis of PCOS based on a combination of symptom codes fulfilling each of the diagnostic criteria remained stable between the years 1995 [216 (95% CI, 215-218), 260 (95% CI, 258 -261) and 234 (95% CI, 233 -236) per 100,000 person-years based on NICHD, Rotterdam and AES criteria, respectively)] and 2014 [213 (95% CI, 213 -214), 295 (95% CI, 294 -296) and 272 (95% CI, 271 -273) per 100,000 person-years based on NICHD, Rotterdam and AES criteria, respectively], but it gradually dropped from 2014 onwards. In the year 2019, the recording of symptoms was the lowest and incidence of PCOS based on the combination of symptom codes fulfilling NICHD criteria was 139 (95% CI, 138 -140), Rotterdam criteria was 200 (95% CI, 200 -201) and AES criteria was 185 (95% CI, 184-186) per 100,000 person-years. Overall, diagnosis based on the Rotterdam criteria was the highest throughout the year, followed by AES criteria, followed by NICHD criteria.

Figure 2.2: Incidence (A) and prevalence (B) trends of PCOS based on: (1) a diagnostic code for PCOS alone, (2) a diagnostic code for PCOS or a combination of symptom codes fulfilling Rotterdam criteria for PCOS diagnosis, (3) a diagnostic code for PCOS or a combination of symptom codes fulfilling Rotterdam criteria for PCOS diagnosis or a code for PCO.

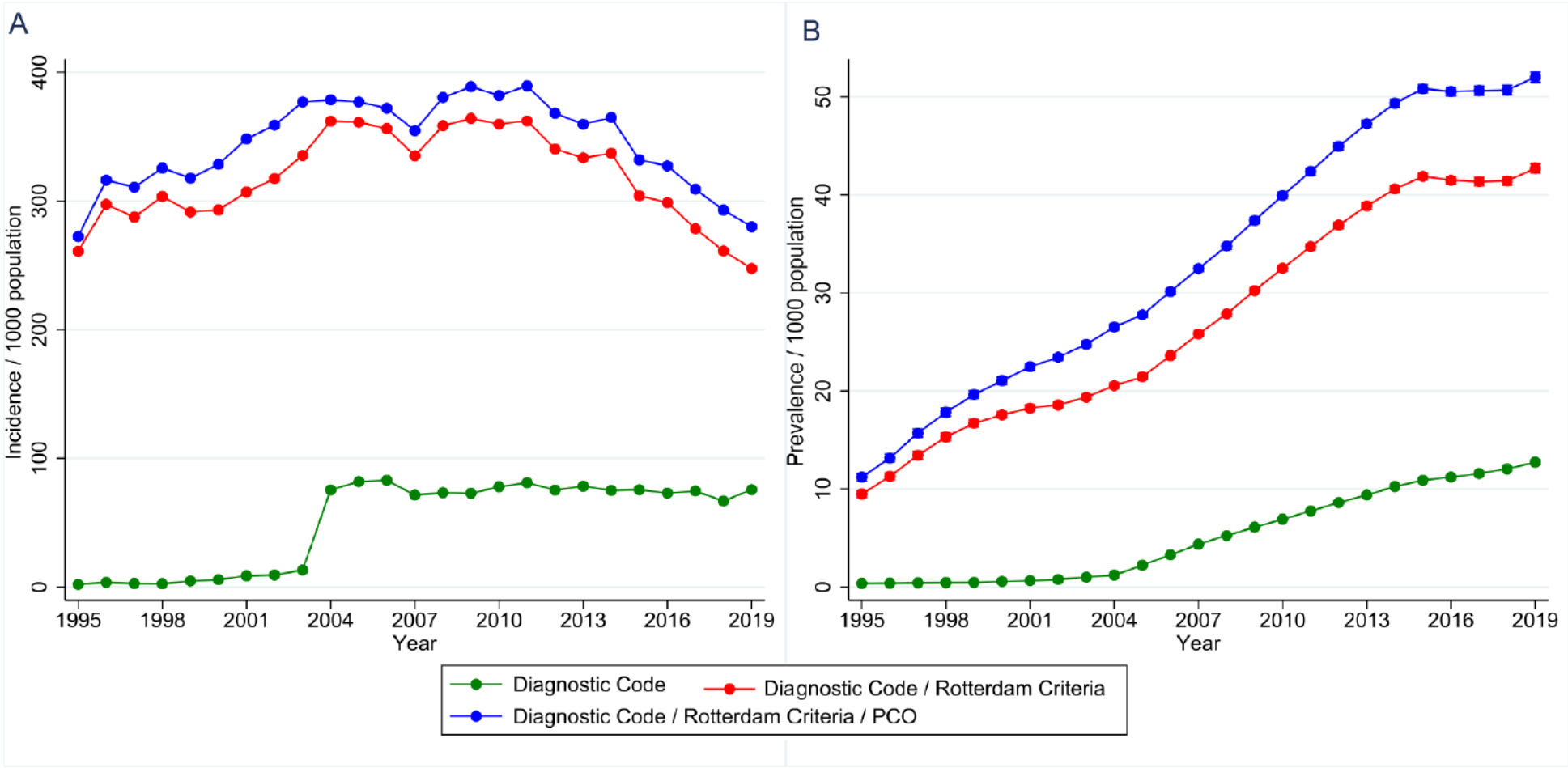


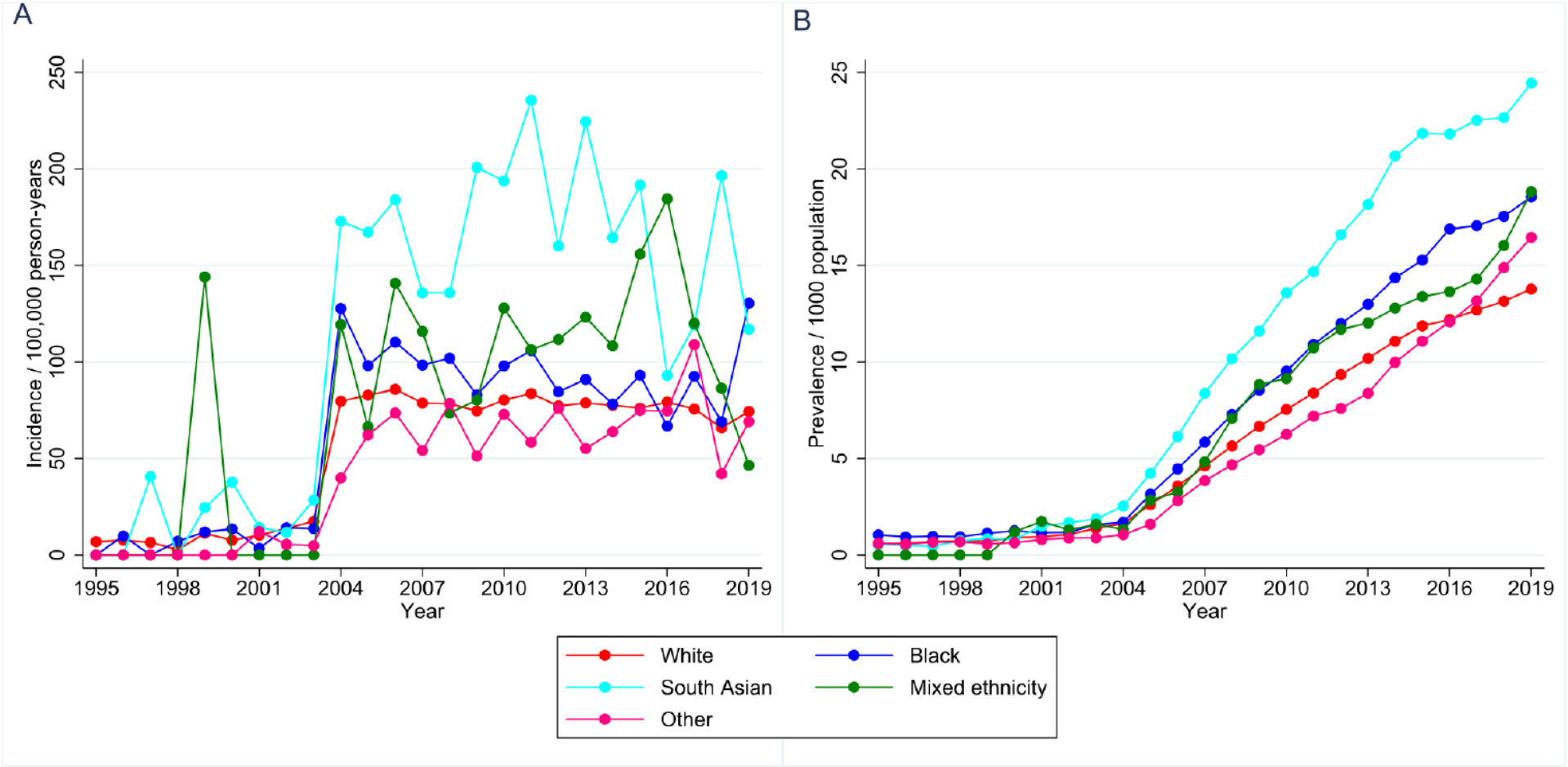
Figure 2.2 presents the incidence trend of PCOS based on the additive definitions considering: (1) diagnostic code or symptoms codes fulfilling Rotterdam criteria, (2) diagnostic code, symptoms codes fulfilling Rotterdam criteria or a code for PCO. The overall difference in the incidence estimates by additionally including (1) Rotterdam criteria, and (2) Rotterdam criteria along with PCO diagnostic codes in the definition of PCOS before the year 2004 was 299 (95% CI, 198– 299) and 332 (95% CI, 332 -332) per 100,000 person-years, respectively. From the year 2004 onwards, the overall difference in the annual incidence estimates by including the additional criteria reduced to 260 (95% CI, 260 -260) and 284 (95% CI, 284 -284) per 100,000 person-years, respectively.

2.4.3 Prevalence trend of PCOS

The prevalence trend estimates of PCOS are presented in **Supplementary 14 a**nd **Figure 2.1.** Based on diagnostic code alone, there was an increasing trend in the prevalence trend of PCOS, with a steeper increase from 2004 onwards. Based on the combination of symptom codes fulfilling the PCOS diagnostic criteria, there was a gradual increase in prevalence from 1995 to 2005, with a steeper increase until 2015, followed by a slight decrease or levelling off. In 2019, the prevalence of PCOS based on diagnostic criteria, symptom code combinations fulfilling NICHD, Rotterdam and AES criteria was 12.7 (95% CI, 12.5-13.0), 23.5 (95% CI, 13.1-23.8), 33.8 (95% CI, 33.3-34.2) and 30.2 (95% CI, 29.8-29.9) per 1,000 reproductive aged women, respectively.

Figure 2.2 presents the prevalence trend of PCOS based on the additive PCOS definitions. The difference in the prevalence estimates by additionally including (1) Rotterdam criteria, and (2) Rotterdam criteria along with PCO diagnostic codes were 9.1 (95% CI, 8.8-9.4) and 10.8 (95% CI, 10.5-11.1) per 1,000 women respectively in the year 1995. This difference gradually increased until 2015 to 31.0 (95% CI, 30.8-31.2) and 39.9 (95% CI, 39.7-40.2) per 1000 women, respectively, following which it remained stable.

Figure 2.3: Incidence and prevalence trends of PCOS based on PCOS diagnostic codes stratified by ethnicity status



2.4.4 Incidence trend of PCOS by ethnicity

Figure 2.3 presents the incidence trend of PCOS based on diagnostics code records stratified by ethnicity. Due to small sample sizes within each of the ethnic subgroups, the incidence rates were prone to large variation across the years. However, the incidence rate of PCOS sharply increased from 2004 onwards for each of the individual ethnic groups. Overall, between 1995 and 2019 the incidence rate of PCOS was highest among the South Asian ethnicity, followed by mixed ethnicity, Black Afro-Caribbean, White Caucasian, and other ethnic minorities [South Asian 157 (95% CI, 156 -158); mixed ethnicity 104 (95% CI, 102 -106); Black Afro-Caribbean 81 (95% CI, 80 -81); White Caucasian 68 (95% CI, 68 -68); other ethnicity 51 (95% CI, 51 -52) per 100,000 person years].

Supplementary 15, Supplementary 16 and **Supplementary 17** presents the incidence trend of PCOS based on combinations of symptom codes fulfilling NICHD, Rotterdam and AES criteria respectively, stratified by ethnicity; they each show similar trends by ethnicity.

2.4.5 Prevalence trend of PCOS by ethnicity

Figure 2.3 presents the prevalence trend of PCOS based on diagnostic codes stratified by ethnicity. The annual point prevalence estimates at the start of each year stratified by ethnicity did not differ distinguishably between ethnic subgroups until the year 2004, primarily due to relatively low levels of diagnosis. Since 2004, the ethnic difference in the prevalence of PCOS has gradually widened. By the year 2019, the prevalence of PCOS based on diagnostic codes among the South Asian community was 24.4 (95% CI, 22.15-26.7) per 1,000 women, followed by Black Afro-Caribbean and mixed ethnicity [18.5 (95% CI, 17.1-20.0) and 18.8 (95% CI, 14.9-22.8) per 1,000 women, respectively]. It was the lowest in other ethnic minority groups and White Caucasian ethnicity [16.5 (95% CI, 14.3-18.6) and 13.8 (95% CI, 13.4-14.2) per 1,000 women, respectively].

Supplementary 15, Supplementary 16 and **Supplementary 17** present the prevalence trend of PCOS stratified by ethnicity, based on combinations of symptom codes fulfilling NICHD, Rotterdam and AES criteria respectively, and the trends were similar based on all PCOS definitions, while scaled highest using Rotterdam criteria, followed by AES and NICHD criteria (**Supplementary 18** and **Supplementary 19**).

Table 2.1: Association between ethnicity, age, body mass index and year of follow-up and risk of PCOS diagnosis

Variables	Adjusted Hazard Ratio (95% CI)
Ethnicity	
Caucasian	Ref
Black Afro-Caribbean	1.33 (1.24-1.43)
South Asian	2.39 (2.21-2.58)
Mixed ethnicity	1.29 (1.06-1.57)
Other ethnic minorities	1.35 (1.19-1.52)
Missing	0.90 (0.87-0.93)
Age Categories	
15-20	Ref
20-25	1.63 (0.58-4.58)
25-30	1.01 (0.16-6.34)
30-35	9.04 (0.89-92.30)
35-40	2.76 (0.13-56.85)
40-45	4.98 (0.11-230.10)
45-50	.
Body mass index	
Underweight (<18.5 kg/m^2)	0.72 (0.64-0.81)
Normal weight (18.5-25 kg/m^2)	Ref
Overweight (25-30 kg/m^2)	1.97 (1.87-2.07)
Obese (30-35 kg/m^2)	5.38 (5.16-5.60)
Missing	0.71 (0.67-0.75)
Year	
1995	Ref
1996	1.7 (0.63-4.59)
1997	1.49 (0.55-4.02)
1998	1.57 (0.60-4.08)
1999	2.73 (1.13-6.59)
2000	2.76 (1.16-6.56)
2001	4.61 (2.01-10.6)
2002	5.93 (2.61-13.47)
2003	7.37 (3.27-16.65)
2004	35.48 (15.90-79.19)
2005	39.1 (17.53-87.22)
2006	38.27 (17.16-85.36)
2007	31.7 (14.21-70.73)
2008	30.98 (13.89-69.13)
2009	30.41 (13.63-67.85)
2010	31.74 (14.23-70.8)
2011	32.95 (14.77-73.5)
2012	29.81 (13.36-66.52)
2013	31.12 (13.95-69.44)
2014	29.41 (13.18-65.64)
2015	30.18 (13.52-67.39)
2016	28.81 (12.89-64.36)

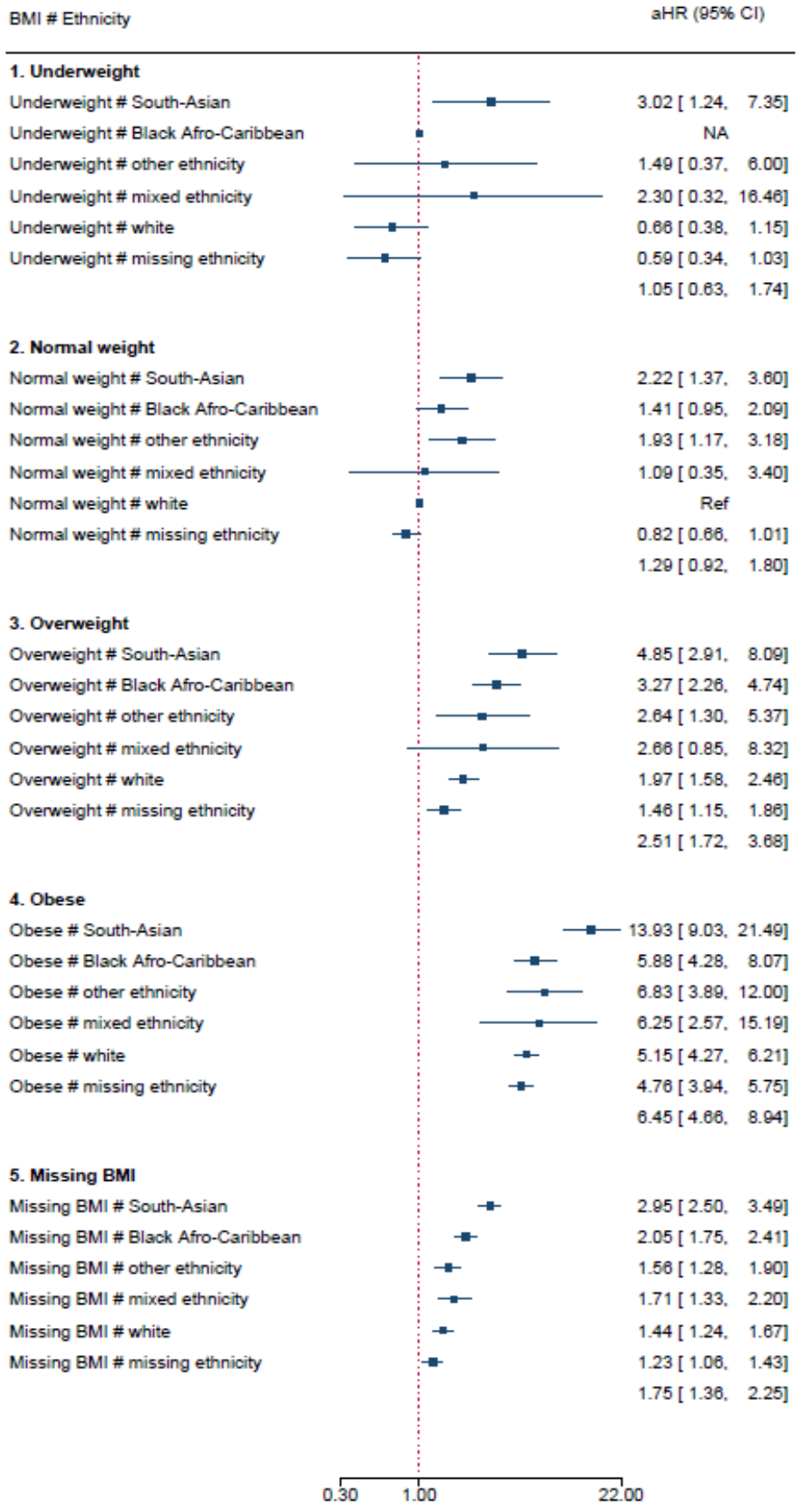
2017	29.33 (13.12-65.54)
2018	26.03 (11.64-58.21)
2019	29.92 (13.38-66.91)
2020	20.49 (9.13-45.97)

2.4.6 Ethnicity, body mass index (BMI), age, and year as risk factors for PCOS diagnosis

Table 2.1 presents the results of the Cox proportional hazards regression model estimating the association between PCOS diagnosis and risk factors including ethnicity, body mass index (BMI), age, and year of follow-up. Compared to women of White Caucasian ethnicity, women of Black Afro-Caribbean, South-Asian, mixed ethnicity and other ethnic minority groups were at a significantly higher risk of PCOS diagnosis [Adjusted hazard ratio (aHR): 1.33 (95% CI, 1.24-1.43), 2.39 (95% CI, 2.21-2.58), 1.29 (95% CI, 1.06-1.57) and 1.35 (95% CI, 1.19-1.52) respectively]. There was no evidence of association between age and PCOS diagnosis. Compared to women with normal BMI, women who were overweight and obese were at increased risk of PCOS diagnosis [aHR: 1.97 (95% CI, 1.87-2.07) and 5.38 (95% CI, 5.16-5.60) respectively]. Notably, women with missing ethnicity records and missing BMI were at a lower risk of diagnosis compared to their corresponding reference groups [0.90 (95% CI, 0.87-0.93) and 0.71 (95% CI, 0.67-0.75), respectively]. Compared to the year 1995, there was a statistically significant increased risk of PCOS diagnosis from the year 1999 onwards, which increased dramatically in the year 2004 [aHR: 35.48 (95% CI, 15.90-79.19)].

In the analysis with the interaction term between ethnicity and BMI subgroups, the results were similar (**Supplementary 20)**. Women of South Asian ethnicity had the highest risk of PCOS diagnosis across all BMI groups, while women with White ethnicity had the lowest (**Figure 2.4)**.

Figure 2.4: Interaction between ethnicity and BMI on the risk of PCOS diagnosis



2.5 Discussion

2.5.1 Summary of findings

In this population-based cohort study using primary care data, a sharp rise in the incidence rate of PCOS diagnosis was found in the year 2004, followed by the incidence rate stabilizing in the subsequent years. Similarly, an increase in the prevalence of PCOS was found throughout the study period, with a sharp increase from 2004 onwards. A decline in PCOS incidence and stabilization of PCOS prevalence was also found based on symptom recording from the year 2014 onwards.

South Asians, followed by Black Afro-Caribbeans, had the highest overall incidence and prevalence of PCOS throughout the study period, and for the first time, the study reports a widening ethnic difference in the prevalence of PCOS over the last two and half decades.

2.5.2 Concurrence with guidelines and literature

The sharp increase in the incidence of PCOS in the year 2004 coincides with the implementation of the broad Rotterdam criteria for PCOS diagnosis as well as the implementation of Quality and Outcomes Framework to improve recording of conditions within primary care. Using the Rotterdam criteria, women without hyperandrogenism or anovulation were diagnosed with PCOS if they had at least one of these two symptoms along with ultrasonographic evidence of polycystic ovaries. This broadened the phenotypic spectrum of PCOS and included those with ovulatory PCOS and non-hyperandrogenic PCOS. According to Sachdeva et al., these phenotypes constitute the milder variants of PCOS, and represent around 17.7% and 3.6% of the total cases of women with a PCOS diagnosis in an Indian population.[92] Notably, while the Androgen Excess Society criteria for PCOS diagnosis was proposed in 2006, the National Institute for Health and Care Excellence (NICE) still recommends Rotterdam criteria for PCOS diagnosis within the UK.[93]

The gradual decline in the incidence rate of PCOS based on a combination of symptom codes may be attributable to the Royal College of Obstetricians and Gynecologists guideline published in 2014 on Long-term Consequences of Polycystic Ovary Syndrome,[94] whereby the guideline stresses auditable standards of 100% accurate diagnosis of PCOS defined by at least two out of the three Rotterdam criteria symptoms.

Corresponding with these findings, previously published literature has described PCOS as an underdiagnosed condition within UK primary care.[1,9] The Quality and Outcomes Framework (QOF) has previously been reported to impact clinician behavior to improve diagnosis, management and care for patients with chronic conditions.[95] While PCOS has not been included as a disease entity in QOF, there has been mass lobbying supporting the inclusion of PCOS within the QOF list of conditions,[96] so that PCOS can be better identified and treated, preventing poor prognosis of the condition.

The findings on ethnic differences in the incidence and prevalence of PCOS also corresponds with the limited literature available. Dayo et al., reported a higher prevalence of PCOS among South Asian ethnicity (3.5%), and lower prevalence among the Chinese community (1.1%), compared to White ethnic women (1.6%) in a multi-ethnic society in Northern California.[84]

2.5.3 Strengths and limitations

This study has many strengths including the large sample size and the representativeness of the data to the UK population to accurately mirror the diagnoses that are routinely made within UK general practices. The study period for this analysis spanned between 1995 to 2019 allowing observation of trends in the incidence and prevalence estimates of PCOS over the last two and half decades, during which various guidelines have been introduced, following the evolution of PCOS diagnosis. While the primary care setting of this study did not allow reporting of the true prevalence of PCOS as observed in the community, it allowed inspection

of changes in PCOS prevalence based on diagnostic guidelines and provided an indication of the extent of missed PCOS diagnoses.

2.5.4 Implications

The analysis indicates that there are high levels of missed PCOS diagnosis in primary care. Addressing this will facilitate improved risk stratification for long term health surveillance and care pathway implementation. Further investigations should be made to elucidate the reasons for lower reporting of PCOS related symptoms from 2014 onwards. Future research should focus on ethnic disparity in the management and access to care for women with PCOS.

Chapter 3 - Impaired glucose regulation among women with PCOS: Incidence and prevalence trends recorded within UK primary care, 1995 - 2019

Manuscript ready for submission

Anuradhaa Subramanian[1], Nicola J Adderley[1], Shakila Thangaratinam[2], Wiebke Arlt[3], Krishnarajah Nirantharakumar[1,4]

[1]Institute of Applied Health Research, University of Birmingham, Birmingham, UK

[2]Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK

[3]National Institute for Health Research (NIHR) Birmingham Biomedical Research Centre, University Hospitals Birmingham NHS Foundation Trust and University of Birmingham, Birmingham, UK

[4]Midlands Health Data Research UK, Birmingham, UK

A.S conceived the idea for the study with guidance from W.A and K.N. A.S extracted the data using DExtER software. A.S designed the study and performed data cleaning and analysis. A.S wrote the initial draft of the manuscript. All authors reviewed and revised the manuscript. All authors read and approved the final manuscript.

3.1 Abstract

3.1.1 Introduction

With the steady increase in the prevalence of Polycystic ovary syndrome (PCOS) over the last two and half decades, it is vital to understand the changing demands of PCOS on health services to manage its' related complications, especially the development of impaired glucose regulation (IGR). Therefore, I aimed to assess time trends in the incidence and prevalence of IGR, type 2 diabetes, and gestational diabetes mellitus (GDM) among women with PCOS.

3.1.2 Methods

Between 1995 and 2020, using the Clinical Practice Research Datalink (CPRD) Gold database and its' associated Pregnancy Register database, women with PCOS exposure ascertained based on either a coded diagnosis of PCOS or a combination of symptom recording indicating PCOS diagnosis based on various diagnostic criteria [National Institute of Child Health and Human Development (NICHD), Rotterdam and Androgen Excess Society (AES)] were included. A series of annual cohort and annual cross-sectional data at the start of each year were obtained based on patient eligibility and PCOS exposure. Annual incident rates and prevalence were estimated for IGR and type 2 diabetes, wherein the outcome diagnosis was ascertained based on Read code diagnosis, HbA1c and fasting blood glucose measurements fulfilling the National Institute for Health and Care and Excellence (NICE) guideline for IGR and type 2 diabetes diagnosis. Annual incidence of GDM was estimated based on Read code diagnosis of GDM among women who are eligible within the pregnancy register

3.1.3 Results

A total of 32,649, 79,654, 112,555 and 100,161 women had a Read code diagnosis of PCOS, and PCOS diagnosed based on NICHD, Rotterdam and AES criteria respectively. Throughout the study period, women with a diagnostic code for PCOS had the highest

incidence rate for IGR (1666 per 100,000 person-years), followed by women with a combination of symptom codes indicating Rotterdam criteria (1017 per 100,000 person-years), AES criteria (992 per 100,000 person-years), and NICHD criteria (930 per 100,000 person-years). A gradual increasing trend in both the incidence rate and prevalence of both IGR and type 2 diabetes was observed among women with PCOS with a sharper increase in the estimates between 2012-2013.

Among pregnancies of women with a diagnostic code of PCOS, combination of symptom codes representing PCOS diagnosis based on NICHD, Rotterdam and AES criteria, the incidence of GDM observed between 2000 and 2019 (during which Pregnancy Register data was available) was on a rise, and the overall incidence was 27.9, 11.7, 13.9 and 12.9 per 1,000 pregnancies respectively.

3.1.4 Conclusion

With the increasing incidence of impaired glucose regulation among women with PCOS, awareness of the effectiveness of existing interventions such as combined oral contraceptives and metformin in prevention against dysglycaemia should be raised.

3.2 Introduction

Polycystic ovary syndrome (PCOS) is a common endocrine disorder with metabolic manifestations throughout a women's lifespan. Several published studies have assessed and established the increased prevalence and incident risk of impaired glucose regulation (IGR) and type 2 diabetes among women with PCOS, irrespective of their bodyweight. [28–30] Overall, PCOS related IGR/type 2 diabetes gives rise to a healthcare related economic burden of £237 million and $1.77 billion in the UK and USA respectively.[32,33] The link between PCOS and development of IGR has been suggested to be driven by androgen excess and insulin resistance underpinning PCOS, and a range of risk factors have been recognized to facilitate the development of IGR, including high body mass index (BMI), belonging to an ethnic minority, environmental exposure to endocrine disruptors, vitamin D deficiency, gut microbiome, and sleep quality.[38] In line with published evidence and guidelines,[3,22,94,97] the National Institute for Health and Care Excellence (NICE) recommends screening of women with PCOS for type 2 diabetes if they are overweight, have a personal history of gestational diabetes, have a family history of type 2 diabetes or if they are not of white Caucasian ethnicity.[35] However, there is no agreed upon consensus or recommendations on screening intervals to suggest continual monitoring for IGR or type 2 diabetes, in order to facilitate timely therapeutic intervention of this high risk population.

In addition to IGR, several studies have also reported an increased risk of gestational diabetes mellitus (GDM) during pregnancy among women with PCOS,[34] but literature suggests that this increased risk may be confounded by higher pre-gravid BMI, fertility treatment and older age at pregnancy. [98,99]

With the steady increase in the prevalence of PCOS over the last two and half decades, it is vital to understand the changing demands of PCOS on health services to manage its' related complications, especially the development of IGR. However, there has been no assessment of

time trends in the incidence and prevalence of IGR, type 2 diabetes, and gestational diabetes among the high-risk group of women with PCOS. Therefore, I aimed to estimate annual incidence rates and prevalence of IGR (a composite of prediabetes and type 2 diabetes) and type 2 diabetes among a representative cohort of women with PCOS within UK primary care between 1995 and 2019. Furthermore, I aimed to estimate annual incidence of GDM among pregnant women with PCOS

3.3 Methods

3.3.1 Study design

Between 1995 and 2019, using the Clinical Practice Research Datalink (CPRD) Gold, a primary care database, I conducted a series of:

(1) annual retrospective open cohort studies to estimate the annual incidence rates of IGR and type 2 diabetes among reproductive aged women with PCOS

(2) cross-sectional studies at the start of every year to estimate the point prevalence estimates of IGR and type 2 diabetes among reproductive aged women with PCOS.

Between 2000 and 2019, using the linked pregnancy register data, I conducted a series of:

(3) annual cohort studies to estimate the annual incidence of GDM among pregnancies of reproductive aged women with PCOS.

3.3.2 Data source

CPRD Gold contains pseudo-anonymised patient medical records from general practices that use Vision system and constitutes a representative data sample of 7% of the UK population. It contains data on (1) demographics including year of birth and sex, (2) diagnoses and symptoms recorded as Read codes, a hierarchical clinical coding system, (3) prescriptions recorded as drug codes, and (4) physical and biochemical measurements such as body mass index and serum testosterone. Pregnancy related data recorded within CPRD Gold between 2000 and

2019 has been used to generate a linked pregnancy register with information on pregnancy start and end dates and outcome of the pregnancy.[100]

3.3.3 Study population

Practices contributing data to CPRD Gold were eligible one year after reporting Up To Standard date, a measure of practice data quality.[85] Women with an acceptable patient flag (indicating research quality data)[85] were eligible for inclusion at the time of recording of PCOS, one year after registration with an eligible practice, or the date when they reached the reproductive age of 15, whichever was the latest. Four base cohorts of women with PCOS were identified, where PCOS exposure was ascertained through recording of one of the following: (1) diagnostic code for PCOS, (2) a combination of symptom codes indicating PCOS based on the National Institute of Child Health and Human Development (NICHD) criteria, (3) a combination of symptom codes indicating PCOS based on the Rotterdam criteria, (4) a combination of symptom codes indicating PCOS based on the Androgen Excess Society (AES) criteria (**Table 1.1**).[10] Further details of PCOS ascertainment criteria are mentioned in section 2.3.3 PCOS definition. Patients with a diagnosis of type 1 diabetes at any time point were excluded.

3.3.4 Follow-up to estimate incidence rate of IGR and type 2 diabetes among women with PCOS:

For each year, women were followed up from January 1st of that year, or the date of their eligibility, whichever was the latest until the earliest of the following time points: (1) December 31st of that year, (2) patient death, (3) patient transfer out of practice, (4) practice ending data contribution to CPRD Gold, or (5) date of ascertainment of IGR or type 2 diabetes.

3.3.5 Impaired Glucose definition:

In accordance with the NICE diagnostic guidelines (**Supplementary 21**),[101] IGR (a composite outcome of prediabetes/type 2 diabetes) was indicated by any of the following: (1) Read code

diagnosis of impaired glucose regulation or type 2 diabetes, (2) record of fasting blood glucose (FBG) > 5.5 mmol/L, or (3) record of Haemoglobin A1c (HbA1c) > 6.0% (42 mmol/mol) . Type 2 diabetes was indicated by any of the following: (1) Read code diagnosis of type 2 diabetes, (2) record of fasting blood glucose (FBG) > 7 mmol/L, or (3) record of HbA1c > 6.5% (47.5 mmol/mol). Gestational diabetes was ascertained by a Read code diagnosis The Read code lists for IGR, type 2 diabetes, GDM and type 1 diabetes are presented in **Supplementary 22, Supplementary 23, Supplementary 24** and **Supplementary 25** respectively.

3.3.6 Statistical Analysis

3.3.6.1 Annual incidence rates of IGR and type 2 diabetes among women with PCOS:

Each year, the incidence rate of IGR and type 2 diabetes was calculated by dividing the total number of at-risk women with PCOS with the outcome of interest during that year divided by the total person-years of follow-up during that year.

3.3.6.2 Point prevalence trends of IGR and type 2 diabetes among women with PCOS:

At the start of each year, the point prevalence was calculated by dividing the total number of eligible women with PCOS with the outcome of interest prior to the start of that year by the total of eligible women with PCOS at the start of that year.

3.3.6.3 Annual incidence of GDM among women with PCOS:

Each year, the denominator cohort consisted of pregnancies (with pregnancy start date in the year of interest) of women with PCOS (recorded prior to the start of pregnancy). Each year, the incidence of GDM was estimated by dividing the total number of pregnancies with a record of GDM between the pregnancy start and end date divided by the total number of pregnancies in the denominator cohort for that year.

All analyses were performed in Stata IC version 17. Data extraction from CPRD Gold and selection of Read code lists was performed using in-house developed software platforms called DExtER[102] and Code Builder respectively.

3.3.7 Ethics

Observational research using CPRD data was approved by National Research Ethics Service Committee. This study has been approved by the Independent Scientific Advisory Committee (Reference: 21_000412). As CPRD data are anonymized, individual patients are not required to give consent for the use of their data.

3.4 Results

3.4.1 Characteristics of the included cohort

During the study period, a total of 32,649, 79,654, 112,555 and 100,161 women were eligible for each of the four cohorts with PCOS diagnostic code, combination of symptom codes indicating PCOS diagnosis based on NICHD criteria, Rotterdam criteria and AES criteria respectively. Overall mean age at eligibility and median follow-up of patients included in any one of the four cohorts were 29.6 (SD, Standard Deviation 8.0) years and 4.3 (IQR, Interquartile range 1.7-8.5) years respectively.**Table 3.1** and **Table 3.2** summarizes age and ethnicity, annually between 1995 and 2019, of the patients with PCOS diagnosis at risk of developing type 2 diabetes and impaired glucose regulation respectively. Over the years, there has been a decrease in mean (SD) age of women with a diagnosis of PCOS at the time of entering the annual cohort. There is also improvement in the recording of ethnicity over time.

Table 3.1: Characteristics of the yearly cohort of patients at risk of type 2 diabetes

Cohort year	Number of patients	Age (Mean (SD)	Ethnicity, n (%)					
			White	Black Afro-Caribbean	South Asian	Mixed ethnicity	Other ethnicity	Missing
1995	4778	26.86 (7.39)	1077 (22.54)	155 (3.24)	46 (0.96)	9 (0.19)	143 (2.99)	3348 (70.07)
1996	6005	26.51 (7.47)	1390 (23.15)	220 (3.66)	58 (0.97)	12 (0.20)	170 (2.83)	4155 (69.19)
1997	8059	26.58 (7.62)	2007 (24.90)	309 (3.83)	84 (1.04)	13 (0.16)	231 (2.87)	5415 (67.19)
1998	10410	26.69 (7.82)	2679 (25.73)	406 (3.90)	127 (1.22)	18 (0.17)	302 (2.90)	6878 (66.07)
1999	13618	27.05 (8.06)	3679 (27.02)	552 (4.05)	199 (1.46)	27 (0.20)	392 (2.88)	8769 (64.39)
2000	17880	27.40 (8.28)	4912 (27.47)	753 (4.21)	257 (1.44)	44 (0.25)	517 (2.89)	11397 (63.74)
2001	22220	27.40 (8.40)	6326 (28.47)	991 (4.46)	360 (1.62)	59 (0.27)	636 (2.86)	13848 (62.32)
2002	27115	27.37 (8.51)	8042 (29.66)	1252 (4.62)	456 (1.68)	73 (0.27)	756 (2.79)	16536 (60.98)
2003	32536	27.24 (8.61)	9989 (30.70)	1561 (4.80)	581 (1.79)	102 (0.31)	883 (2.71)	19420 (59.69)
2004	39089	27.19 (8.80)	12501 (31.98)	2028 (5.19)	737 (1.89)	117 (0.30)	990 (2.53)	22716 (58.11)
2005	45208	26.98 (8.94)	14978 (33.13)	2476 (5.48)	892 (1.97)	146 (0.32)	1119 (2.48)	25597 (56.62)
2006	50045	26.66 (9.04)	17213 (34.40)	2817 (5.63)	1091 (2.18)	167 (0.33)	1224 (2.45)	27533 (55.02)
2007	54033	26.41 (9.14)	20361 (37.68)	3162 (5.85)	1330 (2.46)	225 (0.42)	1302 (2.41)	27653 (51.18)
2008	58293	26.17 (9.25)	24085 (41.32)	3566 (6.12)	1666 (2.86)	290 (0.50)	1440 (2.47)	27246 (46.74)
2009	62084	25.85 (9.34)	27582 (44.43)	3915 (6.31)	1974 (3.18)	350 (0.56)	1543 (2.49)	26720 (43.04)
2010	65553	25.60 (9.42)	30423 (46.41)	4317 (6.59)	2316 (3.53)	410 (0.63)	1660 (2.53)	26427 (40.31)
2011	67612	25.31 (9.55)	32459 (48.01)	4710 (6.97)	2649 (3.92)	446 (0.66)	1723 (2.55)	25625 (37.90)
2012	69059	25.11 (9.66)	33681 (48.77)	4883 (7.07)	2976 (4.31)	480 (0.70)	1810 (2.62)	25229 (36.53)
2013	69821	24.90 (9.77)	34772 (49.80)	5016 (7.18)	3026 (4.33)	513 (0.73)	1812 (2.60)	24682 (35.35)
2014	66843	24.65 (9.91)	33289 (49.80)	4783 (7.16)	2798 (4.19)	496 (0.74)	1664 (2.49)	23813 (35.63)
2015	61027	24.54 (10.01)	30583 (50.11)	4430 (7.26)	2529 (4.14)	424 (0.69)	1412 (2.31)	21649 (35.47)
2016	50381	24.37 (10.12)	25114 (49.85)	3587 (7.12)	1882 (3.74)	346 (0.69)	1128 (2.24)	18324 (36.37)
2017	44370	24.30 (10.20)	22639 (51.02)	3141 (7.08)	1786 (4.03)	301 (0.68)	978 (2.20)	15525 (34.99)
2018	40271	24.21 (10.30)	20441 (50.76)	2758 (6.85)	1652 (4.10)	273 (0.68)	891 (2.21)	14256 (35.40)
2019	37755	24.07 (10.46)	19205 (50.87)	2428 (6.43)	1452 (3.85)	266 (0.70)	822 (2.18)	13582 (35.97)

Table 3.2: Characteristics of the yearly cohort of patients at risk of impaired glucose regulation

Cohort year	Number of patients	Age (Mean (SD)	Ethnicity, n (%)					
			White	Black Afro-Caribbean	South Asian	Mixed ethnicity	Other ethnicity	Missing
1995	4778	26.86 (7.39)	1077 (22.54)	155 (3.24)	46 (0.96)	9 (0.19)	143 (2.99)	3348 (70.07)
1996	6004	26.51 (7.47)	1389 (23.13)	220 (3.66)	58 (0.97)	12 (0.20)	170 (2.83)	4155 (69.20)
1997	8055	26.58 (7.62)	2004 (24.88)	309 (3.84)	84 (1.04)	13 (0.16)	231 (2.87)	5414 (67.21)
1998	10402	26.68 (7.82)	2675 (25.72)	406 (3.90)	127 (1.22)	18 (0.17)	301 (2.89)	6875 (66.09)
1999	13599	27.04 (8.06)	3673 (27.01)	552 (4.06)	199 (1.46)	27 (0.20)	391 (2.88)	8757 (64.39)
2000	17852	27.40 (8.28)	4904 (27.47)	753 (4.22)	257 (1.44)	44 (0.25)	515 (2.88)	11379 (63.74)
2001	22168	27.39 (8.40)	6316 (28.49)	991 (4.47)	360 (1.62)	59 (0.27)	632 (2.85)	13810 (62.30)
2002	27046	27.37 (8.51)	8029 (29.69)	1250 (4.62)	454 (1.68)	73 (0.27)	752 (2.78)	16488 (60.96)
2003	32441	27.23 (8.61)	9964 (30.71)	1557 (4.80)	578 (1.78)	101 (0.31)	880 (2.71)	19361 (59.68)
2004	38925	27.17 (8.80)	12456 (32.00)	2016 (5.18)	729 (1.87)	116 (0.30)	984 (2.53)	22624 (58.12)
2005	44975	26.96 (8.93)	14917 (33.17)	2457 (5.46)	882 (1.96)	145 (0.32)	1110 (2.47)	25464 (56.62)
2006	49719	26.64 (9.04)	17119 (34.43)	2790 (5.61)	1071 (2.15)	164 (0.33)	1213 (2.44)	27362 (55.03)
2007	53634	26.38 (9.14)	20242 (37.74)	3125 (5.83)	1304 (2.43)	221 (0.41)	1290 (2.41)	27452 (51.18)
2008	57806	26.14 (9.25)	23934 (41.40)	3519 (6.09)	1636 (2.83)	287 (0.50)	1422 (2.46)	27008 (46.72)
2009	61547	25.81 (9.34)	27399 (44.52)	3864 (6.28)	1944 (3.16)	347 (0.56)	1517 (2.46)	26476 (43.02)
2010	64909	25.55 (9.41)	30168 (46.48)	4255 (6.56)	2268 (3.49)	406 (0.63)	1634 (2.52)	26178 (40.33)
2011	66887	25.26 (9.54)	32155 (48.07)	4639 (6.94)	2587 (3.87)	442 (0.66)	1697 (2.54)	25367 (37.93)
2012	68235	25.06 (9.66)	33342 (48.86)	4796 (7.03)	2902 (4.25)	475 (0.70)	1778 (2.61)	24942 (36.55)
2013	68877	24.84 (9.77)	34374 (49.91)	4916 (7.14)	2936 (4.26)	503 (0.73)	1778 (2.58)	24370 (35.38)
2014	65797	24.59 (9.91)	32841 (49.91)	4669 (7.10)	2703 (4.11)	489 (0.74)	1624 (2.47)	23471 (35.67)
2015	59991	24.46 (10.01)	30137 (50.24)	4301 (7.17)	2421 (4.04)	420 (0.70)	1368 (2.28)	21344 (35.58)
2016	49485	24.29 (10.12)	24747 (50.01)	3482 (7.04)	1794 (3.63)	341 (0.69)	1088 (2.20)	18033 (36.44)
2017	43490	24.22 (10.20)	22274 (51.22)	3031 (6.97)	1690 (3.89)	295 (0.68)	935 (2.15)	15265 (35.10)
2018	39442	24.13 (10.30)	20110 (50.99)	2663 (6.75)	1548 (3.92)	269 (0.68)	848 (2.15)	14004 (35.51)
2019	36897	23.98 (10.46)	18862 (51.12)	2335 (6.33)	1365 (3.70)	258 (0.70)	777 (2.11)	13300 (36.05)

3.4.2 Incidence trend of IGR and type 2 diabetes

The incidence rate trend estimates of IGR and type 2 diabetes among reproductive aged women with PCOS are presented in **Supplementary 26**, **Supplementary 27**, **Figure 3.1**, **Figure 3.2, Figure 3.3** and **Figure 3.4.**

Throughout the study period, women with a diagnostic code for PCOS had the highest incidence rate for IGR (1666 per 100,000 person-years), followed by women with a combination of symptom codes indicating Rotterdam criteria (1017 per 100,000 person-years), AES criteria (992 per 100,000 person-years), and NICHD criteria (930 per 100,000 person-years). A gradual increasing trend in the incidence rate of IGR throughout the study period was observed among women with PCOS in all four cohorts (average yearly percentage increase in the incidence rate of IGR: 8.6%, 12.0%, 11.5% and 10.7% among women with a diagnostic code for PCOS, and women with a combination of symptom codes indicating PCOS based on NICHD, Rotterdam and AES criteria, respectively). Due to the lower sample size in the cohort of women identified using PCOS diagnostic codes, several temporal spikes were observed in the incidence rates over the study period; however, an overall increasing trend was observed, as for other definitions of PCOS. A sharp increase in the incidence of IGR was observed between 2012-2013 (37.6%, 43.4%, 41.1% and 43.4% across the four cohorts) and 2017-2018 (12.2%, 26.4%, 22.5% and 22.6% across the four cohorts) among women with PCOS.

Similar to the incidence rate of IGR, the overall incidence rate of type 2 diabetes throughout the study period was highest among the cohort of women with a diagnosis of PCOS (1,314 per 100,000 person-years), followed by the cohort of women with a combination of symptom codes indicating PCOS diagnosis based on Rotterdam, AES and NICHD criteria (802, 783 and 729 per 100,000 person-years, respectively).

Figure 3.1: (A) Incidence (rate) trend of impaired glucose regulation and type 2 diabetes per 1,000 person-years and incidence trend of gestational diabetes mellitus per 1,000 pregnancies of women with a diagnostic code for PCOS, and (B) prevalence trend of impaired glucose regulation and type 2 diabetes per 1,000 women with a diagnostic code for PCOS

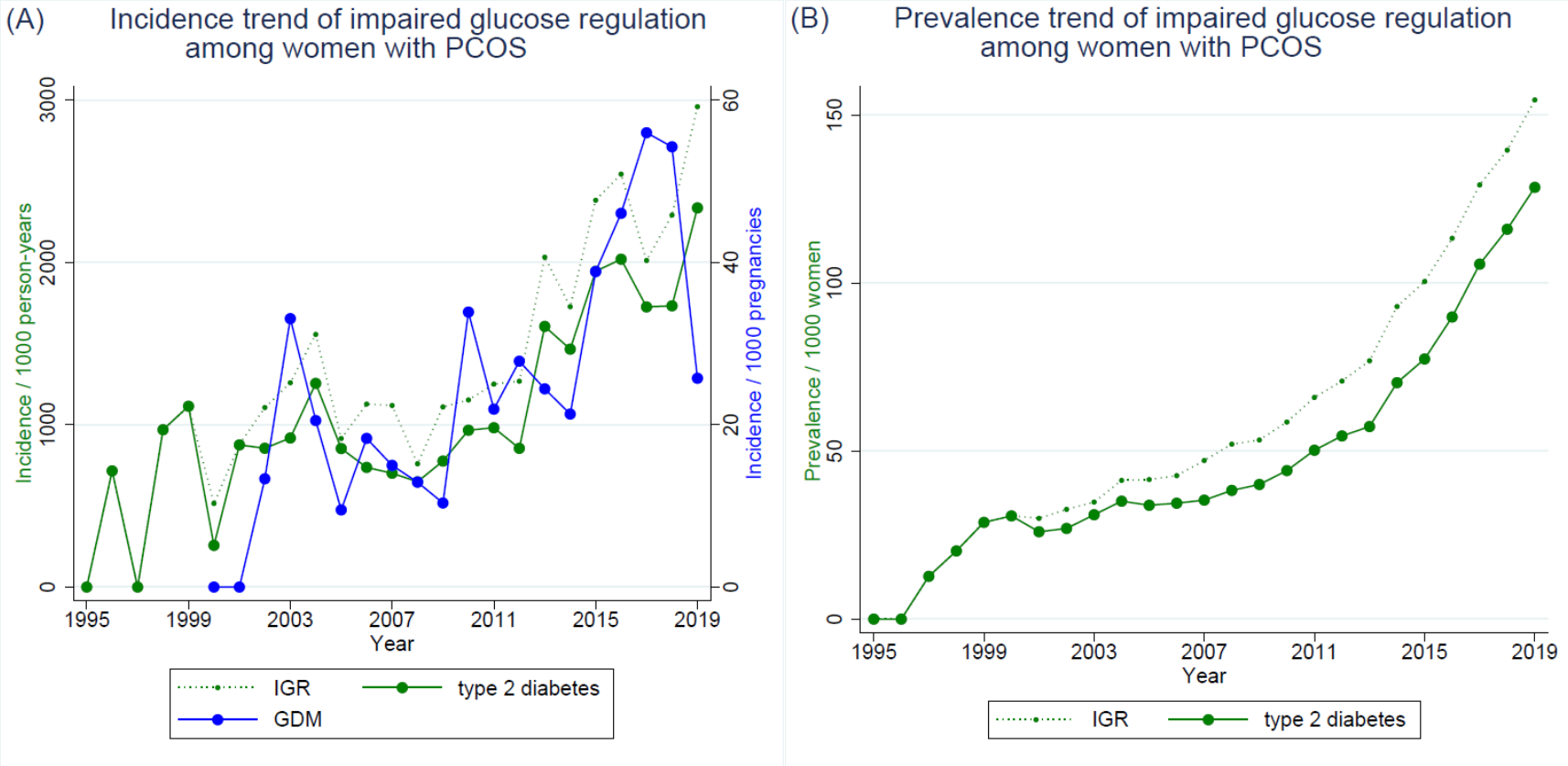


Figure 3.2: (A) Incidence (rate) trend of impaired glucose regulation and type 2 diabetes per 1,000 person-years and incidence trend of gestational diabetes mellitus per 1,000 pregnancies of women with combination of symptoms codes indicating PCOS diagnosis based on NICHD criteria and (B) prevalence trend of impaired glucose regulation and type 2 diabetes per 1,000 women with combination of symptoms codes indicating PCOS diagnosis based on NICHD criteria

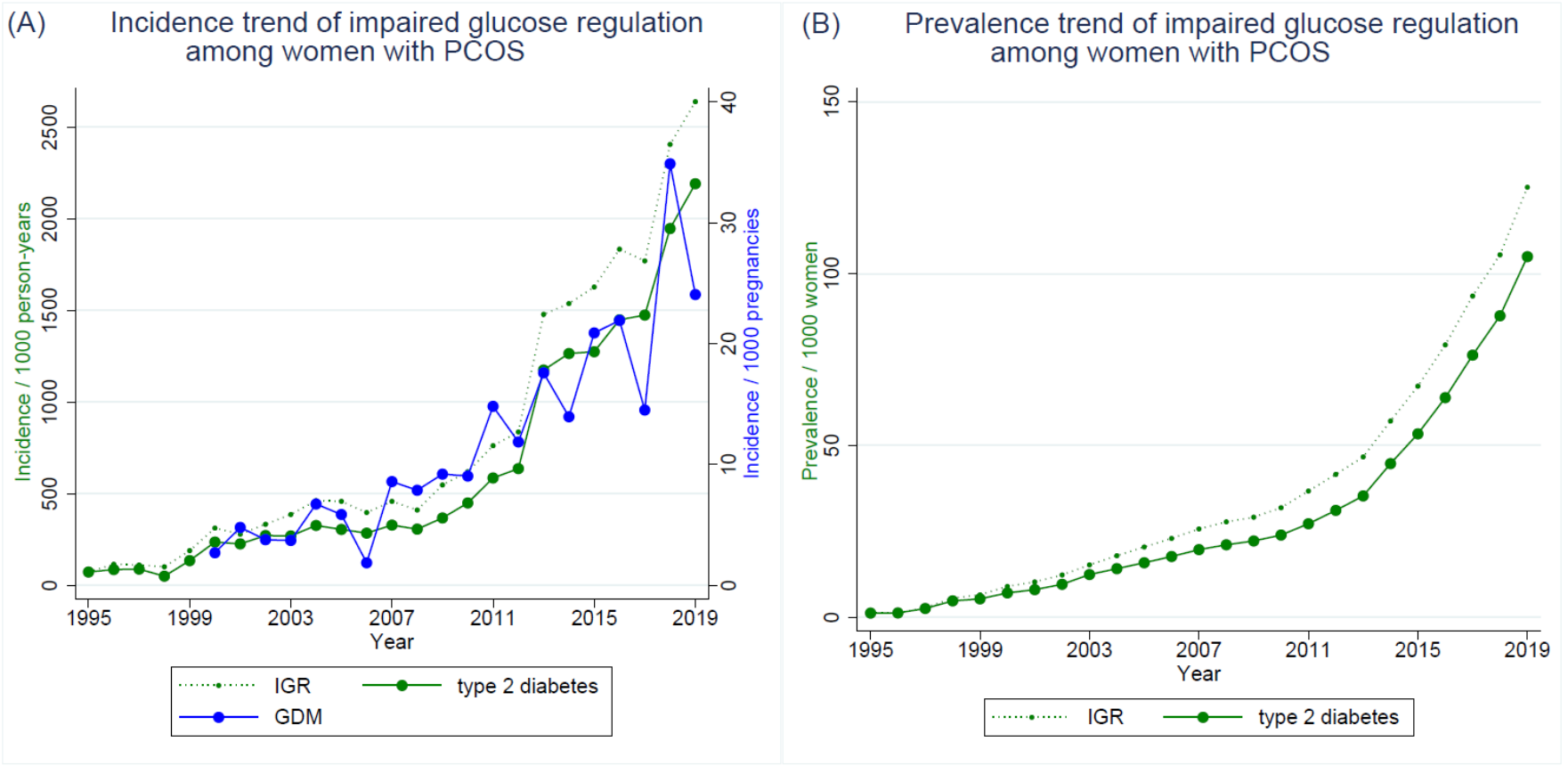


Figure 3.3: A) Incidence (rate) trend of impaired glucose regulation and type 2 diabetes per 1,000 person-years and incidence trend of gestational diabetes mellitus per 1,000 pregnancies of women with combination of symptoms codes indicating PCOS diagnosis based on Rotterdam criteria, and (B) prevalence trend of impaired glucose regulation and type 2 diabetes per 1,000 women with combination of symptoms codes indicating PCOS diagnosis based on Rotterdam criteria

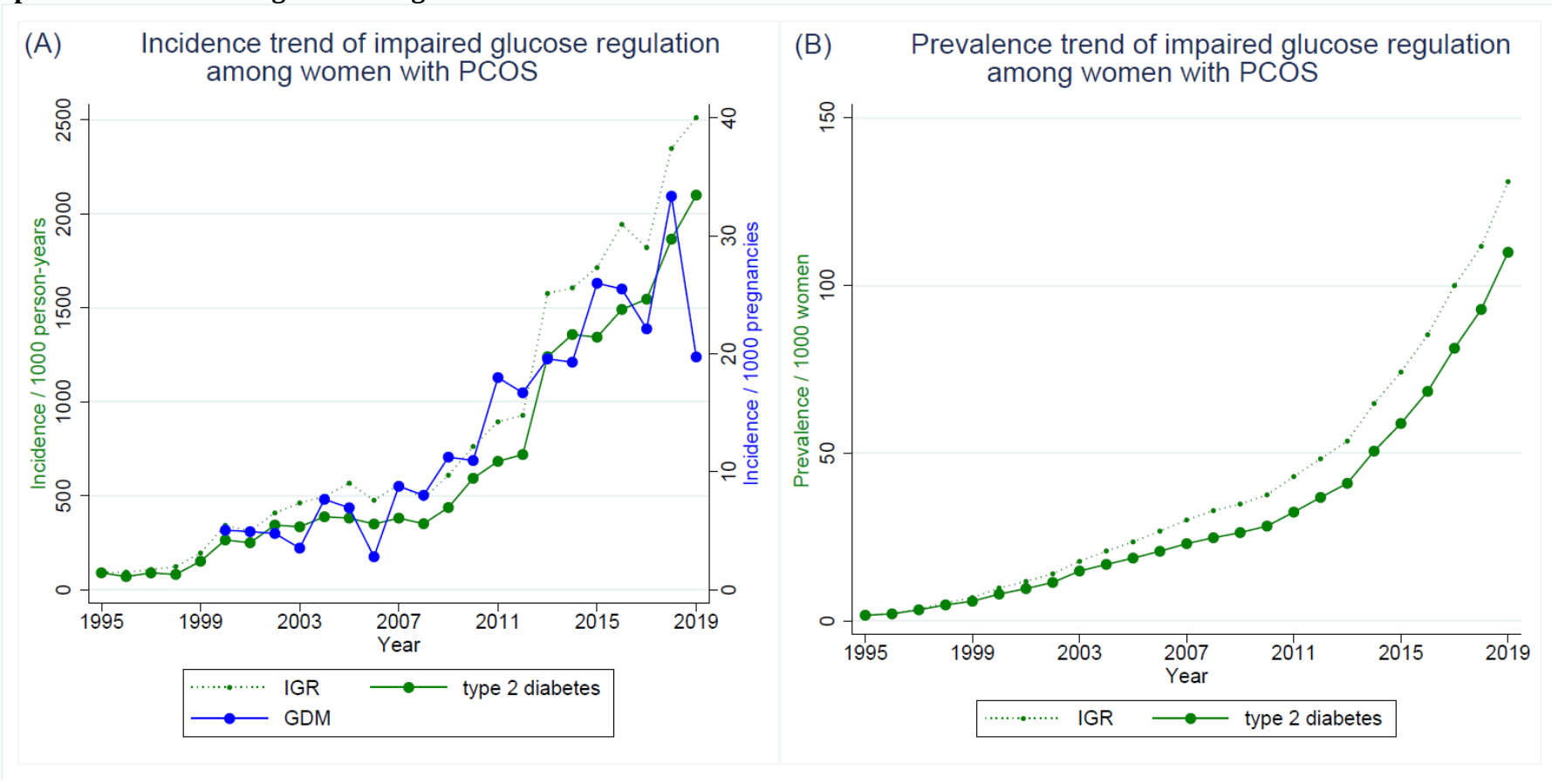
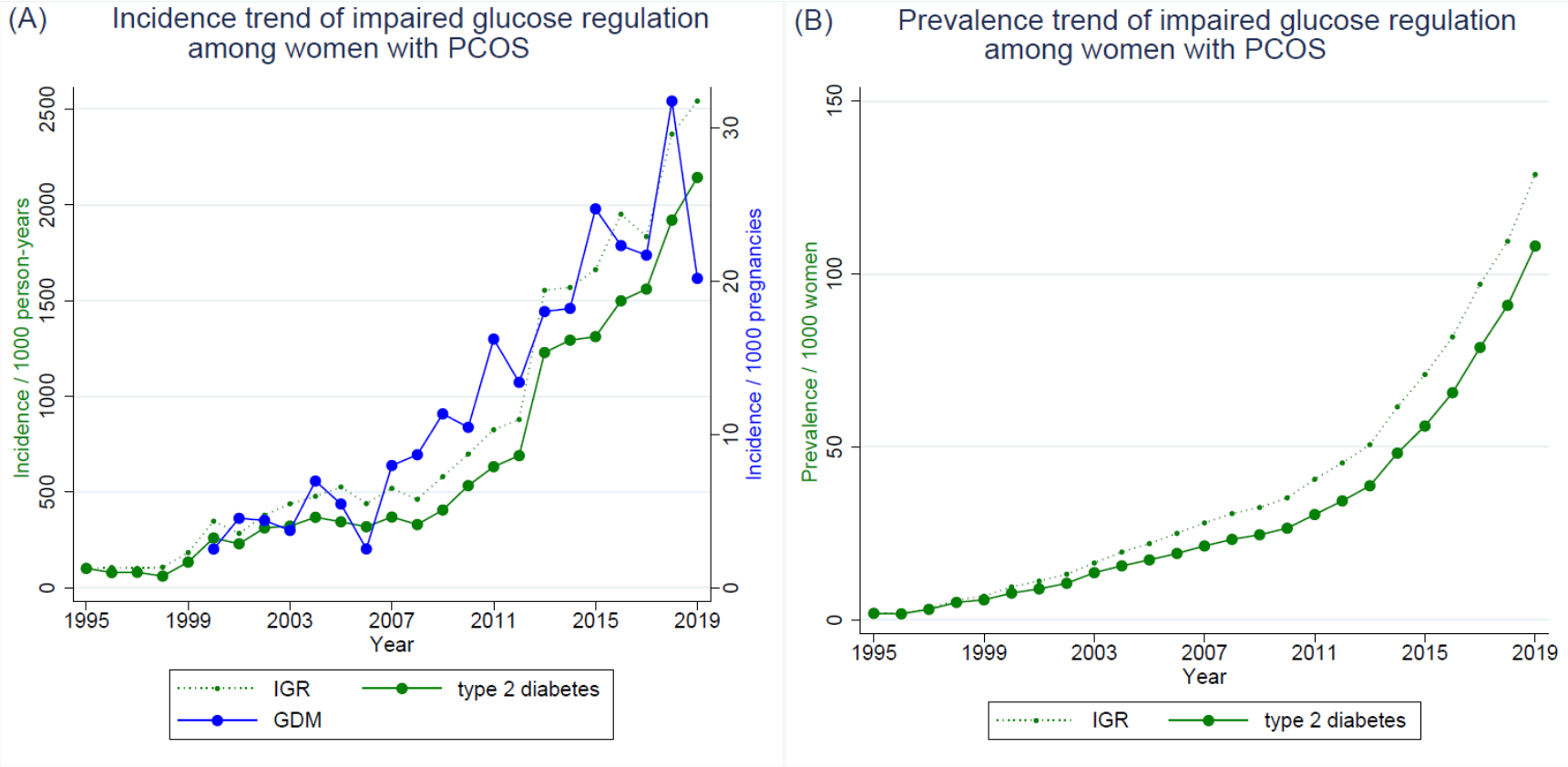


Figure 3.4: (A) Incidence (rate) trend of impaired glucose regulation and type 2 diabetes per 1,000 person-years and incidence trend of gestational diabetes mellitus per 1,000 pregnancies of women with combination of symptoms codes indicating PCOS diagnosis based on AES criteria, and (B) prevalence trend of impaired glucose regulation and type 2 diabetes per 1,000 women with combination of symptoms codes indicating PCOS diagnosis based on AES criteria



3.4.3 Prevalence trend of IGR and type 2 diabetes

Over the last two decades, prevalence of IGR and type 2 diabetes has been on a gradual rise from 28.8 to 154.6 and 28.8 to 128.5 (between 1999 and 2019) per 1,000 women with a diagnosis of PCOS, respectively (**Supplementary 28, Supplementary 29** and **Figure 3.1)**. A similar increasing trend in the prevalence of IGR and type 2 diabetes was observed for all of the other three cohorts of women with PCOS identified using a combination of symptom codes representing diagnosis based on NICHD criteria (IGR: 6.6 to 125.1 and type 2 diabetes: 5.4 to 104.9 per 1,000 women), Rotterdam criteria (IGR:7.0 to 130.9 and type 2 diabetes: 5.9 to 109.9 per 1,000 women), and AES criteria (IGR: 6.9 to 128.9 and type 2 diabetes: 5.8 to 108.1 per 1,000 women) (**Supplementary 28, Supplementary 29, Figure 3.2, Figure 3.3** and **Figure 3.4**).

3.4.4 Incidence trend of GDM

Among pregnancies of women with a diagnostic code of PCOS, combination of symptom codes representing PCOS diagnosis based on NICHD, Rotterdam and AES criteria, the overall incidence of GDM observed between 2000 and 2019 was 27.9, 11.7, 13.9 and 12.9 per 1,000 pregnancies respectively (**Supplementary 30, Figure 3.1, Figure 3.2, Figure 3.3** and **Figure 3.4)**. Between 2000 and 2019, the incidence of GDM among the four base cohorts of women with PCOS increased from 0 to 25.7, 2.7 to 24.1, 5.1 to 19.7 and 2.5 to 20.8 per 1,000 pregnancies, respectively.

3.5 Discussion

3.5.1 Summary of findings

In this primary care-based cohort study of women with PCOS identified through diagnostic codes and a combination of symptom codes, an increase in both the incidence and prevalence of IGR and type 2 diabetes was found over the last two and half decades. Furthermore, an increase in the incidence of GDM was also found among pregnancies of women with PCOS

over the last two decades. Women with a diagnostic code for PCOS had a higher incidence and prevalence estimates of IGR, type 2 diabetes and GDM overall, followed by women with a combination of symptom codes indicating PCOS diagnostic code based on Rotterdam criteria, AES criteria, and NICHD criteria, suggesting diagnostic coding of patients with more severe phenotype by primary care physicians. A more prominent increase in the incidence of both IGR and type 2 diabetes was observed between the years 2012-2013 and 2017-2018.

The sharp increase between 2012 and 2013 coincides with the publication of the public health guideline (PH38) in 2012 on identification of patients at high risk of type 2 diabetes or prediabetes.[103] The guideline recommends active screening of high-risk patients, including those with PCOS, for type 2 diabetes within primary and secondary care.

The sharp increase between 2017 and 2018 coincides with the launch of diabetes prevention week between 16th and 22nd of April 2018,[104] during which campaigns were run nationwide to raise awareness of type 2 diabetes, its' at-risk groups and associated complications. The campaign organized by NHS Diabetes Prevention Programme with events primarily run locally by general practices raised awareness among the public through posters, leaflets, and promotional videos wherein the public was urged to take up the free NHS Health Check to assess their risk of type 2 diabetes.

3.5.2 Comparison with existing literature

The primary care-based Danish cohort study by Pal et. al. suggests an increasing incidence of type 2 diabetes between 2009 and 2013 (4.98 to 5.06 per 1,000 person-years), and a decreasing trend until 2018 (3.56 per 1,000 person-years) in the general population of patients registered with a general practice[105]. Similarly, another study in the UK, using data from CPRD shows an increasing trend in the incidence of type 2 diabetes between 1991 and 2002 (169 to 448 per 100,000 population), a dip in the incidence until 2006 (376 per 100,000 population), followed by another rise in incidence until 2010 (515 per 100,000 population).[106] In another study by

McManus et. al., evaluating the impact of the NHS Diabetes Prevention Programme in the general population,[107] the authors have suggested that the incidence of type 2 diabetes during the years 2018 and 2019 are significantly lower than the counterfactual estimates in the absence of the Programme. Despite the stable/decreasing trend in the incidence of type 2 diabetes among the Danish and UK general population post-2013 and post-2002 respectively, a steady increase in type 2 diabetes incidence among the sub-cohort of women with PCOS was observed. This might indicate either a true increase in incidence or an increase in screening and capturing of patients with type 2 diabetes among the high-risk of cohort of women with PCOS. In a study by Dabelea et al., using data form Kaiser Permanente of Colorado GDM Screening Program, the prevalence of GDM has also been observed to be on a rise among all pregnant women.[108] The study reported highest incidence among the Asian population within the study period between 1994 and 2002 (6.3 to 8.6%), followed by African American (2.5 to 4.6%), Hispanic (2.8 to 3.4%) and non-Hispanic white (1.9 to 3.4%) population, following similar ethnic distribution for PCOS prevalence reported in Chapter 2.

3.5.3 Recommendations from guidelines

The Royal College of Obstetricians and Gynecologists' guideline on long-term consequences of PCOS recommends screening for impaired glucose regulation among women with PCOS, especially if they are overweight/obese or have other risk-factors such as age above 40 years, personal history of gestational diabetes or have a family history of type 2 diabetes.[94] It recommends oral glucose tolerance test when it is feasible, and HbA1c test in the absence of resources or patients' unwillingness. Furthermore, the guideline recommends annual screening for type 2 diabetes if a woman is recorded as having impaired fasting glucose or impaired glucose regulation.

Finally, NICE recommends risk assessment for the development of GDM during antenatal appointment. The following risk factors are considered, including BMI above 30 kg/m^2,

previous history of delivery of a baby weighing more than 4.5kg, previous history of GDM, family history of diabetes and belonging to an ethnic subgroup predisposing the women to a high prevalence of diabetes such as South Asian or Black Afro-Caribbean ethnicity.[109]

3.5.4 Strengths and limitations

The major strength of this study is the use of real-world data from a representative sample of patients with diagnosed PCOS within primary care, and large sample size. This study is also the first to report on the incidence and prevalence trends of impaired glucose regulation among women with PCOS. However, the real-world setting from which routinely collected data was extracted for this study also has caveats such as under-recording of both the exposure (PCOS) and the outcome (IGR), and changes in reporting and documentation of these conditions over time. Therefore, it is not possible to confirm true incidence trends from the trends in diagnosed incidence and prevalence as captured in this study. Finally, sample sizes during the early years of the study period were considerably smaller due to limited use of electronic health records in this period.

3.5.5 Implications

With the increasing incidence of impaired glucose regulation among women with PCOS, awareness of the effectiveness of existing interventions such as combined oral contraceptives and metformin for prevention of dysglycaemia should be raised. Resources for research should be allocated to create novel effective pharmacological/non-pharmacological interventions for treatment of PCOS.

3.5.6 Conclusion

The incidence of type 2 diabetes, IGR and GDM has been on a rise in the high-risk cohort of women with PCOS. The increasing incidence may reflect a true increase in incidence or improved screening and better capture of impaired glucose regulation among women with PCOS.

Chapter 4 - Polycystic ovary syndrome and risk of adverse obstetric outcomes: a retrospective population-based matched cohort study in England

Published article

BMC Medicine 298 (2022) < https://doi.org/10.1186/s12916-022-02473-3 >

Anuradhaa Subramanian[1], Siang Ing Lee[1], Katherine Phillips[1], Konstantinos A. Toulis[1], Punith Kempegowda[2], Michael W. O'Reilly[3], Nicola J Adderley[1], Shakila Thangaratinam[2,4], Wiebke Arlt[2,5], Krishnarajah Nirantharakumar[1,6]

[1]Institute of Applied Health Research, University of Birmingham, Birmingham, UK

[2]WHO Collaborating Centre for Global Women's Health, Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK

[3]Endocrinology and Metabolism Unit, Department of Medicine, Royal College of Surgeons in Ireland (RCSI), University of Medicine and Health Sciences, Dublin, Republic of Ireland

[4]Birmingham Women's and Children's NHS Foundation Trust, Birmingham, UK

[5]National Institute for Health Research (NIHR) Birmingham Biomedical Research Centre, University Hospitals Birmingham NHS Foundation Trust and University of Birmingham, Birmingham, UK

[6]Midlands Health Data Research UK, Birmingham, UK

A.S conceived the idea for the study with guidance from S.T, W.A and K.N. A.S extracted and merged data from multiple sources to generate the analysis-ready dataset. A.S designed the study and performed data cleaning and analysis. A.S wrote the initial draft of the manuscript with support from N.J.A, and K.A.T. All authors reviewed and revised the manuscript. All authors read and approved the final manuscript.

4.1 Abstract

4.1.1 Background

Polycystic ovary syndrome (PCOS) affects up to one in five women of childbearing age. Observational studies assessing the association between maternal PCOS, and adverse obstetric outcomes have reported varying results, depending on patient population, diagnostic criteria for PCOS and covariates accounted for in their analyses. I aimed to assess the risk of obstetric outcomes among a population-based representative cohort of women with PCOS compared to an age matched cohort of women without PCOS.

4.1.2 Methods

A retrospective cohort study was conducted of pregnancies of women in England aged 15-49 years identified from the Clinical Practice Research Datalink (CPRD) GOLD pregnancy register and linked Hospital Episodes Statistic (HES) data between March 1997 and March 2020. Pregnancies from the register that had a linked HES delivery record were included. Linked CPRD primary care data was used to ascertain maternal PCOS exposure prior to pregnancy. To improve detection of PCOS, in addition to PCOS diagnostic codes, codes for (1) polycystic ovaries or (2) hyperandrogenism and anovulation together were also considered. Sensitivity analysis was limited to only pregnant women with a diagnostic code for PCOS.

Primary outcomes ascertained from linked HES data were: (1) preterm delivery (gestation <37 weeks); (2) mode of delivery; (3) high (>4000g) or low birthweight (<2500g); and (4) stillbirth. Secondary outcomes were: (1) very preterm delivery (<32 weeks), (2) extremely preterm delivery (<28 weeks), (3) small and (4) large for gestational age.

Conditional logistic regression models were performed adjusting for age, ethnicity, deprivation, dysglycaemia, hypertension, thyroid disorders, number of babies born at index pregnancy, and pre-gravid BMI. Multiple imputation was performed for missing outcome data.

4.1.3 Results

27,586 deliveries with maternal PCOS were matched for age (±1 year) to 110,344 deliveries without maternal PCOS. In the fully adjusted models, maternal PCOS was associated with an increased risk of (1) preterm birth [aOR: 1.11 (95% CI, 1.06-1.17)], and (2) emergency caesarean, elective caesarean and instrumental vaginal compared to spontaneous delivery [aOR: 1.10 (95% CI, 1.05-1.15), 1.07 (95% CI, 1.03-1.12) and 1.04 (95% CI, 1.00-1.09), respectively]. There was absence of association with low birthweight, high birthweight, and stillbirth. In the sensitivity analysis, the association with preterm birth [aOR: 1.31 (95% CI, 1.13-1.52)], emergency caesarean [aOR: 1.15 (95% CI, 1.02-1.30)], and elective caesarean [aOR: 1.03 (95% CI, 1.02-1.03)] remained.

While there was no significant association with any of the secondary outcomes in the primary analysis, in the sensitivity analysis maternal PCOS was associated with increased risk of extremely preterm delivery [aOR: 1.86 (95% CI, 1.31-2.65)], and lower risk of small for gestational age babies [aOR: 0.74 (95% CI, 0.59-0.94)].

4.1.4 Conclusions

Maternal PCOS was associated with increased risk of preterm and caesarean delivery. Association with low birthweight may be largely mediated by lower gestational age at birth.

4.2 Introduction

Polycystic ovary syndrome (PCOS) is a common yet underdiagnosed endocrine disorder,[1,9] with a diagnosed prevalence of 10%;[6] it is estimated that half of women with PCOS are undiagnosed.[8] Consensus criteria for diagnosis of PCOS require presence of two out of the following three features: (1) biochemical evidence or clinical manifestations of androgen excess such as hirsutism and hair loss, (2) chronic oligo-/anovulation and (3) polycystic ovarian morphology on ultrasound.[3] The adverse clinical phenotype is largely driven by a complex interplay between insulin resistance and androgen excess.[110] PCOS is considered a lifelong metabolic disorder[111] with a plethora of adverse risks during and following pregnancy,[112] and even posing intergenerational risks to the children of women with PCOS.[113] These risks may be attributed to the biochemical features of PCOS or several other co-existing risk factors such as high BMI, or comorbidities that are commonly seen among women with PCOS.[114]

Several systematic reviews have pooled together findings from observational studies examining the association between maternal PCOS and the risk of a range of obstetric outcomes. However, these reviews suggest varying results across the primary studies that they included owing to methodological heterogeneity,[55–57] which included differences in terms of source population, criteria employed for PCOS ascertainment, and confounders matched and adjusted for in their design and analysis respectively. Several of these primary studies are further limited in terms of outdated data, their sample size,[58,59] and restrictive selection of pregnant women who have undergone assisted reproduction[60,61] within their studies.

Furthermore, socio-demographic factors such as high BMI, deprivation and minority ethnic background, as well as metabolic disturbances such as insulin resistance, hypertension and thyroid disorders, may exacerbate the severity of PCOS.[9,30,36,37,115,116] The existing literature is limited in terms of comprehensively identifying, assessing and accounting for these confounders/mediators.

Therefore, in order to overcome the limitations of the observational studies in the existing literature, an age matched retrospective cohort study of pregnant women was performed using a population representative, UK primary care-based data source, to identify the risk of adverse obstetric outcomes including preterm birth, different mode of delivery, high and low birthweight, and stillbirth in women with PCOS compared to those without. Furthermore, confounders agreed a priori were adjusted for in a series of regression models, adding covariates step by step to identify the extent of confounding conferred by each risk factor.

4.3 Methods

4.3.1 Study design and data source

A retrospective open cohort study of pregnant women identified from primary care records [Clinical Practice Research Datalink (CPRD) GOLD Pregnancy Register], with their delivery recorded in secondary care [linked Hospital Episode Statistics (HES)] between 1997 and 2020, was performed to determine the incidence of adverse obstetric outcomes among women with PCOS in comparison to women without PCOS.

CPRD GOLD contains representative data from 7% of the general practices across the UK, covering 20 million patients from 973 practices. It contains pseudo-anonymized patient-level data on demographics, symptoms, diagnoses, drug prescriptions, physical measurements, and laboratory test results. Furthermore, patient-level data can be linked to other data sources such as HES data and deprivation data, via a trusted third party.[85] The linkage of databases aided capture of information on exposure (PCOS) from primary care, the obstetric outcomes from HES maternity tail and important potential confounders from both primary and secondary care. Symptoms and diagnoses are recorded within CPRD GOLD using Read codes, a hierarchical clinical coding system. Using maternity, antenatal and delivery health records within CPRD GOLD, pregnancy episodes and their outcomes are identified through a validated algorithm,[100]

which formulated the CPRD GOLD Pregnancy Register and formed the source cohort for this study.

4.3.2 Study population

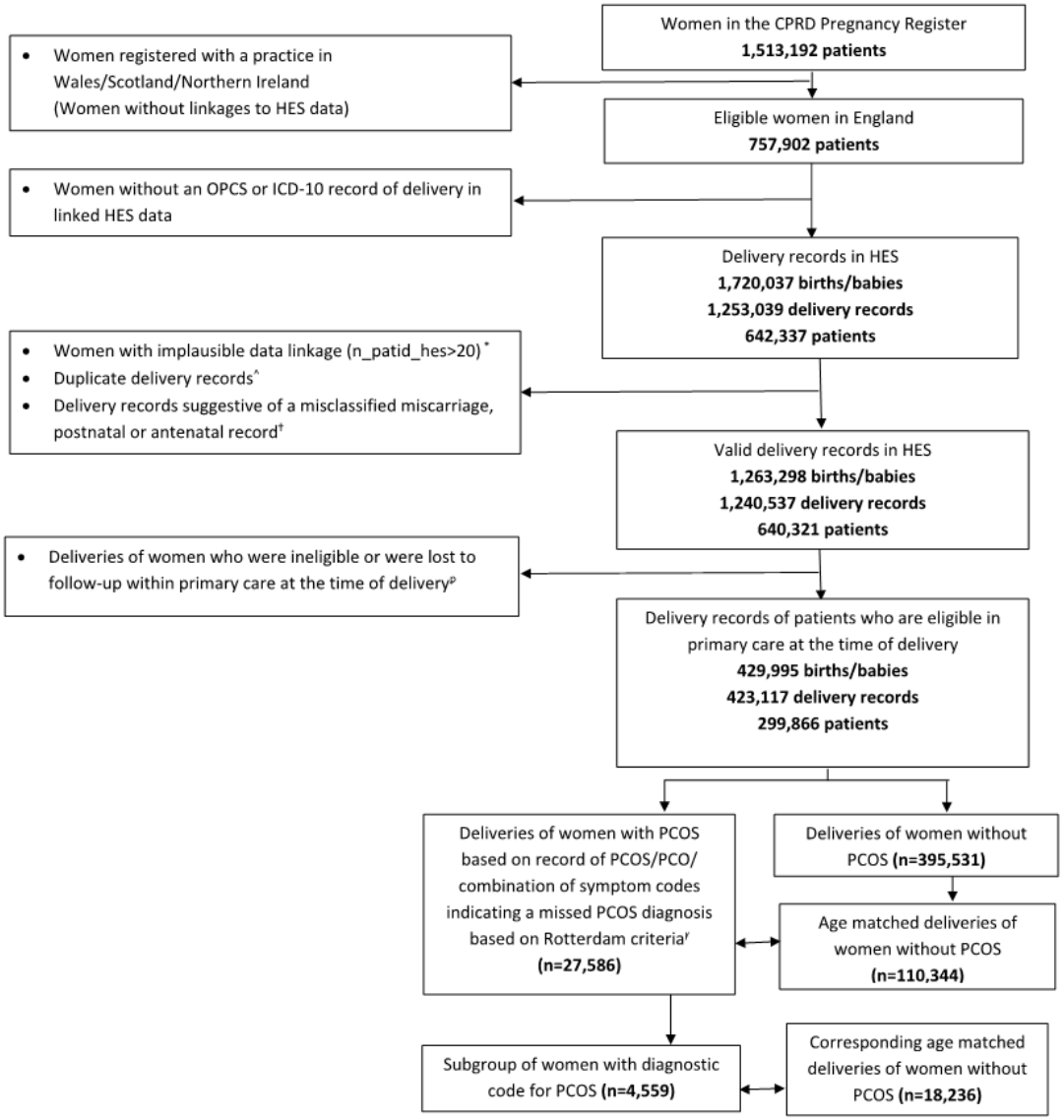
Pregnant women were included from the CPRD GOLD Pregnancy Register if they were registered at a general practice in England and had a record of delivery from linked HES data [containing information on admissions to National Health Service (NHS) hospitals in England]. Deliveries formed the unit of analysis in this study and an index date was assigned to each eligible delivery record. Women with implausible data linkage (where a patient record in HES is linked to more than 20 patient records across 20 different primary care practices) were excluded. Furthermore, delivery records were excluded if they were: (1) duplicates, or (2) misclassified miscarriage, postnatal or antenatal record. Delivery records were considered misclassified miscarriages if the reported gestational age was less than 23 weeks. If two deliveries were recorded within 180 days of each other for the same patient, one of the delivery records was considered as a misclassified antenatal or postnatal record. Finally, delivery records were excluded if women were ineligible or were lost to follow-up within primary care at the time of delivery. Women were considered ineligible within primary care if they (1) did not have an acceptable patient flag within CPRD GOLD (indicating sufficient data quality), (2) did not have a minimum registration period of one year with an eligible general practice on delivery date (practices were considered eligible one year after the "up-to-standard" date, a flag for sufficient practice data quality), or (3) were aged <15- or >49 years on delivery date.

Once linked, the mother's PCOS exposure status for each delivery record was ascertained from primary care prior to the index date (date of delivery). PCOS was defined as a Read code record of PCOS. Due to underdiagnosis of PCOS within primary care, records of polycystic ovaries (PCOs),[26,30] or a combination of symptom codes indicating a missed PCOS diagnoses based on Rotterdam criteria [(1) anovulation and (2) biochemical or symptomatic presentation of

hyperandrogenism; a Read code record of hair loss or hirsutism and a recorded measure of serum testosterone level ≥2.0 nmol/L were considered as symptomatic and biochemical presentation of hyperandrogenism, respectively] were also considered.

For each delivery record of women with PCOS (in a random order), four control delivery records of women without PCOS were selected from a pool of age-matched (±1 year) pregnant women without replacement. Cohort selection for this study is described in **Figure 4.1**.

Figure 4.1: Flow chart describing cohort selection



* Number of primary care patient records linked to the same HES patient record is large (n_patid_hes>20). This linkage may not be reliable and therefore these patients are excluded

^ (1) In case of more than 9 births during the same delivery with missing birthweight data, only the first birth is included and the rest are considered duplicates; (2) In case of multiple births, if all babies have the same birthweight recorded, then only one of the babies is included and the rest are considered duplicates; (3) If the number of births reported within a delivery does not match with the number of birth records within a delivery, excess birth records are considered duplicates. Duplicates are excluded

† Delivery records are considered as misclassified miscarriages if the reported gestational age is less than 23 weeks; Delivery records are considered as misclassified antenatal or postnatal records if two deliveries are recorded within 180 days of each other for the same patient, and the record with missing birthweight is considered misclassified.

P (1) Patients without an acceptable patient flag within CPRD GOLD (indicating sufficient data quality); (2) Patients without a minimum registration period of one year with an eligible general practice on delivery date (Practices were considered eligible one year after the "up-tostandard" date, a flag for sufficient practice data quality); (3) Patients aged <15- or >49 on delivery date; (4) Patients transferred out of practice, or their registered practice stopped contributing data to CPRD GOLD on their date of delivery.

r Rotterdam criteria: (1) anovulation and (2) biochemical or symptomatic presentation of hyperandrogenism; Read code record of hair loss or hirsutism and a recorded measure of serum testosterone level ≥ 2.0 nmol/L was considered as symptomatic and biochemical presentation of hyperandrogenism respectively.

4.3.3 Outcomes

Four primary outcomes were considered and identified from HES data: (1) preterm birth, (2) mode of delivery, (3) high or low birthweight, (4) stillbirth.

Gestational age recorded within the HES maternity tail at the time of delivery and relevant ICD-10 codes were used to identify the outcome preterm birth (gestational age at birth <37 weeks). Based on Operating Procedure Codes Supplement (OPCS) codes and ICD-10 codes, mode of delivery was classified into one of the following four categories as a categorical outcome variable: (1) emergency caesarean section, (2) elective or other unspecified caesarean section, (3) instrumental vaginal delivery, (4) spontaneous or other unspecified vaginal delivery (reference category). Based on birthweight(s) recorded in the maternity tail, delivery was classified as high or low birthweight delivery if at least one of the babies born in that delivery was above 4000 grams or below 2500 grams, respectively. In addition, a record of the relevant ICD-10 code was used to identify a high birthweight baby. Stillbirth outcomes were identified using relevant ICD-10 codes and from maternity tail records.

As secondary outcomes, gestational age was classified to identify very preterm (<32 weeks) and extremely preterm (<28 weeks) delivery. Small and large for gestational age babies (birthweight <10^{th} and >90^{th} centile, respectively) were identified using the INTERGROWTH 21^{st} project,[117] and their software tools, by comparing the birthweight and gestational age recorded in HES data to the international anthropometric standards.

4.3.4 Explanatory variables

Risk factors or features of PCOS that are also obstetric risk factors were considered as possible explanatory variables and adjusted for them in this analysis in a step-by-step manner. This included age, ethnicity, deprivation, impaired glucose regulation based on a diagnosis of type 2 diabetes or prediabetes, diagnosis of hypertension, thyroid disorders, number of babies born within the delivery, and pre-gravid body mass index (BMI). For the outcomes low and high

birthweight and mode of delivery, gestational age was further considered as an explanatory variable.

Ethnicity was identified using relevant Read codes from primary care records and was categorized as (1) white Caucasian, (2) South Asian, (3) black Afro-Caribbean and (4) mixed or multiple ethnic group or (5) other ethnic minority group. Primary care linked English index of multiple deprivation (IMD) data provided a relative measure of deprivation based on seven different domains.[118] Type 2 diabetes was identified from primary care through relevant Read Codes, record of HbA1c ≥48 mmol/L (≥6.5%) or fasting blood glucose >7 mmol/L. Impaired glucose regulation was identified through relevant Read codes, HbA1c ≥42 mmol/L (≥6.0%) or fasting blood glucose ≥5.5 mmol/L. Diagnoses of hypertension and thyroid disorders were identified from primary care through Read code records. The number of babies born during that delivery was derived from linked HES maternity tail records. Pre-gravid BMI was identified as the latest BMI measured in primary care at least a year before index date and was categorized according to WHO standards as under/normal weight (<25 kg/m^2), overweight (25-30 kg/m^2) and obese (≥30 kg/m^2). A separate missing category was created for those with missing data on ethnicity, deprivation, number of babies born within the delivery and pre-gravid BMI.

4.3.5 Statistical Analysis

Deliveries were the unit of this analysis. Baseline explanatory variables were described using appropriate summary statistics stratified by exposure to maternal PCOS. Mean with standard deviation (SD) and median with interquartile range (IQR) were provided for continuous variables as appropriate. Frequency and percentage were provided for categorical variables.

Multiple imputation using chained equation was performed to impute missing delivery related data that were essential to compute outcome variables.[119–121] Missing values were imputed 31 times (since gestational age was missing among 31% of the women in the study) using linear

(for gestational age and birthweight outcomes), logistic (for stillbirth outcome and sex of the baby) and multinomial logistic (for the categorical delivery method outcome) regression as appropriate using the independent variables age, BMI, impaired glucose regulation, deprivation and the number of babies delivered. Conditional logistic or multinomial logistic regression models were used to provide unadjusted and adjusted odds ratios (ORs) for the binary and nominal categorical outcome variables (mode of delivery), respectively, among women with PCOS compared to women without PCOS. Robust confidence intervals were estimated after accounting for the intragroup correlation of multiple deliveries of a woman throughout her reproductive age. Explanatory variables were included in a step-by-step manner in the regression model, resulting in a fully adjusted model.

A sensitivity analysis was performed restricting to women with a coded diagnosis of PCOS only and their corresponding matched controls. All analyses were performed in Stata IC version 15. Two-sided P values were obtained for all tests, and a P value <0.05 was considered as statistically significant. Selection of Read, ICD-10 and OPCS code lists was performed using an inhouse developed software platform called Code Builder, with systematic searching of existing code lists, and through clinical knowledge and discussion methods used in previous publications.[122] The list of codes used for exposure ascertainment are provided in

Supplementary *2*, Supplementary 3, Supplementary 4, Supplementary 5, Supplementary 6 and **Supplementary 7.** The list of codes used for outcome ascertainment are provided in **Supplementary 31, Supplementary 32, Supplementary 33, Supplementary 34** and **Supplementary 35**. The study results are reported as per the RECORD (REporting of studies Conducted using Observational Routinely-collected health Data) statement.

4.4 Results

4.4.1 Cohort selection

Out of the 1,513,192 women identified within the CPRD GOLD Pregnancy Register, 757,902 women were eligible for linkage to HES. Of these women, 642,337 had a record of delivery (n=1,253,039) within HES linked data based on OPCS and ICD-10 records. After excluding patients and delivery records as outlined in the methods section above (also describe din **Figure 4.1**), a final eligible cohort of 423,117 delivery records from 299,866 patients were identified. From the eligible cohort of delivery records, 27,586 (6.5%) were for women with a coded diagnosis of PCOS/PCO or a combination of symptom codes indicating a missed PCOS diagnosis based on Rotterdam criteria; these deliveries formed the exposed cohort for the primary analysis. From a pool of 395,531 control delivery records, an unexposed cohort of 110,344 was selected, matched for maternal age. In the sensitivity analysis, 4,559 (1.1%) deliveries by women who had a specifically coded diagnosis for PCOS, and their corresponding matched controls (18,236 deliveries) were included.

4.4.2 Baseline characteristics

The mean (SD) age at delivery of women with and without PCOS was 30.86 (5.38) and 30.85 (5.33), respectively (**Table 4.1**).

Table 4.1: Baseline characteristics of deliveries of women with PCOS and age matched controls

	Primary analysis	
Variables	**Deliveries of women with PCOS***	**Age matched deliveries of women without PCOS**
All	**(n=27586)**	**(n=110344)**
Age at delivery [Mean (SD)]	30.86 (5.38)	30.85 (5.33)
Age at delivery [Median (IQR)]	30.00 (26.00-34.00)	31.00 (27.00-34.00)
Age categories, n (%)		
14 - 19 years	467 (1.69)	1802 (1.63)
20 - 29 years	11537 (41.82)	45596 (41.32)
30 - 39 years	14357 (52.04)	58313 (52.85)
40 - 50 years	1225 (4.44)	4633 (4.20)
Pre-gravid BMI [Mean (SD)]	26.54 (6.38)	25.11 (5.43)
Pre-gravid BMI [Median (IQR)]	24.00 (21.00-30.00)	23.00 (21.00-27.00)
BMI Categories, n (%)		
<25 kg/m2	13055 (47.32)	59799 (54.19)
25-29.9 kg/m2	6493 (23.54)	25103 (22.75)
30-34.9 kg/m2	3667 (13.29)	10423 (9.45)
35-39.9 kg/m2	1882 (6.82)	4023 (3.65)
≥40 kg/m2	1013 (3.67)	2002 (1.81)
Missing	1476 (5.35)	8994 (8.15)
IMD, n (%)		
1 (Most deprived)	3334 (12.09)	12989 (11.77)
2	2795 (10.13)	11052 (10.02)
3	2706 (9.81)	11215 (10.16)
4	2637 (9.56)	11031 (10.00)
5	2973 (10.78)	11782 (10.68)
6	2578 (9.35)	10165 (9.21)
7	2547 (9.23)	10557 (9.57)
8	2696 (9.77)	10319 (9.35)
9	2693 (9.76)	10353 (9.38)
10 (Least deprived)	2607 (9.45)	10793 (9.78)
Missing	19 (0.07)	81 (0.07)
Ethnicity, n (%)		
White	13343 (48.37)	50894 (46.12)
South Asian	1465 (5.31)	3638 (3.30)
Black Afro-Caribbean	1567 (5.68)	5315 (4.82)
Mixed or multiple ethnicity	170 (0.62)	651 (0.59)
Others	721 (2.61)	2561 (2.32)
Missing	10320 (37.41)	47285 (42.85)
Record of symptoms and measurements at baseline, n (%)		
PCO	12706 (46.06)	0 (0)
Hair Loss	2898 (10.51)	2645 (2.40)

Hirsutism	1825 (6.62)	645 (0.58)
Anovulation	17852 (64.71)	10845 (9.83)
High Testosterone (serum testosterone level ≥2.0 nmol/L)	3250 (11.78)	467 (0.42)
Other comorbidities, n (%)		
Type 2 diabetes	675 (2.45)	1259 (1.14)
Prediabetes	1123 (4.07)	2038 (1.85)
Hypertension	489 (1.77)	1230 (1.11)
Thyroid disorders	1105 (4.01)	2403 (2.18)
Pregnancy related variables	1123 (4.07)	2038 (1.85)
Number of babies at the delivery, n (%)		
1	27163 (98.47)	108446 (98.28)
2	409 (1.48)	1864 (1.69)
3	14 (0.05)	30 (0.03)
4	0 (0)	4 (0.00)

*Record of PCOS/ PCO/ conglomeration of symptom codes indicating a missed PCOS diagnosis based on Rotterdam criteria [two of the three symptoms recorded: (1) PCO, (2) anovulation, and (3) biochemical or symptomatic presentation of hyperandrogenism]. Read code record of hair loss or hirsutism and a recorded measure of serum testosterone level ≥2.0 nmol/L was considered as symptomatic and biochemical presentation of hyperandrogenism respectively.
PCOS: Polycystic Ovary Syndrome; PCO: Polycystic ovaries; SD: Standard Deviation; IQR: Interquartile Range; BMI: Body Mass Index; IMD: Index of Multiple Deprivation

Compared to women without PCOS, women with PCOS had higher pre-gravid BMI [mean (SD): 26.54 (6.38) vs 25.11 (5.43)], were more likely to be deprived [IMD most deprived decile (1): 12.09% vs 11.77%] and were more likely to be from an ethnic minority background [South Asian (5.31% vs 3.30%) and black Afro-Caribbean (5.68% vs 4.82%)]. As expected, women with PCOS were more likely to have a record of PCOS-related symptoms such as hair loss (10.51% vs 2.40%), hirsutism (6.62% vs 0.58%), anovulation (64.71% vs 9.83%), and serum testosterone ≥2.0 nmol/L (11.78% vs 0.42%). Women with PCOS were also more likely to have metabolic disturbances including comorbidities such as type 2 diabetes (2.45% vs 1.14%), prediabetes (4.07% vs 1.85%), hypertension (1.77% vs 1.11%), and thyroid disorders (4.01% vs 2.18%) (**Table 4.1**). The baseline characteristics of deliveries of women with a diagnostic code for PCOS and their maternal age-matched deliveries of women without PCOS is presented in **Supplementary 36**.

Table 4.2: Risk of primary obstetric outcomes among women with PCOS compared to women without PCOS

Outcomes	Deliveries of women with PCOS* (n=20,7586)	Age matched deliveries of women without PCOS (n=110,344)
Preterm (<37 weeks of gestational age at delivery)		
Number of patients	27586	110344
Outcome events, n (%)	2104 (7.63%)	7520 (6.82%)
Unadjusted OR (95% CI)	1.13 (1.07-1.19)	
Adjusted OR (95% CI) (Model 1)	1.12 (1.07-1.17)	
Adjusted OR (95% CI) (Model 2)	1.09 (1.03-1.14)	
Adjusted OR (95% CI) (Model 3)	1.11 (1.05-1.17)	
Adjusted OR (95% CI) (Model 4)	1.11 (1.06-1.17)	
Model of delivery		
Number of patients	27586	110344
Outcome events, n (%)		
Emergency CS	3473 (12.59%)	12073 (10.94%)
Elective/Other/Unspecified CS	4211 (15.26%)	15279 (13.85%)
Instrumental Vaginal	3077 (11.15%)	12573 (11.39%)
Spontaneous/Other/Unspecified Vaginal	16825 (60.99%)	70419 (63.82%)
Unadjusted OR (95% CI)		
Emergency CS	1.20 (1.15-1.26)	
Elective/Other/Unspecified CS	1.15 (1.11-1.20)	
Instrumental Vaginal	1.02 (0.98-1.07)	
Spontaneous/Other/Unspecified Vaginal	Ref	
Adjusted OR (95% CI) (Model 1)		
Emergency CS	1.20 (1.15-1.26)	
Elective/Other/Unspecified CS	1.15 (1.10-1.19)	
Instrumental Vaginal	1.02 (0.98-1.07)	
Spontaneous/Other/Unspecified Vaginal	Ref	
Adjusted OR (95% CI) (Model 2)		
Emergency CS	1.17 (1.12-1.23)	
Elective/Other/Unspecified CS	1.13 (1.08-1.18)	
Instrumental Vaginal	1.02 (0.98-1.07)	
Spontaneous/Other/Unspecified Vaginal	Ref	
Adjusted OR (95% CI) (Model 3)		
Emergency CS	1.18 (1.12-1.23)	
Elective/Other/Unspecified CS	1.13 (1.09-1.18)	
Instrumental Vaginal	1.02 (0.98-1.07)	
Spontaneous/Other/Unspecified Vaginal	Ref	
Adjusted OR (95% CI) (Model 4)		
Emergency CS	1.11 (1.06-1.16)	
Elective/Other/Unspecified CS	1.08 (1.04-1.13)	
Instrumental Vaginal	1.04 (0.99-1.09)	
Spontaneous/Other/Unspecified Vaginal	Ref	

Adjusted OR (95% CI) (Model 5)		
Emergency CS	1.10 (1.05-1.15)	
Elective/Other/Unspecified CS	1.07 (1.03-1.12)	
Instrumental Vaginal	1.04 (1.00-1.09)	
Spontaneous/Other/Unspecified Vaginal	Ref	
High birthweight >4 kg (for at least one of the babies)		
Number of patients	27586	110344
Outcome events, n (%)	2709 (9.82%)	10632 (9.64%)
Unadjusted OR (95% CI)	1.02 (0.97-1.07)	
Adjusted OR (95% CI) (Model 1)	1.03 (0.99-1.08)	
Adjusted OR (95% CI) (Model 2)	1.03 (0.98-1.07)	
Adjusted OR (95% CI) (Model 3)	1.02 (0.98-1.07)	
Adjusted OR (95% CI) (Model 4)	0.96 (0.92-1.00)	
Adjusted OR (95% CI) (Model 5)	0.97 (0.92-1.01)	
Low birthweight <2.5 kg (for at least one of the babies)		
Number of patients	27586	110344
Outcome events, n (%)	1627 (5.90%)	5903 (5.35%)
Unadjusted OR (95% CI)	1.11 (1.05-1.18)	
Adjusted OR (95% CI) (Model 1)	1.10 (1.04-1.16)	
Adjusted OR (95% CI) (Model 2)	1.08 (1.02-1.14)	
Adjusted OR (95% CI) (Model 3)	1.10 (1.03-1.17)	
Adjusted OR (95% CI) (Model 4)	1.13 (1.06-1.20)	
Adjusted OR (95% CI) (Model 5)	1.03 (0.95-1.13)	
Stillbirth		
Number of patients	27586	110344
Outcome events, n (%)	122 (0.44%)	471 (0.43%)
Unadjusted OR (95% CI)	1.04 (0.85-1.26)	
Adjusted OR (95% CI) (Model 1)	1.03 (0.85-1.25)	
Adjusted OR (95% CI) (Model 2)	1.02 (0.85-1.24)	
Adjusted OR (95% CI) (Model 3)	1.01 (0.84-1.22)	
Adjusted OR (95% CI) (Model 4)	0.99 (0.81-1.21)	

*Record of PCOS/PCO/combination of symptom codes indicating a missed PCOS diagnosis based on Rotterdam criteria [(1) anovulation and (2) biochemical or symptomatic presentation of hyperandrogenism; Read code record of hair loss or hirsutism and a recorded measure of serum testosterone level ≥2.0 nmol/L was considered as symptomatic and biochemical presentation of hyperandrogenism respectively].

PCOS: Polycystic Ovary Syndrome; CS: Caesarean Section; OR: Odds Ratio

Model 1: Adjusted for age, ethnicity, and deprivation

Model 2: Adjusted for age, ethnicity, deprivation, baseline dysglycaemia, hypertension and thyroid disorders

Model 3: Adjusted for age, ethnicity, deprivation, baseline dysglycaemia, hypertension, thyroid disorders, and numbers of babies born at the delivery

Model 4: Adjusted for age, ethnicity, deprivation, baseline dysglycaemia, hypertension, thyroid disorders, numbers of babies born at the delivery, and pre-gravid body mass index

Model 5: Adjusted for age, ethnicity, deprivation, baseline dysglycaemia, hypertension, thyroid disorders, numbers of babies born at the delivery, pre-gravid body mass index, and gestational age

Table 4.3: Risk of secondary obstetric outcomes among women with PCOS compared to women without PCOS

Outcomes	Deliveries of women with PCOS*	Age matched deliveries of women without PCOS
Very Preterm (<32 weeks of gestational age at delivery)		
Number of patients	27586	110344
Outcome events, n (%)	619 (2.24%)	2244 (2.03%)
Unadjusted OR (95% CI)	1.11 (1.01-1.22)	
Adjusted OR (95% CI) (Model 1)	1.09 (0.99-1.19)	
Adjusted OR (95% CI) (Model 2)	1.07 (0.97-1.18)	
Adjusted OR (95% CI) (Model 3)	1.07 (0.97-1.18)	
Adjusted OR (95% CI) (Model 4)	1.07 (0.97-1.18)	
Extremely preterm (<28 weeks of gestational age at delivery)		
Number of patients	27586	110344
Outcome events, n (%)	272 (0.99%)	909 (0.82%)
Unadjusted OR (95% CI)	1.20 (1.04-1.39)	
Adjusted OR (95% CI) (Model 1)	1.16 (1.01-1.33)	
Adjusted OR (95% CI) (Model 2)	1.14 (0.99-1.31)	
Adjusted OR (95% CI) (Model 3)	1.13 (0.98-1.29)	
Adjusted OR (95% CI) (Model 4)	1.13 (0.98-1.29)	
Large for gestational age >90th percentile (for at least one of the babies)		
Number of patients	27586	110344
Outcome events, n (%)	4922 (17.84%)	18593 (16.85%)
Unadjusted OR (95% CI)	1.07 (1.03-1.11)	
Adjusted OR (95% CI) (Model 1)	1.08 (1.05-1.12)	
Adjusted OR (95% CI) (Model 2)	1.06 (1.03-1.10)	
Adjusted OR (95% CI) (Model 3)	1.06 (1.03-1.10)	
Adjusted OR (95% CI) (Model 4)	1.00 (0.97-1.04)	
Small for gestational age <10th percentile (for at least one of the babies)		
Number of patients	27586	110344
Outcome events, n (%)	1113 (4.03%)	4305 (3.90%)
Unadjusted OR (95% CI)	1.04 (0.97-1.11)	
Adjusted OR (95% CI) (Model 1)	1.01 (0.94-1.09)	
Adjusted OR (95% CI) (Model 2)	1.01 (0.94-1.09)	
Adjusted OR (95% CI) (Model 3)	1.00 (0.93-1.08)	
Adjusted OR (95% CI) (Model 4)	1.03 (0.96-1.11)	

*Record of PCOS/PCO/combination of symptom codes indicating a missed PCOS diagnosis based on Rotterdam criteria [(1) anovulation and (2) biochemical or symptomatic presentation of hyperandrogenism; Read code record of hair loss or hirsutism and a recorded measure of serum testosterone level ≥2.0 nmol/L was considered as symptomatic and biochemical presentation of hyperandrogenism respectively].

PCOS: Polycystic Ovary Syndrome; CS: Caesarean Section; OR: Odds Ratio
Model 1: Adjusted for age, ethnicity, and deprivation
Model 2: Adjusted for age, ethnicity, deprivation, baseline dysglycaemia, hypertension and thyroid disorders
Model 3: Adjusted for age, ethnicity, deprivation, baseline dysglycaemia, hypertension, thyroid disorders, and numbers of babies born at the delivery
Model 4: Adjusted for age, ethnicity, deprivation, baseline dysglycaemia, hypertension, thyroid disorders, numbers of babies born at the delivery, and pre-gravid body mass index

4.4.3 Risk of primary obstetric outcomes among women with PCOS compared to their age-matched controls

4.4.3.1 Preterm birth

Among the delivery records of women with and without a pre-existing diagnosis of PCOS, 7.63% (n=2,104) and 6.82% (n=7,520) of them were delivered preterm, resulting in 13% increased crude odds of preterm delivery among women with PCOS compared to women without PCOS [OR 1.13 (95% CI, 1.07-1.19)] (**Table 4.2**). There was marginal attenuation of the increased odds with adjustment for covariates [aOR: 1.11 (95% CI, 1.06-1.17)]. For the secondary outcomes of preterm delivery, among the delivery records of women with and without PCOS, 2.24% and 2.03% of deliveries were before 32 weeks of gestational age and 0.99% and 0.82% were before 28 weeks of gestational age, respectively **(Table 4.3)**. There were 11% and 20% increased crude odds of delivery before 32 and 28 weeks of gestational age, respectively [OR 1.11 (95% CI, 1.01-1.22) and 1.20 (95% CI, 1.04-1.39)], among women with PCOS compared to women without PCOS. There was marginal attenuation in the effect size at each step when serially adjusting for covariates, which resulted in increased odds of both outcomes among women with PCOS compared to women without PCOS, although statistically insignificant in the final model [aOR: 1.07 (95% CI, 0.97-1.18) and 1.13 (95% CI, 0.98-1.29) for delivery <32 and <28 weeks of gestational age, respectively]. In the sensitivity analysis including a sub-cohort of deliveries by women with a diagnostic code for PCOS and their corresponding maternal age-matched control deliveries, the odds ratios were more pronounced for delivery less than 37, 32 and 28 weeks of gestational age [gestational age <37

weeks aOR: 1.31 (95% CI, 1.13-1.52); gestational age <32 weeks aOR: 1.42 (95% CI, 0.88-2.31); gestational age <28 weeks aOR: 1.86 (95% CI, 1.31-2.65)]. **(Supplementary 37** and **Supplementary 38)**

4.4.3.2 Mode of delivery

Compared to deliveries of women without PCOS, delivery of women with PCOS were more likely to occur by caesarean section (emergency: 12.59% vs 10.94%, elective/other/unspecified: 15.26% vs 13.85%) and less likely to occur by vaginal delivery (instrumental: 11.15% vs 11.39%, spontaneous/other/unspecified: 60.99% vs 63.82%). When serially adjusting for covariates, marginal attenuation in the effect estimate was observed, with the highest drop observed when adjusting for pre-gravid BMI. In the fully adjusted model, compared to spontaneous/other/unspecified vaginal delivery, delivery of women with PCOS were 4% at higher odds of being an instrumental vaginal delivery [aOR: 1.04 (95% CI, 1.00-1.09)], 7% at higher odds of being elective/other/unspecified caesarean section [aOR: 1.07 (95% CI, 1.03-1.12)] and 10% at higher odds of being emergency caesarean section [aOR: 1.10 (95% CI, 1.05-1.15)] compared to women without PCOS (**Table 4.2**). In the sensitivity analysis, among deliveries of women with a diagnostic code for PCOS and their matched delivery records, the increased odds for instrumental vaginal delivery was no longer evident and for elective/other/unspecified caesarean section was less pronounced [aOR: 1.00 (95% CI, 1.00-1.00) and 1.03 (95% CI, 1.02-1.03), respectively], while there was a more pronounced increased odds of emergency caesarean section delivery [aOR: 1.15 (95% CI, 1.02-1.30)] **(Supplementary 37)**.

4.4.3.3 Birthweight

The proportion of at least one of the babies in a single delivery being born with high birthweight (>4000g) did not differ significantly between delivery records of women with and without PCOS [9.82% vs 9.64%, OR: 1.02 (95% CI, 0.98-1.07), aOR: 0.97 (95% CI, 0.92-1.01)]. The

proportion of low birthweight (<2500g) was significantly higher among deliveries of women with PCOS compared to women without PCOS (5.90% vs 5.35%), with an 11% increase in the crude odds of low birthweight [OR: 1.11 (95% CI, 1.05-1.18)]. However, this was insignificant in the fully adjusted model [aOR: 1.03 (95% CI, 0.95-1.13)] (**Table 4.2**).

In the sensitivity analysis, in the fully adjusted model, there was no increased risk of either high or low birthweight of babies born to mothers with PCOS compared to mothers without PCOS [aOR: 1.00 (95% CI, 0.88-1.13) and 1.03 (95% CI, 0.77-1.37), respectively] **(Supplementary 37)**.

When standardizing the birthweight using INTERGROWTH 21st project tools to consider the outcomes large and small for gestational age (LGA and SGA), there was a significant association between maternal PCOS and LGA babies in the unadjusted model [uOR: 1.07 (95% CI, 1.03-1.11)], which became non-significant when adjusting for pre-gravid BMI. There was no statistically significant association between maternal PCOS and odds of either LGA or SGA in the fully adjusted analysis [aOR: 1.00 (95% CI, 0.97-1.04) and 1.03 (95% CI, 0.96-1.11), respectively] (**Table 4.3**). In the fully adjusted sensitivity analysis, there was no significant association between maternal PCOS and LGA [aOR: 1.08 (95% CI, 0.99-1.18)], similar to the primary analysis; however, there was 26% lower odds of SGA in deliveries among women with PCOS compared to women without PCOS [aOR: 0.74 (95% CI, 0.59-0.94)] **(Supplementary 38)**.

4.4.3.4 Stillbirth

Among women with and without PCOS, the proportion of deliveries with stillbirth was 0.44% and 0.43%, respectively, and there was no significant difference in the crude or adjusted odds of stillbirth in either the primary or sensitivity analysis [aOR: 0.99 (95% CI, 0.81-1.21) and 0.52 (95% CI, 0.27-1.02), respectively].

4.5. Discussion

4.5.1 Summary of findings

In this retrospective cohort study of hospital-based delivery records, women with PCOS were found to be at an increased risk of preterm delivery and caesarean section compared to women without PCOS, even after accounting for several confounders including sociodemographic variables, pre-existing maternal conditions such as dysglycaemia, hypertension, and thyroid disorders, number of babies born at the delivery and pre-gravid BMI. Furthermore, women with PCOS were found to be crudely at an increased risk of delivering small babies weighing below 2.5 kg, however the association disappeared after adjustment for gestational age. This was further supported by the absence of evidence of increased risk of babies born small for gestational age, suggesting that lower birthweight of babies born to mothers with PCOS was mediated by their lower gestational age at delivery. This also highlights the importance of standardising birthweight against gestational age using anthropometric reference data to define optimal fetal growth outcomes as opposed to using absolute birthweight. An increased risk of babies born large for gestational age was found among women with PCOS, but the association became insignificant with adjustment for pre-gravid BMI, suggesting that LGA is mediated by maternal pre-gravid BMI. There was no evidence of association between maternal PCOS and the risk of stillbirth.

4.5.2 Strengths and limitations

This study has many strengths including large sample size, and population-based data collected from primary care records and hospital episode statistics birth records. One of the limitations might be the underdiagnosis of PCOS within the data source used. It is notable that across different settings, women with PCOS experience long delays in diagnosis and tend to report their symptoms multiple times prior to a diagnosis.[3] Women with a diagnostic code for PCO, or a combination of symptom codes indicating a missed PCOS diagnosis based on the

Rotterdam criteria were therefore included, which constituted 83% of the exposed women included in the primary analysis. This higher estimate of missed PCOS diagnosis in comparison to the literature[8,9]may have introduced misclassification within the PCOS exposure group. Therefore, a sensitivity analysis was performed including only women with a diagnostic code for PCOS and their age matched controls. Women with a diagnostic code for PCOS within primary care may reflect those with a severe phenotype associated with the combination of menstrual irregularity and androgen excess, who consulted their general practitioners for treatment and management.[1] In agreement with this, the results of the sensitivity analysis, restricted to women with a diagnostic code for PCOS and their matched controls, suggest a more profound and significant odds ratio for preterm, very preterm and extremely preterm delivery compared to results from the primary analysis.

A limitation of the study is the missing outcome data, for which multiple imputation was performed. Furthermore, information on some of the confounders including maternal education level, primigravidity were unavailable within the data source used. Another limitation of this study is the restriction of the eligible cohort to deliveries recorded within the hospital setting, thereby missing deliveries that happened elsewhere such as in non-NHS hospitals or in the home setting. This may affect the generalizability of this study's findings. However, 96% of deliveries in England are recorded within HES data.[123]

Another limitation of the study is the absence of data on mode of conception; it was therefore not feasible to evaluate any effect modification attributable to in-vitro fertilization when assessing the association between PCOS and risk of obstetric outcomes. The increased risk of obstetric outcomes among women with PCOS observed in this study may therefore be attributable to a combination of exposures to PCOS and in-vitro fertilization, a prevalent mode of conception among women with PCOS.

Pregnancy induced complications or gestational weight gain were not adjusted for in the analysis of this study as these constitute intermediates between pre-pregnancy risk factors and obstetric outcomes. It is well established that women with PCOS are at an increased risk of developing antepartum complications such as gestational diabetes, pregnancy induced hypertension and pre-eclampsia.[124] Considering the increased risk of preterm delivery conferred by these pregnancy complications,[125,126] it is possible that pregnancy complications formed the interlink between maternal PCOS and the risk of preterm delivery. Furthermore, caesarean section may be considered for the management of women presenting with suspected or established preterm labor.[127] This complex biological pathway mediated by pregnancy induced complications could potentially explain the increased risk of preterm and operative delivery observed in this study.

4.5.3 Comparison with existing literature

This study is in agreement with existing reviews[41,112,128] and a recent Swedish nationwide cohort study[129] and confirms the association between maternal PCOS and preterm birth of varying degree. However, the adjusted odds ratios observed in this study for preterm birth is modest compared to the odds ratios reported in the literature. This may be attributed to several factors including differences in the source population, exposure definition and residual confounding. Furthermore, genome-wide association studies have indicated a genetic polymorphism (EBF-1 gene) to be associated with both women's likelihood of delivery preterm[130] and progression of PCOS,[131] providing a plausible genetic explanation to the finding. In addition, a dysregulated hypothalamic-pituitary-adrenal (HPA) axis, as observed in both women with PCOS[132] and manifested during stress,[133] has been associated with a modest increased risk of spontaneous preterm delivery, further supporting these findings. This study is also in agreement with reviews and cohort studies that suggest an increased risk of caesarean delivery.[60,112] The findings of absence of significant association of maternal PCOS

with stillbirth is supported by Roos et.al.,[60] while a more recent study by Valgeirsdottir et.al.,[134] suggests a 50% increased risk of stillbirth among women with PCOS, although the exposure ascertainment within the study suffers from misclassification due to inclusion of women with anovulation as well as women with PCOS.

4.5.4 Implications

With a PCOS diagnosis, women have expressed concerns about infertility and pregnancy,[135] and would benefit from the awareness of their pregnancy and delivery related risks, and evidence based surveillance and care to avert these risks. Future research is needed to understand the pathophysiological underpinnings of maternal PCOS on the risk of obstetric outcomes, so that interventions can be designed to reduce these risks.

4.5.5 Conclusion

Women with PCOS are at an increased risk of obstetric outcomes including preterm and operative delivery. Association with low birthweight maybe mediated by lower gestational age at delivery.

Chapter 5 - Increased COVID-19 infections in women with polycystic ovary syndrome: a population-based study

Published article

European Journal of Endocrinology 184.5 (2021): 637-645. < https://doi.org/10.1530/EJE-20-1163>

Anuradhaa Subramanian[1*], Astha Anand[1*], Nicola J. Adderley[1], Kelvin Okoth[1], Konstantinos A. Toulis[1], Krishna Gokhale[1], Christopher Sainsbury[1], Michael W. O'Reilly[2,3], Wiebke Arlt[3,4], Krishnarajah Nirantharakumar[1,5]

* Joint first authors

[1] Institute of Applied Health Research, University of Birmingham, Birmingham, UK

[2] Royal College of Surgeons in Ireland (RCSI) University of Medicine and Health Sciences, Dublin, Republic of Ireland

[3] Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK

[4] National Institute for Health Research (NIHR) Birmingham Biomedical Research Centre, University Hospitals Birmingham NHS Foundation Trust and University of Birmingham, Birmingham, UK

[5] Midlands Health Data Research UK, Birmingham, UK

A.S conceived the idea for the study with guidance from W.A and K.N. A.S extracted the data using DExtER software. A.S designed the study and performed data cleaning and analysis. A.S wrote the initial draft of the manuscript with support from A.A, N.J.A, M.W.O and K.A.T. All authors reviewed and revised the manuscript. All authors read and approved the final manuscript.

5.1 Abstract

5.1.1 Objective

Several recent observational studies have linked metabolic co-morbidities to an increased risk from COVID-19. Here I investigated whether women with PCOS are at an increased risk of COVID-19 infection.

5.1.2 Design

Population-based closed cohort study between 31st January 2020 and 22nd July 2020 in the setting of a UK primary care database (The Health Improvement Network, THIN).

5.1.3 Methods

Main outcome was incidence of COVID-19 coded as suspected or confirmed by the primary care provider. Cox proportional hazards regression model was used with stepwise inclusion of explanatory variables (age, body mass index, impaired glucose regulation, androgen excess, anovulation, vitamin D deficiency, hypertension, and cardiovascular disease) to provide unadjusted and adjusted hazard ratios (HR) of COVID-19 infection among women with PCOS compared to women without PCOS.

5.1.4 Results

A total of 21,292 women were identified with a coded diagnosis of PCO/PCOS and were matched randomly to 78,310 control women for age and general practice. The crude COVID-19 incidence was 18.1 and 11.9 per 1,000 person-years among women with and without PCOS, respectively. Age-adjusted Cox regression analysis suggested a 51% higher risk of COVID-19 among women with PCOS compared to women without PCOS [HR: 1.51 (95% CI, 1.27-1.80), p<0.001]. After adjusting for age and BMI, HR reduced to 1.36 (95% CI, 1.14-1.63), p=0.001. In the fully adjusted model, women with PCOS had a 28% increased risk of COVID-19 [aHR: 1.28 (95% CI, 1.05-1.561), p=0.015].

5.1.5 Conclusion

Women with PCOS are at an increased risk of COVID-19 infection and should be specifically encouraged to adhere to infection control measures during the COVID-19 pandemic.

5.2 Introduction

The novel severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) reached pandemic status in March 2020, with a consequent severe impact on international healthcare systems and the global economy.[66] The resulting coronavirus disease 2019 (COVID-19) causes mild symptoms in most cases, but the incidence of severe illness, respiratory failure and mortality in high-risk groups has led to mandated quarantine measures and economic shutdown across the globe in order to protect capacity within health systems and intensive care unit.[67,68] Multiple large observational studies have shown that those with metabolic risk factors such as diabetes, obesity and cardiovascular disease are at higher risk of severe COVID-19 infection.[67,68,136,137] Shielding strategies are recommended for older patients and for those with significant comorbidities that place them in a high-risk bracket for severe COVID-19 infection, including being immunocompromised or pregnant, or for those with health conditions such as diabetes, heart, liver and lung disease.

Women with polycystic ovary syndrome (PCOS) have recently been highlighted as an overlooked and potentially high-risk population for COVID-19 complications.[70] PCOS is a lifelong metabolic condition of women, typically associated in most cases with androgen excess, anovulatory infertility and polycystic ovarian morphology on ultrasound.[3–5] PCOS has an estimated population prevalence between 8-16% of all women, depending on the population studied.[6,7] Women with PCOS are at significantly increased risk of type 2 diabetes mellitus (T2DM),[29,138,139] non-alcoholic fatty liver disease (NAFLD)[26] and cardiovascular disease.[140] PCOS prevalence is also notably higher in black and South Asian women than in white women,[141] the former appear to have a higher risk of severe COVID-19.[142–146] Whilst younger age and female sex are typically associated with a lower overall risk of severe COVID-19 infection and mortality,[67,136] patients with PCOS may represent a distinct subgroup of women

at higher-than-average risk of adverse COVID-19-related outcomes. It is therefore imperative to determine whether PCOS is linked to COVID-19 susceptibility.

It was hypothesized that women with PCOS are at a higher risk of development of COVID-19 compared to an age-matched control population. The aim was to examine the incident risk of reported suspected/confirmed COVID-19 in women with PCOS in the UK utilizing a large primary care database, in comparison to matched population controls.

5.3 Methods

5.3.1 Study design and data source

A population-based retrospective closed cohort study to determine the incident risk of COVID-19 infection in women with PCOS in comparison to women without PCOS was conducted in The Health Improvement Network (THIN) database. THIN is an anonymized longitudinal primary care electronic medical records database from 365 active general practices in the UK. The records include patient demographics data, symptoms, diagnoses, drug prescriptions, physical measurements, and laboratory tests results. Symptoms and diagnoses are recorded using Read codes, a hierarchical coding system.[147] Researchers at University of Birmingham have previously conducted studies examining long-term outcomes of women with PCOS using the THIN database.[26,44]

5.3.2 Study population

Women aged 18 and above were included if they had a minimum registration period of one year with an eligible general practice to maximize completeness of baseline records. Patient age at study entry (31st of January 2020) was not restricted to reproductive age considering the lifelong metabolic disturbances associated with this condition. Women with a coded diagnosis of PCOS or polycystic ovaries (PCO) before study entry were included in the PCOS cohort. For the purposes of this study, women with a coded diagnosis of PCO were considered as women with PCOS as previous studies have highlighted these codes have been interchangeably

recorded in the primary care electronic medical records in the UK.[9] Read codes for PCOS and PCO are listed in

Supplementary *2* and **Supplementary 7** respectively. Women who are pregnant at study entry were excluded from cohort selection as they are more likely to be tested for COVID-19, due to systematic screening during admission for delivery,[148] which could affect the primary outcome. For every woman with PCOS, four women without a diagnostic code for PCOS/PCO were selected, matched for age- (±1 year) and general practice location.

5.3.3 Outcome and follow-up

The primary outcome was a composite of suspected or confirmed diagnosis of COVID-19 in primary care; Read codes are listed in **Supplementary 39**. According to NHS Guidance and Standard Operating Procedures for Primary Care, and UK Faculty of Clinical Informatics guidelines, confirmed COVID-19 codes represents a positive RT-PCR test, while a suspected COVID-19 code represents a symptomatic presentation of COVID-19 and/or contact history with a confirmed patient.[149] Considering the wide unavailability of RT-PCR tests outside of the hospital setting until relatively later in the initial wave of COVID-19 pandemic in the UK, most cases of COVID-19 in primary care are coded as "suspected". All women included in the study were followed up from 30th January 2020 (index date) until patient exit date: A patient was considered to exit the study at the earliest of the suspected/confirmed COVID-19 infection documentation date or the patient being lost to follow-up (i.e., patient deregistration from the practice or patient death) or study end date (22nd July 2020, last date of data provided by Cegedim, the THIN data provider).

5.3.4 Explanatory variables

PCOS features that overlap with COVID-19 infection risk were considered as explanatory variables, which included age, BMI, impaired glucose regulation, androgen excess, anovulation (lack of regular ovulation or symptomatic sequelae of anovulation) , vitamin D

deficiency at baseline, and concurrent diagnosis of hypertension and cardiovascular diseases, informed by previously identified COVID-19 risk factors.[70] Age was categorized into 10-year age bands: 18-30, 30-40, 40-50, 50-60 and 60+ years. BMI was considered as a continuous variable. Multiple imputation using chained equations and predictive mean matching was performed to replace missing BMI values. Impaired glucose regulation was categorized as either (1) absence of diabetes, (2) pre-diabetes or (3) diabetes (identified by either Read code records or the HbA1c measurement at baseline (42 to 47 mmol/mol (6.0-6.4%) for pre-diabetes and ≥48 mmol/mol (>=6.5%) for diabetes). Androgen excess was defined as the latest serum testosterone measurement ≥2.0 nmol/L at baseline and/or the presence of hirsutism. Vitamin D deficiency was identified by Read codes. Cardiovascular disease was defined as a composite of ischemic heart disease, heart failure, stroke, transient ischemic attack, and peripheral vascular disease.

5.3.5 Statistical Analysis

Description of baseline variables are provided using appropriate summary statistics stratified as PCOS and non-PCOS group. Mean with standard deviation (SD) and median with interquartile range (IQR) were provided for continuous variables as appropriate. Frequency and percentage were provided for categorical variables. T-test and chi-squared test were used to test for statistically significant differences in the baseline variables between PCOS and non-PCOS.

A Cox proportional hazards regression model was used to provide unadjusted and adjusted hazard ratios (HRs) of the primary outcome among women with PCOS compared to women without PCOS after stepwise inclusion of the explanatory variables in the Cox model, culminating with a fully adjusted model.

Two sets of sensitivity analyses were performed to assess the robustness of the findings. Firstly, the exposed cohort was restricted to patients with a coded diagnosis of PCOS only (instead of

PCOS/PCO) and performed the Cox regression analysis along with their matched controls. Secondly, the analyses were restricted to patients of reproductive age (18-50) at study entry and through the study period.

All analyses were performed in Stata IC version 15. Two-sided P values were obtained for all tests, and a P value <0.05 was considered as statistically significant. Selection of Read code lists was performed using methods used in previous publications.[150]

5.3.6 Ethics

The THIN data collection scheme and research carried out using THIN data were approved by the NHS South-East Multicentre Research Ethics Committee in 2003. Under the terms of the approval, studies must undergo independent scientific review. Approval for this study was obtained from the THIN Scientific Review Committee in September 2020 (SRC protocol reference 20-010)

5.4 Results

5.4.1 Characteristics of the cohort of women with PCOS and their age-matched controls

As of 31st January 2020, 326 practices out of 365 practices qualified for inclusion, with 1,012,944 registered women aged 18 and above. A total of 8,103 women with a coded diagnosis of PCOS and 13,189 additionally with a coded diagnosis of PCO were identified. From a pool of 969,162 women eligible to be in the control population, a total of 78,310 women were randomly selected as controls, matched for age and GP surgery location.

The mean (SD) age at study entry of the women with and without PCOS was 39.3 (11.1) and 39.5 (11.3), respectively (**Table 5.1**). Among the women with PCOS, the mean (SD) age at diagnosis of PCOS and mean (SD) duration after the diagnosis of PCOS at study entry were 27.0 (7.0) years and 12.4 (8.9) years, respectively.

Table 5.1: Baseline characteristics of women with PCOS and age-matched controls.

	Exposed (PCO/PCOS) (n=21,292)	Unexposed (n=78,310)	P-value
Age [Mean (SD)]	39.3 (11.1)	39.5 (11.3)	0.030¥
Age [Median (IQR)]	38.5 (30.5-46.5)	38.5 (30.5-47.5)	
Age categories			0.009§
18 - 30 years	4697 (22.1)	17403 (22.2)	
30 - 40 years	7170 (33.7)	25591 (32.7)	
40-50 years	5658 (26.6)	20736 (26.5)	
50-60 years	2951 (13.9)	11358 (14.5)	
>60 years	816 (3.8)	3222 (4.1)	
BMI [Mean (SD)]	31.0 (8.4)	27.1 (6.7)	<0.001¥
BMI [Median (IQR)]	29.7 (24.4-36.2)	25.50 (22.3-30.4)	
BMI Categories			<0.001§
Normal/Underweight (<25 kg/m^2)	5530 (26.0)	31671 (40.4)	
Overweight (25-30 kg/m^2)	4494 (21.1)	18112 (23.1)	
Obese (>30 kg/m^2)	9538 (44.8)	17837 (22.8)	
Missing	1730 (8.1)	10690 (13.7)	
Androgen Excess*	4849 (22.8)	1399 (1.8)	<0.001§
Testosterone ≥2.0 nmol/L	2552 (12.0)	665 (0.9)	<0.001§
Hirsutism	2838 (13.3)	786 (1.0)	<0.001§
Anovulation	5867 (27.6)	5770 (7.4)	<0.001§
IGR categories			<0.001§
Absence of IGR	18,767 (88.14)	74,590 (95.25)	
Pre-diabetes	873 (4.10)	1673 (2.14)	
Diabetes	1,652 (7.76)	2,047 (2.61)	
Vitamin D Deficiency	627 (3.0)	1398 (1.8)	<0.001§
Hypertension	2023 (9.5)	4404 (5.6)	<0.001§
Composite CVD ¬	45 (1.6)	984 (1.3)	<0.001§
Ischemic Heart Disease	175 (0.8)	484 (0.6)	0.001§
Stroke/TIA	128 (0.6)	381 (0.5)	0.038§
Heart Failure	51 (0.2)	139 (0.2)	0.066§
Peripheral Vascular Disease	31 (0.2)	91 (0.1)	0.277§

*(Hirsutism /Testosterone≥2.0 nmol/L)
¬(Ischemic Heart Disease/Stroke/TIA/Heart Failure/Peripheral Vascular Disease)
BMI-Body Mass Index; IGR-Impaired Glucose Regulation; TIA-Transient Ischemic Attack
¥ - P-value obtained from t-test comparing means of the variable between the two groups.
§ - P-value obtained from chi-squared test comparing the percentage of women in each category between the two groups.

As anticipated, there were significantly higher levels of all characteristic features of PCOS among the women with PCOS than in the matched controls (**Table 5.1**). Out of the women with a record of BMI at baseline (91.9% and 86.3% among women with and without PCOS), women with PCOS had significantly higher BMI compared to women without PCOS (mean (SD): 31.0 (8.4) vs 27.1 (6.7), $p<0.001$). Androgen excess, defined as a coded diagnosis of hirsutism and/or the latest recorded serum testosterone measurement ≥2.0 nmol/L prior to study entry, was recorded for 22.8% and 1.8% of women with and without PCOS, respectively ($p<0.001$). A coded diagnosis of anovulation at baseline was recorded for 27.6% and 7.4% of the women with and without PCOS, respectively ($p<0.001$). At baseline, approximately 7.8% and 4.1% of the women with PCOS had diabetes and pre-diabetes, respectively, while only 2.6% and 2.1% of the control women had records of these conditions ($p<0.001$). Women with PCOS were more likely to be vitamin D deficient (3.0% vs 1.8%, $p<0.001$), hypertensive (9.5% vs 5.6%, $p<0.001$), or have cardiovascular disease at baseline (1.6% vs 1.3%, $p<0.001$).

Table 5.2: Risk of suspected/confirmed COVID-19 among women with PCOS compared to women without PCOS

	Exposed	Unexposed
Primary Analysis	**(n=21,292)**	**(n=78,310)**
Outcome events, n (%)	180 (0.85)	438 (0.56)
Person-years	9,967	36,727
Crude Incidence Rate/1000 PY	18.06	11.93
Unadjusted Hazard ratio (95% CI)	1.52 (1.27-1.80) p<0.001	
Adjusted Hazard ratio (95% CI)*	1.28 (1.05-1.56) p=0.015	
Sensitivity Analysis		
Restriction of exposure to PCOS codes only	**(n=8,103)**	**(n=29,711)**
Outcome events, n (%)	69 (0.85)	160 (0.54)
Person-years	3,788	13,926
Crude Incidence Rate/1000 PY	18.21	11.49
Unadjusted Hazard ratio (95% CI)	1.59 (1.20-2.10) p=0.001	
Adjusted Hazard ratio (95% CI)*	1.38 (0.99-1.92) p=0.056	
Restriction of cohort to women of reproductive age (18-50)	**(n=17,525)**	**(n=63,775)**
Outcome events, n (%)	152 (0.87)	353 (0.55)
Person-years	8,180	29,546
Crude Incidence Rate/1000 PY	18.58	11.95
Unadjusted Hazard ratio (95% CI)	1.56 (1.29-2.88) p<0.001	
Adjusted Hazard ratio (95% CI)*	1.30 (1.05-1.62) p=0.018	

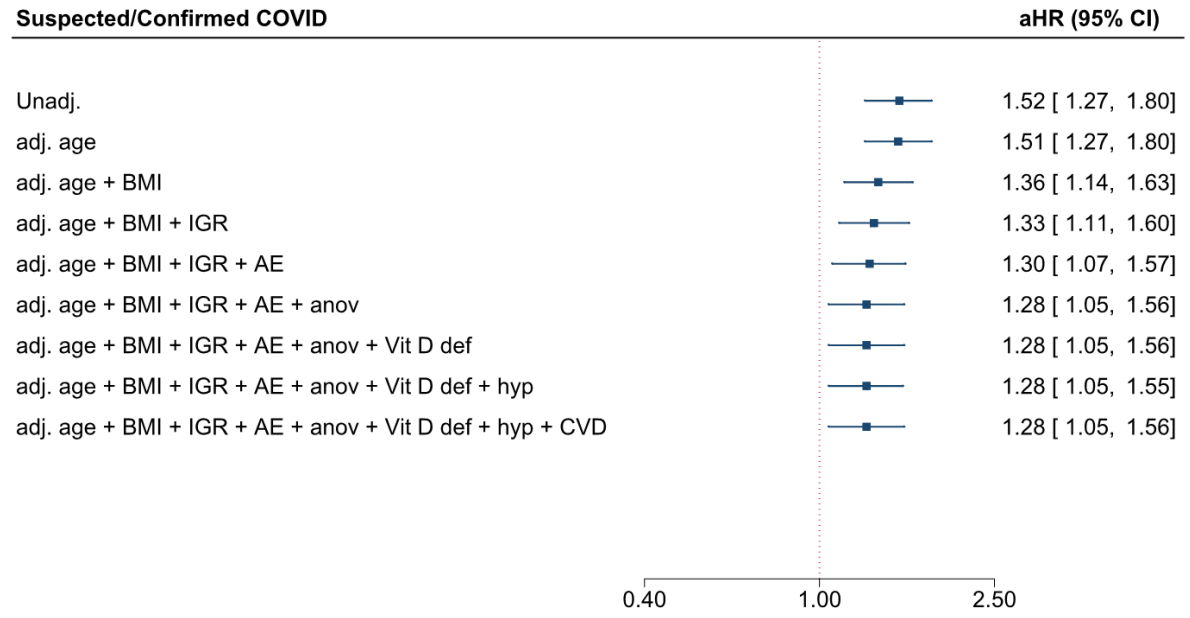
*Adjustment for age category, body mass index, impaired glucose regulation, androgen excess, anovulation, vitamin D deficiency, hypertension, and cardiovascular diseases

5.4.2 Risk of COVID-19 among women with PCOS compared to their age-matched controls, after adjustment for PCOS features

Among the women with and without PCOS, 0.9% (n=180) and 0.6% (n=438), respectively, had a record of suspected/confirmed COVID-19 in their primary care records during a cumulative follow-up of 9,967 and 36,727 person-years, respectively (**Table 5.2**). Confirmed COVID-19 codes were only present in 0.1% (n=14) and 0.1% (n=70) of women with and without PCOS respectively. This provided a crude COVID-19 incidence rate of 18.1 and 11.9 per 1,000 person-years among the women with and without PCOS, respectively. An age adjusted Cox regression analysis suggested a 51% higher risk of suspected/confirmed COVID-

19 among women with PCOS compared to women without PCOS [1.51 (95% CI, 1.27-1.80), p<0.001] (**Figure 5.1**).

Figure 5.1: Risk of confirmed/suspected COVID-19 among women with PCOS after serial adjustment for PCOS features.



BMI=Body Mass Index; IGR=Impaired Glucose Regulation; AE=Androgen Excess; anov=Anovulation; hyp=Hypertension; CVD=Cardiovascular disease

After adjusting for BMI and age, the hazard ratio reduced to 1.36 (95% CI, 1.14-1.63), p=0.001 (**Figure 5.1**). When additionally adjusting for impaired glucose regulation, the hazard ratio was marginally further reduced to 1.33 (95% CI, 1.11-1.60), p=0.002. Following this, in a series of further stepwise adjustments for androgen excess and anovulation, the hazard ratios were reduced to 1.30 (95% CI, 1.07-1.57), p=0.008 and 1.28 (95% CI, 1.05-1.56), p=0.014 p=0.018, respectively. Additional adjustment for vitamin D deficiency, hypertension and cardiovascular disease made no difference to the effect estimate. In the fully adjusted model, women with PCOS had a 28% increased risk of suspected/confirmed COVID-19 compared to women without PCOS [aHR: 1.28 (95% CI, 1.05-1.56), p=0.015].

When restricting the exposure ascertainment to codes specific to PCOS only, i.e., excluding PCO codes, a 37% increased risk of suspected/confirmed COVID-19 was observed among women with PCOS (n=8,103) compared to their matched controls (n=29,711), although, the results did not reach statistical significance [aHR: 1.38 (95% CI, 0.99-1.92), p=0.056].

In the sensitivity analysis restricting to reproductive aged women, the results suggest that women with PCOS between the age of 18-50 years (n=17,525) have a 30% increased risk of suspected/confirmed COVID-19 compared to women without PCOS matched for age and general practice (n=63,775) [aHR: 1.30 (95% CI, 1.05-1.62), p=0.018] **(Table 5.2**).

Table 5.3: Risk factors for confirmed/suspected COVID-19 from the fully adjusted model

Risk factors	Adjusted Hazard ratio
PCOS	1.28 (1.05-1.56), p=0.015
Age category	
18-30 years	Reference standard
30-40 years	0.89 (0.71-1.06), p=0.286
40-50 years	1.03 (0.82-1.29), p=0.785
50-60 years	0.89 (0.68-1.18), p=0.428
≥60 years	0.41 (0.23-0.74), p=0.003
BMI	1.02 (1.01-1.03), p<0.001
Androgen excess	1.11 (0.83-1.50), p=0.478
Anovulation	1.06 (0.84-1.35), p=0.594
Impaired Glucose Regulation	
Absence of IGR	Reference standard
Pre-diabetes	1.31 (0.86-2.00), p=0.215
Diabetes	1.36 (0.96-1.93), p=0.085
Vitamin D deficiency	1.61 (1.05-2.47), p=0.029
Hypertension	1.19 (0.88-1.62), p=0.258
Cardiovascular Disease	1.88 (1.12-3.17), p=0.017

5.4.3 Risk factors for COVID-19 among all women

In the fully adjusted model, there was lower risk of reported suspected/confirmed COVID-19 among women aged ≥60 years compared to women aged 18-30 [aHR: 0.41 (95% CI, 0.23-0.74), p=0.001] and 2% higher risk with every unit (kg/m2) increase in BMI [aHR: 1.02 (95% CI, 1.01-1.03), p<0.003] (**Table 5.3**). Furthermore, there was higher risk of suspected/confirmed COVID among women who had vitamin D deficiency [aHR: 1.61 (95% CI, 1.05-2.47), p=0.029] or cardiovascular disease [1.88 (95% CI, 1.12-3.17), p=0.017] at baseline. Risk was also higher in the presence of prediabetes and type 2 diabetes, but this did

not reach statistical significance [aHR 1.31 (95% CI, 0.86-2.00), p=0.215 and 1.36 (95% CI, 0.96-1.93), p=0.085, respectively].

5.5 Discussion

5.5.1 Summary of findings

In this retrospective cohort study spanning the first wave period of the COVID-19 pandemic in the UK, it was found that a diagnosis of PCOS confers a 51% increased risk of development of confirmed or suspected COVID-19 infection compared to the background age-matched female population. A higher observed susceptibility to COVID-19 infection (26%) in the PCOS cohort persisted even after adjustment for individual cardio-metabolic risk factors known to cluster within PCOS, which have recently been directly linked to increased COVID-19 susceptibility including obesity, impaired glucose regulation and androgen excess.[67,70,151] These data support an independent relationship between a diagnosis of PCOS and risk of COVID-19 infection, however the precise pathophysiological mechanisms underpinning this association are not clear.

5.5.2 Implication

PCOS is a lifelong condition associated with severe health consequences in women, including a significantly increased risk of T2DM, NAFLD and cardiovascular disease.[26,140] To my knowledge, this is the first publication since the pandemic outbreak that has demonstrated an increased susceptibility to COVID-19 infection in women with PCOS. Given the high prevalence of PCOS in the population, these findings need to be considered when designing public health policy and advice as our understanding of COVID-19 evolves. Before the onset of the COVID-19 pandemic, women with PCOS had low rates of satisfaction with access to and provision of healthcare services in relation to their condition.[152] Women with PCOS consistently report fragmented care, delayed diagnosis and a perception of poor clinician understanding of their condition as major factors contributing to this dissatisfaction.[153] Women

suffering from this condition may fear, with some degree of justification, that an enhanced risk of COVID-19 infection will further compromise timely access to healthcare and serve to increase the sense of disenfranchisement currently experienced by many patients. The pandemic has already dramatically altered our current healthcare delivery models, and although the increased rollout of virtual consultations and methods of delivering remote healthcare have been commendable, for many patients with PCOS these will not be an appropriate substitute for the traditional clinician-patient live consultation. The risk of mental health problems including low self-esteem, anxiety and depression is significantly higher in women with PCOS than the background female population, and advice on strict adherence to social distancing needs to be tempered by the associated risk of exacerbating these underlying problems.

5.5.3 Plausible biological mechanisms

PCOS is a pro-inflammatory state, and it has been hypothesized that inflammation may underpin many of the cardio-metabolic abnormalities in this disorder.[154] Increased circulating levels of pro-inflammatory mediators, including highly sensitive C-reactive protein (hsCRP), tumor necrosis factor (TNF)-alpha, procalcitonin and interleukin-18 (IL-18), have been reported in women with PCOS,[155,156] and although more pronounced in the context of obesity, these associations persist even after correction for total fat mass. Pro-inflammatory cytokines are implicated in adipose tissue dysfunction and inflammation,[157] and have been implicated in the pathophysiology of insulin resistance and diabetes.[158] Severe COVID-19 infection, with associated respiratory failure requiring oxygen therapy or admission to intensive care for intubation and ventilation, has also been linked with an exaggerated systemic inflammatory response, which can trigger catastrophic acute respiratory distress syndrome (ARDS) with associated multi-organ failure and high mortality. It is conceivable that women with PCOS, who have been demonstrated to have low-grade inflammation beyond that observed in simple

obesity , are potentially at increased risk of severe COVID-19 infection because of this underlying pro-inflammatory predisposition.[159,160]

The link between COVID-19 infection and androgens merits further discussion. Androgen excess is a cardinal feature of PCOS and identified as a primary driver of increased risk of T2DM and NAFLD in affected women.[26,27] Significant gender differences have been observed in COVID-19 outcomes, with a higher likelihood of hospitalization and death in men reported in multiple studies.[159] Intriguingly, androgen deprivation therapy in men treated for prostate cancer was associated with a significantly reduced risk of SARS-COV-2 infection compared to those treated with alternative disease regimens in a recent study;[160] a preliminary report from Spain has linked more severe infection with androgenic alopecia in male patients.[161] Conversely, a limited number of small studies have also linked low serum testosterone at baseline in hospitalised men to an increased risk of intensive care unit (ICU) admission and death;[162,163] indeed, it is intriguing that the metabolic complications associate with male hypogonadism mirror those of women with androgen excess.[164] Early in vitro studies suggest that transmembrane serine protease 2 (TMPRSS2), which is highly regulated by androgens, is a critical enzyme mediating the entry of the SARS-CoV-2 into cells.[165] It is reasonable to speculate that women with PCOS and androgen excess are at increased susceptibility of infection through this mechanism. Whilst androgen excess was not identified as a major contributor to COVID-19 susceptibility in this PCOS cohort, it is likely to be the subject of increased clinical research interest in the months and years ahead. In addition, it is a limitation of this study that the diagnosis of androgen excess was based on surrogate parameters, hirsutism, and serum testosterone concentrations. Testosterone has not been systematically measured in the PCOS cohort of this study, with no data available on 11-oxygenated androgens, the predominant circulating androgens in PCOS.[166] Androgens are important modulators of immune function,[167] and very recent observations have highlighted that peripheral blood

mononuclear cells preferentially activate 11-oxygenated androgens and that natural killer cells, the prime innate defence against viral infection, represent the major site of this intracrine androgen activation.[168]

5.5.4 Strengths and limitations

Strengths of this study include a large sample size from a dataset generalizable to the UK population and the study period covers the majority of the COVID-19 pandemic duration in the UK to date (at the time of manuscript submission for publication). The proportion of missing information was low, and a range of potential confounders were adjusted for in a stepwise series of regression models. However, there are several important limitations. The data quality is dependent on accurate coding by general practitioners and primary care administrative staff; there is a possibility of miscoding of the PCOS/PCO diagnosis and recording of suspected or confirmed COVID-19 may be incomplete.

A considerable limitation in this study was the restriction of PCOS ascertainment using clinical codes recorded by a general practitioner. Endocrinological evaluation is more likely to be performed by a specialist in secondary care, while the GP may limit to the coding of a confirmed diagnosis. Therefore, prevalence of PCOS observed in primary care setting are usually under recorded rather than over diagnosed. Importantly, it was not possible to explore or adjust for the effect of patient ethnicity or socioeconomic status, as this data was unavailable. While a number of important confounders were adjusted for, there remains a possibility of unmeasured confounding. Also, confounders such as androgen excess and impaired glucose regulation were restricted to clinical coding and available measurements such as serum testosterone and HbA1c, which may not have captured the complete picture of metabolic disturbances. Finally, during the first wave of COVID-19 in the UK there was no widespread testing in primary care, with a COVID-19 test generally only being performed if a patient was

admitted to hospital; a combination of confirmed and clinically suspected COVID-19 infections were presented.

5.5.5 Conclusion

In conclusion, this study shows that women with PCOS are at an increased risk of COVID-19 infection, and except for obesity the adjustment for potentially confounding factors did not mitigate this risk, pointing at inherent PCOS-specific factors. Future studies should explore the potentially critical role of androgens in conveying this risk and assess in more detail the contribution of ethnicity and socioeconomic deprivation. Based on the results, women with PCOS should be specifically encouraged to adhere to the recommended infection control measures for the duration of the COVID-19 pandemic.

Chapter 6 - Combined oral contraceptive pills and risk of impaired glucose regulation among women with Polycystic ovary syndrome: a UK primary care based pharmacoepidemiological study

The contents of this chapter have been published as part of a larger study

Diabetes care vol. 44,12 (2021): 2758-2766. < https://doi.org/10.2337/dc21-0437 >

A.S conceived the idea for the study with guidance from W.A and K.N. A.S extracted the data with support from K.M.G. A.S designed the study with guidance from K.N. A.S performed the analysis. A.S wrote the initial draft of the chapter with support from N.J.A and K.A.T.

6.1 Abstract

6.1.1 Introduction

Polycystic ovary syndrome (PCOS) is characterized by androgen excess and irregular menses; androgens are drivers of increased metabolic risk in women with PCOS. Combined oral contraceptive pills (COCPs) are used in PCOS both for cycle regulation and to reduce the biologically active androgen fraction. I examined COCP use and risk of dysglycemia (prediabetes and type 2 diabetes) in women with PCOS.

6.1.2 Methods

Using a large U.K. primary care database [The Health Improvement Network (THIN); 3.7 million patients from 787 practices], I carried out a nested pharmacoepidemiological case-control study to investigate COCP use in relation to dysglycemia risk (2,407 women with PCOS with and without a diagnosis of dysglycemia during follow-up). Conditional logistic regression was used to obtain adjusted odds ratios (aORs).

6.1.3 Results

Women with PCOS and COCP use had a reduced dysglycemia risk [aOR: 0.72, (95% CI, 0.59–0.87)].

6.1.4 Conclusion

In this study, limited by its retrospective nature and the use of routinely collected electronic general practice record data, which does not allow for exclusion of the impact of prescription-by-indication bias, women with PCOS exposed to COCPs had a reduced risk of dysglycemia. Future prospective studies should be considered for further understanding of these observations and potential causality.

6.2 Introduction

Polycystic ovary syndrome (PCOS) is the most common endocrine disorder in reproductive aged women.[3] Despite a lack of consensus around the diagnostic criteria for PCOS, the most recent consensus among experts is that androgen excess is the cardinal feature of PCOS.[3] PCOS has been associated with several adverse metabolic outcomes including impaired glucose regulation (IGR), obstructive sleep apnoea (OSA) and non-alcoholic fatty liver disease (NAFLD), and literature has revealed androgen excess as the mediator or crucial explanatory factor in the association between PCOS and metabolic outcomes.[26,30,44]

Findings from the retrospective cohort study conducted as part of Chapter 3 of this doctoral thesis have suggested a rise in the incidence of new onset type 2 diabetes and IGR over the last two decades among women with a history of PCOS. Recent data suggests that the risk of diabetes is up to 4-fold higher in women with PCOS, independent of body mass index (BMI), and is diagnosed on average four years earlier than in the background population.[29] A systematic review, with a meta-analysis of 15 and 12 studies looking at the association between PCOS and incident risk of impaired glucose tolerance (IGT) and T2DM respectively, estimated an increase in the prevalence of both IGT and T2DM by approximately 3-fold (3.3 and 2.9 respectively) among women with PCOS compared to non-PCOS women.[28] Circulating androgen burden is closely correlated with insulin resistance and the likelihood of an abnormal oral glucose tolerance test in women with PCOS.[169] Androgen excess is increasingly implicated as a potential risk factor for the onset of hyperglycaemia and other metabolic disturbances in women, and testosterone levels have been shown to predict incident diabetes in population studies.[27] Studies examining the impact of antiandrogen therapy on insulin sensitivity and diabetes risk in PCOS have been for the most part small scale and heterogeneous in nature, with conflicting results.[170,171]

Oral contraceptive pills (OCPs) is used as a contraceptive method of choice by 26% of reproductive aged women in the UK.[172] Lifestyle changes aside, OCPs, more specifically combined oral contraceptive pills (COCPs) are considered to be the mainstay of treatment for menstrual symptoms such as dysmenorrhea, menorrhagia, menstrual migraine and pelvic pain and dermatological symptoms such as acne, hirsutism, and hair loss among women with PCOS. COCPs display a multifaceted and synergistic anti-androgenic mechanism of action; Mainly, at the level of hypothalamus, COCPs consistently inhibit gonadotropin-releasing hormone (GnRH) secretion, leading to an analogous inhibition of both follicle-stimulating hormone (FSH) (to a lesser extent) and luteinizing hormone (LH) (to a greater extent).[173] At the level of pituitary, COCPs disrupt the mid-cycle LH surge.[173] These actions lead to a significant decrease in ovarian secretion of testosterone and other androgens. Additionally, COCPs produce a dose-dependent stimulation of sex hormone-binding globulin (SHBG) production, thereby increasing the high-affinity binding of testosterone to SHBG leading to a decrease in bioactive, free testosterone. In line with this, in a systematic analysis of experimental studies, with a combined total of 1495 reproductive aged women, meta-analysing the results suggested an average decrease in free circulating testosterone by 0.49 nmol/L [MD: -0.49 (95% CI, -0.55, -0.32) nmol/L], and an average increase in SHBG by 99.08 nmol/l [MD: 99.08 (95% CI, 86.43,111.73)], among women who were users of OCPs compared to non-users.[174] At the periphery, COCPs also limit adrenal androgen production and inhibit the peripheral conversion of testosterone to the more potent androgen dihydrotestosterone (DHT), thereby inhibiting DHT binding to nuclear androgen receptors.[175]

Notably, within the population of women with PCOS, there is an unmet need for quality evidence on the long-term metabolic effects of COCP use. In some studies, COCP use has been reported to raise insulin levels and impair glucose sensitivity in the short-term.[176,177] However,

the longer-term clinical implications of reduced androgen levels offered by oral contraceptives, especially among women with PCOS have not been fully elucidated.

Several myths and misconceptions have also been barriers for use of OCP among all women. In a study by Molloy et. al., the authors found that women who believed that "Taking a break from the long-term use of the OCP is a good idea" were 52% less likely to be users of OCPs, after adjusting for several key covariates including marital status, education level and access to medical services [aOR: 0.48 (95% CI, 0.28-0.81)].[178] During an online event "PCOS & The Pill – a discussion marking International Women's Day" on 5th of March 2020, researchers (including myself), women with/without PCOS who have used oral contraceptive pills and experts on PCOS discussed and raised awareness of myths and perceptions that are often cited without evidence such as risk of sterility, weight gain, and depression due to oral contraceptive use.[179]

The two most prescribed types of OCPs include combined oral contraceptive pills (COCPs), followed by progestogen only pills (POPs). COCPs are comprised of an oestrogen component (oestradiol, ethinyloestradiol, or oestetrol) and a progestogen component with varying degrees of antiandrogenic properties, while POPs are comprised of only the progestogen component.

Considering the lack of evidence of protective metabolic effects of COCP use, and the myths revolving around OCP use, there is a need for a methodologically well conducted study to look at the association between COCP use and the risk of IGR among women with PCOS. Further to evaluating the effectiveness of COCPs in the real world, it is also important to estimate the effectiveness of COCPs with and without an antiandrogenic progestin component such as low dose cyproterone acetate. In this chapter, I aimed to determine if COCPs have a protective effect on the incidence of IGR among women with PCOS using a nested case-control designed pharmaco-epidemiological study.

6.3 Methods

6.3.1 Data Source

Data was derived from The Health Improvement Network (THIN) database, a UK general practice database comprising electronic medical records of over 17 million patients from 787 practices, of which around 3.7 million are currently registered with their practice. Read codes, a hierarchical coding system, is used by general practitioners to document patients' clinical, symptom and management information including referrals during each patient encounter. THIN database additionally documents information on sociodemographic data, laboratory investigations such as haemoglobin A1C (HbA1c), 2-hour oral glucose tolerance test (OGTT) tests, and testosterone, physical measurements such as body mass index, lifestyle variables such as smoking status and alcohol consumption, and drug prescriptions. THIN data resembles the population structure of the UK in terms of age, prevalence of certain chronic conditions such as type 2 diabetes and national mortality rates[180].Therefore, the results of this study may be generalisable to the UK population. THIN data has also been used to study several conditions including PCOS,[26] IGR[27] and type 2 diabetes.[181] Notably, the implementation of the Quality and Outcomes Framework (QOF), whereby general practitioners are financially rewarded for appropriate documentation and management of patients with specific chronic conditions, has resulted in improved recording of type 2 diabetes.[182]

6.3.2 Study Period

The study period spanned between 1st January 2000 and 31st January 2017.

6.3.3 Study Design

Nested within a retrospective open cohort study to identify a base cohort of women with PCOS, this pharmacoepidemiological study was a case-control study of women with IGR and their age- (±2 years), BMI- (±2 kg/m2), PCOS diagnosis date- (±2 years), and time of PCOS

diagnosis- (before or after patient eligibility to enter the cohort) matched control women without IGR.

6.3.4 Study Population

Source population was restricted to women registered for at least a year to general practices that met the quality criteria: (1) Practices with acceptable mortality reporting[183] recorded in the previous year or before and (2) practices with minimum 1 year usage of the Vision system for documenting patient medical records. Women became eligible to enter the base cohort if they were aged between 18 and 50 and had a Read code diagnosis of PCOS or polycystic ovaries (PCO). Due to the under recording of PCOS within primary care,[9] similarity in the nomenclature between PCOS and PCO, and consistent with the definition for PCOS from previous studies,[26,44,184] PCO was considered in addition to PCOS to define the base cohort. Women with a record of IGR at baseline (prior to the patient eligibility or prior to the diagnosis of PCOS/PCO, whichever was the latest) were excluded. Furthermore, women with a prescription of insulin or oral diabetic medications including metformin at baseline were excluded since such prescriptions can be considered as a proxy or an indicator of IGR at baseline. Women with a Read code diagnosis of type 1 diabetes at any time during or prior to their registration with the practice were excluded.

6.3.5 Outcome

Record of outcome was identified by Read codes indicating IGR or diabetes mellitus (except type 1 or gestational diabetes mellitus). Read codes that were non-specific to the type of diabetes were still included since these codes were more likely to be indicative of type 2 than type 1 or gestational diabetes mellitus. An alternative means of identifying the outcome was through proxy blood glucose measurements in the range of IGR. These included HbA1c measures $\geq$ 42 mmol/mol, fasting blood glucose $\geq$ 6 mmol/L, random blood glucose $\geq$ 11.1

mmol/L, blood glucose 2 hours after consumption of 75g glucose dissolved in 250 to 300 ml of water ≥ 7.8 mmol/L and/or an additional health data flag suggesting abnormality in the test.

6.3.6 Selection of matched controls

Women without a Read code diagnosis or a proxy measure of IGR formed the pool of potential controls for this study. For every case woman, 1 control woman was randomly selected from the pool of potential controls after matching for age (± 2 years), BMI (± 2 kg/m^2), date of PCOS diagnosis (± 2 years) and incidence of PCOS diagnosis (i.e., a binary record of whether PCOS/PCO diagnosis was made after or prior to patient eligibility).

Sensitivity analysis was performed where the methods for selection of controls were altered. Firstly, risk set sampling was performed, wherein control women were identified from a group of 'at-risk' patients at the index date of the case women. This implied that patients who were cases were allowed to be a control for another case at any time prior to their IGR diagnosis. Secondly, patients were allowed to be selected as controls for more than one case patient (i.e., sampling with replacement).

6.3.7 Exposure and Exposure window

The primary exposure was a binary variable considered as at least one COCP prescription within the pre-specified exposure time window.

Two secondary exposures were considered to explore the effect of anti-androgenic progestin component and a dose response between COCP and the risk of IGR. For the first secondary exposure, women were categorised into one of the three exposure categories of COCP prescription: (1) no prescription, (2) prescription of OCP with a progestin component without anti-androgenic properties (such as levonorgestrel, gestodene desogestrel, norethisterone and norgestimate) and (3) prescription of COCP with anti-androgenic progestin component (such as drospirenone and cyproterone acetate). If a women had prescriptions of COCPs with and without anti-androgenic progestin within the exposure window, then the women was

categorized into the third exposure category of prescription of COCP with anti-androgenic progestin component. For the second secondary exposure, the number of times prescriptions for OCPs were dispensed within the exposure window was counted. The number of prescriptions dispensed was then used to categorize women into one of the three exposure categories of OCP prescription: (1) no prescription, (2) ≤ 3 prescription within the exposure window, and (3) > 3 prescription within the exposure window. Three prescription count was considered as a cut-off count since it was the median number of COCP prescriptions. Prior to estimating prescription counts, multiple prescriptions that were dispensed on the same day were considered as duplicates and duplicates were disregarded.

The exposure window extended from one year prior to cohort entry up to six months prior to index date. Cohort entry was considered as the latest of the following: (1) one year after practice acceptable mortality rate, (2) one year after practice Vision installation date, (3) one year after patient registration with an eligible practice, (4) patient turning age 18, (5) date of Read code diagnosis of PCOS/PCO). Prescriptions up to one year before cohort entry were considered due to delay in the diagnosis of PCOS within primary care.[8] Index date was considered as the date of recording of the outcome for cases. The same date was assigned as the index date for the matched control women in order to avoid time-window bias.[185] Prescriptions only until 6 months prior to the index date were considered since any prescription after this period may not be attributed towards a protective effect against the development of a long-term metabolic condition that is perceived as a continuum of glucose dysregulation.

6.3.8 Covariates

Covariates were selected based on biological plausibility for confounding and were obtained at the time of cohort entry. These included age, BMI category, socio-economic status, ethnicity, smoking status, concurrent diagnosis of hypothyroidism or hypertension, prescription of lipid lowering medication (a proxy for dyslipidaemia), and prescriptions of metformin and other

isolated anti-androgenic drug prescriptions (such as high dose cyproterone acetate, flutamide, finasteride and spironolactone) within the same exposure window.

As per the WHO recommendation for obesity measurement, latest BMI at baseline (in kg/m^2) was categorized as <25 (underweight to normal BMI range - reference category), 25 to 30 (overweight range) and ≥30 kg/m^2 (obesity range). Socio-economic status was categorized based on Townsend quintiles from 1 to 5, 1 denoting affluent status (reference category) and 5 denoting the most deprived. Ethnicity was categorized based on the UK 2011 census classification as: (1) white Caucasian (reference category), (2) black Afro-Caribbean, (3) South Asian, (4) mixed race and (5) other ethnic minority such as Chinese and Middle Eastern. Smoking status was categorized as: (1) currently smoking, (2) discontinued smoking and (3) never smoked.

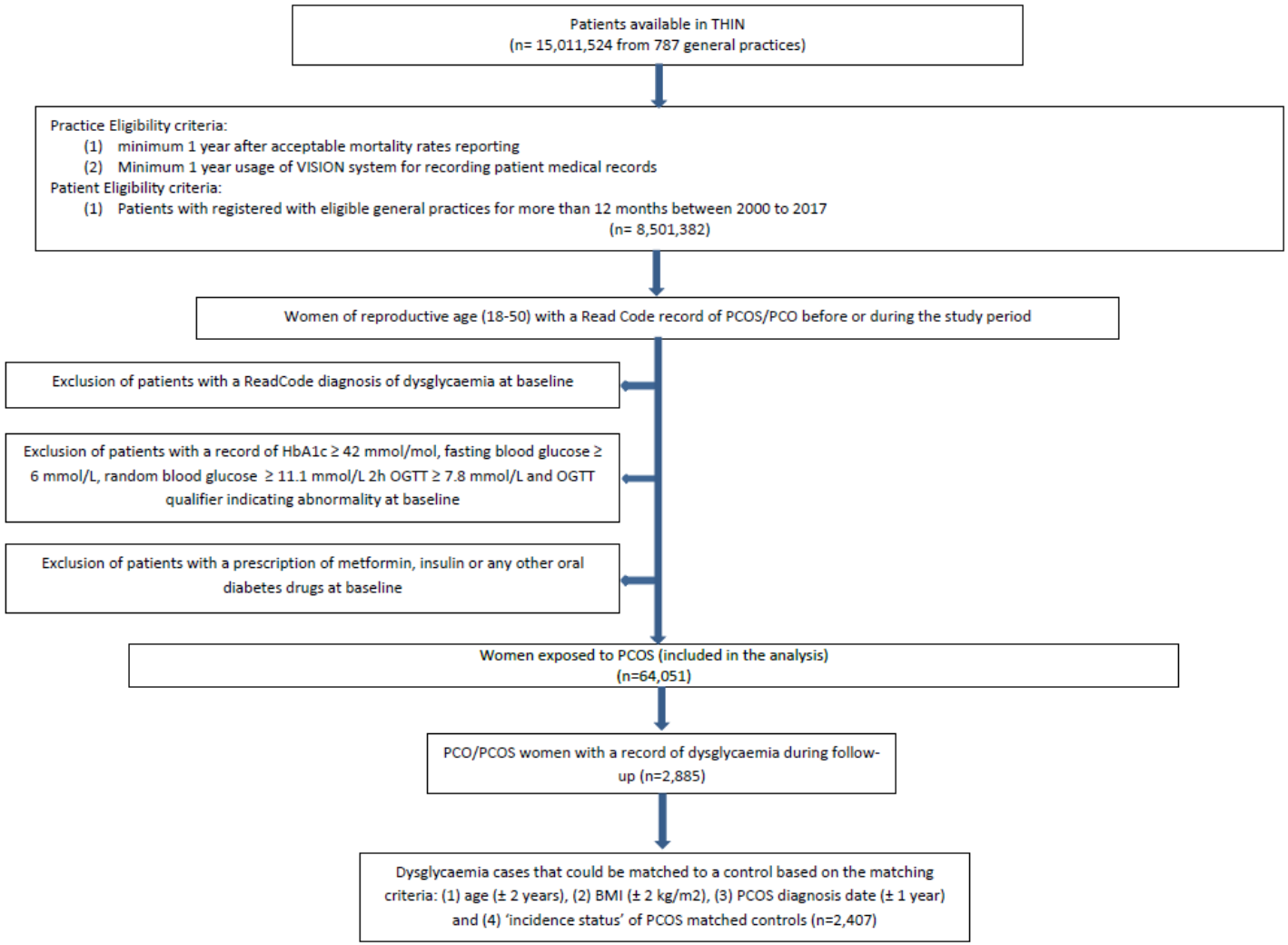
6.3.9 Statistical Analysis

Baseline characteristics of cases and controls were described using appropriate summary statistics such as mean (SD), median (IQR), and n (%). To account for matching, a conditional logistic regression model was run to estimate unadjusted and adjusted odds ratios for impaired glucose tolerance among women prescribed OCP (index exposure category) compared to those in the reference exposure category.

6.3.10 Ethical approval

Approval from the South-East Multi-centre Research Ethics Committee was obtained in 2003. Further registration and authorization for the conduct of this specific study was obtained from the relevant Scientific Review Committee (17THIN026) as per requirement.

Figure 6.1: Flowchart for cohort selection



6.4 Results

6.4.1 Characteristics of the base cohort

A total of 15 million patients were available from 787 general practices. Of these, 64,051 women were eligible for inclusion in the base cohort and had a coded diagnosis of either PCOS or PCO, without IGR at baseline. Cohort selection is described in a flowchart in **Figure 6.1**. Baseline characteristics of women included in the base cohort is presented in **Table 6.1**. Their mean age and median BMI were 30.4 (Standard Deviation, SD) 7.0 years and 25.9 (Interquartile Range, IQR 22.2-31.9) respectively. Overall, 16.0%, 4.7%, 42.8% and 16.7% of the women had missing information on BMI, smoking status, ethnicity, and socioeconomic deprivation measure, respectively. Among those with a BMI measure, over 55.6% of them had their BMI recorded in the overweight or obesity range (>25 kg/m^2). Of those with available information on the described covariate, women in the base cohort were predominantly of white ethnicity (83.5%) and non-smokers (61.1%). At baseline, 2.2% and 2.4% of them had a record of hypertension and hypothyroidism, respectively. Regarding PCOS related symptoms, 9.1%, 3.4% and 25.3% had a recording of hirsutism, hair loss and anovulation, respectively, at baseline. At least one COCP prescription was made prior to the index date in 43.3% of the women in the included cohort, of which 22.5% were prescriptions with anti-androgenic progestin component. High dose anti-androgen such as cyproterone and other anti-androgenic drugs were prescribed to 17.2% and 0.1% of the women in the base cohort, respectively. Lipid lowering drugs were prescribed to 0.6% of women in the base cohort at baseline.

Table 6.1: Baseline characteristics of women in the base cohort with a diagnosis of PCOS/PCO

	(n=64,051)
Age [Mean (SD)]	30.42 (7.04)
BMI [Median (IQR)]	25.9 (22.2-31.9)
BMI Categories	
Normal/Underweight (<25 kg/m^2)	23881 (37.28)
Overweight (25-30 kg/m^2)	12685 (19.80)
Obese (>30 kg/m^2)	17249 (26.93)
Missing	10236 (15.98)
Smoker categories	
Non-Smoker	37311 (58.25)
Discontinued Smoker	9044 (14.12)
Smoker	14674 (22.91)
Missing	3022 (4.72)
Ethnicity	
White Caucasian	30597 (47.77)
Black Afro-Caribbean	1464 (2.29)
Chinese/middle eastern/other ethnic minorities	582 (0.91)
South Asian	3085 (4.82)
Mixed Race	897 (1.40)
Missing	27426 (42.82)
Townsend	
1 (least deprived)	11270 (17.60)
2	10280 (16.05)
3	12064 (18.83)
4	11530 (18.00)
5 (most deprived)	8182 (12.77)
Missing	10725 (16.74)
Baseline comorbidities	
Hypertension	1420 (2.22)
Hypothyroidism	2172 (3.39)
PCOS related Conditions	
Hirsutism	5810 (9.07)
Hair loss	2203 (3.44)
Anovulation	16226 (25.33)
Baseline drug use	
COCP*	27768 (43.35)
COCP without anti-androgen	25481 (39.78)
COCP with anti-androgenic progestin	14437 (22.54)
High dose/isolated anti-androgen prescription	
Cyproterone	11069 (17.28)
Other anti-androgen drugs^	42 (0.07)
Lipid lowering medication	410 (0.64)

COCP* - Combined Oral Contraceptive Pills

^ flutamide/finasteride/spironolactone

Note: Women with IGR or glucose lowering drug prescription at baseline not included in the cohort

6.4.2 Selection of cases and controls

Within the base cohort, a total of 2,885 women had an incident recording of IGR during a total follow-up of 299,551 person-years, equivalent to IGR incidence rate of 96.3 per 10,000 person-years among women with PCOS. Of these 2,885 women, 2,407 women (cases) were able to be matched to 2,407 women (controls) for age (± 2 years), BMI (± 2 kg/m^2), PCOS diagnosis date (± 2 years), and incident/prevalent status of PCOS diagnosis. **Table 6.2** presents the baseline characteristics of case and control women.

6.4.3 Characteristics of cases and controls

The mean age at index, i.e., age at the date of IGR recording among the cases and the same assigned date for the matched control women, were 38.9 (SD 8.3) and 28.8 (SD 8.2) years respectively. The mean age at PCOS diagnosis among the case and control women were 28.8 (SD 6.1) and 28.8 (SD 6.0) years respectively. Mean BMI was similar between the case and control women [32.72 (SD 6.98) and 32.49 (SD 7.03) kg/m^2 respectively]. Compared to control women, cases were more like to be deprived: among cases, percentage of women with Townsend deprivation score 1 (least deprived) and 5 (most deprived) were 14.6% and 16.9%, while among control patients, it was 20.0% and 12.3% respectively. Compared to control women, cases were more likely to be smokers (26.5% vs 20.8%), from an ethnic minority background (Mixed race: 2.6% vs 0.9%; Black Afro-Caribbean: 3.3% vs 1.7%; South Asian: 10.0% vs 3.2%). Compared to control women, cases were more likely to have concurrent diagnosis of hypertension (25.9% vs 11.6%) and hypothyroidism (10.6% vs 7.8%) at baseline. Similarly, cases were more likely to have at least one prescription of metformin (17.3% vs 13.7%), lipid lowering drugs (6.2% vs 4.9%) and high dose/isolated anti-androgen drug prescription (1.7% vs 1.0%).

Table 6.2: Baseline characteristics of case and control women

Variable	Women with IGR (cases) (n=2407)	Women without IGR (controls) (n=2407)
Age at index [Mean (SD)]	38.89 (8.32)	38.80 (8.22)
Age at PCOS diagnosis [Mean (SD)]	28.84 (6.08)	28.76 (6.00)
BMI [Mean (SD)]	32.72 (6.98)	32.59 (7.03)
BMI Categories		
Normal/Underweight (<25 kg/m^2)	293 (12.17)	309 (12.83)
Overweight (25-30 kg/m^2)	448 (18.61)	444 (18.45)
Obese (>30 kg/m^2)	1290 (53.59)	1278 (53.1)
Missing	376 (15.62)	376 (15.62)
Townsend		
1 (least deprived)	351 (14.58)	481 (19.98)
2	359 (14.91)	436 (18.11)
3	473 (19.65)	457 (18.99)
4	471 (19.57)	420 (17.45)
5 (most deprived)	408 (16.95)	295 (12.26)
Missing	345 (14.33)	318 (13.21)
Smoker categories		
Non-Smoker	1306 (54.26)	1354 (56.25)
Discontinued Smoker	295 (12.26)	362 (15.04)
Smoker	639 (26.55)	501 (20.81)
Missing	167 (6.94)	190 (7.89)
Ethnicity		
White Caucasian	999 (41.5)	1099 (45.66)
Mixed Race	38 (1.58)	21 (0.87)
Chinese/middle eastern/other ethnic minorities	21 (0.87)	13 (0.54)
Black Afro-Caribbean	80 (3.32)	40 (1.66)
South Asian	241 (10.01)	77 (3.2)
Missing	1028 (42.71)	1157 (48.07)
Concurrent Conditions at baseline		
Hypertension	623 (25.88)	179 (11.59)
Hypothyroidism	256 (10.64)	188 (7.81)
Prescription of drugs within the exposure time window		
Contraceptives		
No Pill	1728 (71.79)	1592 (66.14)
COCP without anti-androgen	301 (12.51)	389 (16.16)
Anti-Androgenic progestin containing COCP*	378 (15.70)	426 (17.70)
High Dose/isolated anti-androgen prescription^	41 (1.70)	23 (0.96)
Metformin	417 (17.32)	330 (13.71)
Lipid lowering drugs	150 (6.23)	119 (4.94)

*Co-Cyprindiol/Drospirenone

^Cyproterone acetate/flutamide/finasteride/spironolactone

6.4.4 COCP and risk of IGR

Compared to women who developed IGR, a higher proportion of control women who did not develop IGR were prescribed COCP at least once during the exposure window (28.2% vs 33.9%) (**Table 6.2**). Mean number of prescriptions that were dispensed among the case and control women within the exposure window were 1.5 (SD 4.0) and 1.9 (SD 4.7) respectively. The odds of developing IGR was lower among those prescribed with COCP during the exposure window compared to those who were not prescribed with COCP [OR: 0.73, (95% CI, 0.64-0.83)]. After adjustment for age, smoking status, BMI category, ethnicity, Townsend deprivation score, base recording of hypertension and hypothyroidism, concurrent baseline prescription of metformin, lipid lowering drugs and high dose/isolates anti-androgenic drugs, there was no change in the effect estimate [aOR: 0.74, (95% CI, 0.64–0.86)] (**Table 6.3**).

Table 6.3: Association between IGR and COCP prescription (overall and according to prescription counts and type of progestin component) – Primary Analysis

	Adjusted OR (95% CI)			
Exposure				
(At least 1 COCP prescription within the exposure window; Yes/No)	0.74 (0.64-0.86)			
COCP dispensed count within the exposure window (continuous variable)		0.98 (0.96-0.99)		
COCP dispensed count within the exposure window (categorical variable)				
No pill			Ref	
Dispensed prescription count < 3			0.80 (0.67-0.96)	
Dispensed prescription count ≥ 3			0.67 (0.55-0.81)	
Contraceptives with or without anti-androgenic progestin component (categorical variable)				
No Pill				Ref
COCP without anti-androgenic progestin				0.72 (0.59-0.87)
COCP with anti-androgenic progestin*				0.76 (0.63-0.91)
Covariates				
Age	1.05 (0.99-1.12)	1.05 (0.99-1.12)	1.05 (0.99-1.12)	1.05 (0.99-1.12)
Smoker				
Non-Smoker	Ref	Ref	Ref	Ref
Discontinued Smoker	0.96 (0.79-1.16)	0.95 (0.79-1.15)	0.96 (0.79-1.16)	0.96 (0.79-1.16)
Smoker	1.47 (1.26-1.71)	1.47 (1.26-1.72)	1.47 (1.26-1.71)	1.46 (1.25-1.71)
Missing	0.93 (0.69-1.25)	0.94 (0.70-1.27)	0.94 (0.70-1.26)	0.93 (0.69-1.25)
BMI category				
Normal Weight (<25 kg/m^2)	Ref	Ref	Ref	Ref

Overweight (25-30 kg/m^2)	1.44 (0.91-2.29)	1.44 (0.91-2.28)	1.42 (0.89-2.26)	1.45 (0.91-2.31)
Obese (>30 kg/m^2)	1.71 (0.93-3.15)	1.72 (0.94-3.16)	1.68 (0.91-3.09)	1.72 (0.93-3.17)
Missing	.	.	.	.
Ethnicity				
White Caucasian	Ref	Ref	Ref	Ref
Mixed Race	2.41 (1.33-4.36)	2.28 (1.26-4.11)	2.38 (1.31-4.30)	2.40 (1.33-4.34)
Chinese/middle eastern/other ethnic minorities	1.71 (0.83-3.54)	1.77 (0.85-3.66)	1.72 (0.83-3.55)	1.71 (0.83-3.53)
Black Afro-Caribbean	2.04 (1.33-3.12)	2.07 (1.35-3.17)	2.04 (1.33-3.13)	2.03 (1.33-3.12)
South Asian	3.4 (2.51-4.59)	3.36 (2.49-4.54)	3.39 (2.51-4.59)	3.39 (2.51-4.59)
Missing	0.98 (0.85-1.12)	0.97 (0.85-1.11)	0.98 (0.85-1.12)	0.98 (0.85-1.12)
Townsend				
1 (Least Deprived)	Ref	Ref	Ref	Ref
2	1.19 (0.96-1.46)	1.18 (0.96-1.46)	1.18 (0.96-1.46)	1.19 (0.96-1.46)
3	1.40 (1.14-1.73)	1.40 (1.14-1.73)	1.41 (1.14-1.73)	1.40 (1.14-1.73)
4	1.41 (1.14-1.74)	1.42 (1.15-1.75)	1.41 (1.14-1.74)	1.41 (1.14-1.74)
5 (Most Deprived)	1.75 (1.39-2.20)	1.75 (1.39-2.20)	1.75 (1.39-2.20)	1.75 (1.39-2.20)
Missing	1.53 (1.22-1.92)	1.53 (1.22-1.91)	1.53 (1.22-1.91)	1.53 (1.22-1.92)
Concurrent diagnoses				
Hypothyroidism	1.36 (1.09-1.69)	1.36 (1.09-1.69)	1.36 (1.10-1.70)	1.35 (1.09-1.68)
Hypertension	2.97 (2.48-3.56)	2.98 (2.49-3.56)	2.96 (2.47-3.55)	2.97 (2.48-3.56)
Concurrent prescriptions				
Metformin	1.51 (1.25-1.82)	1.50 (1.24-1.81)	1.51 (1.25-1.82)	1.50 (1.24-1.81)
High dose/ isolated anti-androgen^	1.76 (0.98-3.15)	1.73 (0.97-3.11)	1.77 (0.99-3.18)	1.74 (0.97-3.13)
Lipid lowering drugs	1.22 (0.58-2.53)	1.19 (0.57-2.47)	1.22 (0.58-2.53)	1.21 (0.58-2.52)

*Co-Cyprindiol/Drospirenone

^Cyproterone acetate/flutamide/finasteride/spironolactone

Table 6.4: Association between IGR and COCP prescription (overall and according to prescription counts and type of progestin component) – Sensitivity Analysis performing risk set sampling and sampling with replacement

	Adjusted OR (95% CI)			
Exposure				
(At least 1 COCP prescription within the exposure window; Yes/No)	0.74 (0.64-0.86)			
COCP dispensed count within the exposure window (continuous variable)		0.97 (0.96-0.99)		
COCP dispensed count within the exposure window (categorical variable)				
No pill			Ref	
Dispensed prescription count < 3			0.82 (0.69-0.98)	
Dispensed prescription count ≥ 3			0.65 (0.54-0.79)	
Contraceptives with or without anti-androgenic progestin component (categorical variable)				
No Pill				Ref
COCP without anti-androgenic progestin				0.77 (0.64-0.93)
COCP with anti-androgenic progestin*				0.72 (0.60-0.86)
Covariates				
Age	1.12 (1.05-1.19)	1.12 (1.05-1.19)	1.12 (1.05-1.19)	1.12 (1.05-1.19)
Smoker				
Non-Smoker	Ref	Ref	Ref	Ref
Discontinued Smoker	0.94 (0.79-1.13)	0.94 (0.78-1.12)	0.94 (0.79-1.13)	0.94 (0.79-1.13)
Smoker	1.30 (1.12-1.51)	1.31 (1.13-1.52)	1.30 (1.12-1.51)	1.30 (1.12-1.51)
Missing	0.97 (0.73-1.3)	0.99 (0.74-1.33)	0.98 (0.73-1.31)	0.97 (0.73-1.30)
BMI category				
Normal Weight (<25 kg/m^2)	Ref	Ref	Ref	Ref

Overweight (25-30 kg/m^2)	1.64 (0.99-2.72)	1.67 (1.01-2.77)	1.63 (0.98-2.7)	1.64 (0.99-2.72)
Obese (>30 kg/m^2)	1.80 (0.95-3.42)	1.82 (0.96-3.45)	1.77 (0.93-3.37)	1.79 (0.95-3.41)
Missing	.	.	.	.
Ethnicity				
White Caucasian	Ref	Ref	Ref	Ref
Mixed Race	1.89 (1.08-3.31)	1.86 (1.06-3.26)	1.89 (1.08-3.31)	1.90 (1.08-3.33)
Chinese/middle eastern/other ethnic minorities	1.86 (0.91-3.80)	1.92 (0.94-3.92)	1.85 (0.90-3.79)	1.85 (0.91-3.79)
Black Afro-Caribbean	1.92 (1.28-2.88)	1.95 (1.30-2.91)	1.93 (1.29-2.88)	1.93 (1.29-2.88)
South Asian	4.00 (2.92-5.47)	4.03 (2.95-5.52)	3.99 (2.91-5.46)	4.01 (2.93-5.49)
Missing	1.04 (0.92-1.19)	1.05 (0.92-1.20)	1.04 (0.91-1.18)	1.04 (0.91-1.19)
Townsend				
1 (Least Deprived)	Ref	Ref	Ref	Ref
2	1.17 (0.96-1.44)	1.18 (0.96-1.45)	1.18 (0.96-1.45)	1.18 (0.96-1.44)
3	1.19 (0.98-1.44)	1.20 (0.98-1.45)	1.20 (0.98-1.46)	1.19 (0.98-1.44)
4	1.35 (1.10-1.65)	1.35 (1.10-1.65)	1.35 (1.11-1.66)	1.35 (1.10-1.65)
5 (Most Deprived)	1.56 (1.25-1.93)	1.55 (1.25-1.92)	1.56 (1.25-1.93)	1.56 (1.26-1.93)
Missing	1.23 (0.99-1.53)	1.24 (1.00-1.54)	1.23 (0.99-1.53)	1.23 (0.99-1.53)
Concurrent diagnoses				
Hypothyroidism	1.14 (0.93-1.39)	1.13 (0.93-1.38)	1.14 (0.94-1.40)	1.14 (0.93-1.39)
Hypertension	2.46 (2.08-2.91)	2.47 (2.09-2.91)	2.46 (2.08-2.91)	2.47 (2.09-2.91)
Concurrent prescriptions				
Metformin	1.46 (1.21-1.75)	1.46 (1.21-1.75)	1.45 (1.21-1.74)	1.46 (1.22-1.76)
High dose/ isolated anti-androgen^	1.84 (1.06-3.18)	1.80 (1.04-3.12)	1.83 (1.06-3.16)	1.86 (1.07-3.22)
Lipid lowering drugs	1.80 (0.79-4.13)	1.77 (0.77-4.07)	1.80 (0.78-4.14)	1.81 (0.79-4.15)

*Co-Cyprindiol/Drospirenone

^Cyproterone acetate/flutamide/finasteride/spironolactone

When considering the exposure COCP as a continuous variable, there was a 2% reduction in the odds of developing IGR with every unit increase in COCP prescription count [aOR: 0.98 (95% CI, 0.96-0.99)] (**Table 6.3**). Considering the categorized prescription count variable, women with a prescription count <3 and ≥3 records of COCP within the exposure window were at 20% and 33% reduced odds of developing IGR compared to women with no prescription of COCP within the exposure window, after adjustment for the covariates considered [aOR: 0.80 (95% CI, 0.67-0.96) and 0.67 (95% CI, 0.55-0.81), respectively] (**Table 6.3**).

Women prescribed with COCP with and without anti-androgenic progestin component had a 24% and 28% reduction in the odds of development of IGR compared to women without a prescription of COCP [aOR: 0.72 (95% CI, 0.59-0.87) and 0.76 (95% CI, 0.63-0.91) respectively] (**Table 6.3**).

The results of the sensitivity analysis where risk set sampling was performed did not differ significantly from the primary analysis (**Table 6.4**).

6.4.5 Other observed risk factors for IGR

Among the covariates considered, women who were smokers [aOR: 1.47 (95% CI, 1.26-1.72), from ethnic minority backgrounds such as mixed race, black Afro-Caribbean, and South Asian [aOR: 2.41 (95% CI, 1.33-4.36), 2.04 (95% CI, 1.33-3.12), and 2.40 (95% CI, 2.51-4.59)], or from a deprived background were at a higher odds of development of IGR compared to their respective reference population groups of non-smokers, white Caucasians, and least deprived women. Women with a concurrent diagnosis of hypothyroidism or hypertension were at an increased risk of development of IGR [aOR: 1.39 (95% CI, 1.09-1.69) and 2.97 (95% CI, 2.48-3.56), respectively]. Women with a concurrent prescription of metformin and high dose/isolated anti-androgen prescription were also observed to have higher odds of

development of IGR [aOR: 1.51 (95% CI, 1.25-1.82) and 1.76 (95% CI, 0.98-3.25) respectively].

6.5 Discussion

6.5.1 Summary of findings

In this retrospective nested case control study, the use of COCP was associated with a reduction in development of IGR among women with PCOS. Furthermore, a dose-response relationship was observed for the number of COCP prescriptions issued within the exposure window that was matched between case and control women (with and without IGR). Finally, no difference in the odds of development of IGR was found between women prescribed COCPs with and without anti-androgenic progestin component. These findings remained in a sensitivity analysis where risk set sampling was performed for selection of control women.

6.5.2 Comparison with existing literature

To the best of my knowledge, this is the first study estimating the effect of COCP use on the risk of development of IGR among women with PCOS. However, small-scale trials and cohort studies have been conducted to look at the efficacy and effectiveness of COCP use on development of type 2 diabetes and change in glucose metabolism such as change in fasting glucose and homeostasis model assessment-estimated insulin resistance (HOMA-IR).

In a 2016 Korean population study of 6,554 postmenopausal women, investigators found that those who took COCP during their reproductive years for >6 months had a 37% increased risk of type 2 diabetes.[186] However, in a more recent study with examination of the National Health and Nutrition Examination Survey (NHANES) database between 2007 and 2018, it was found that COCP use in >6,000 women aged 35–50 years who met matching criteria was associated with a 29% reduced risk of type 2 diabetes compared with the risk in never users.[187] For the first cohort of the Nurses' Health Study, 2,276 healthy women were followed for a median of 12 years from 1976, with findings that risk of type 2 diabetes was increased by 10% in women

with previous COCP use compared with those who never took the medication;[188] however, these data reflect the use of older COCP preparations with higher ethinylestradiol concentrations between the 1970s and 1990s.

A systematic review by Halperin et. al., pooled together findings from 29 and 11 cohort studies with a combined sample size of 584 and 265 women with PCOS, and found that there was no association between oral contraceptives and either of the IGR related proxy measures - fasting glucose and HOMA-IR, respectively.[189] However, meta regression analysis conducted within the review suggested BMI could be a confounder in the association between oral contraceptives and risk of IGR.

Another narrative review by Medeiros looked at the association between COCPs and change in fasting glucose and three-hour OGTT response among women with PCOS.[75] The primary RCTs included in the review reported contradictory findings for COCP use ranging from a null effect on fasting glucose,[190–192] to a non-significant reduction[193] or a significant 8.0% reduction in fasting glucose[194] among the various trials. Similarly, this review reported contradictory findings for the effect of COCP on HOMA-IR among the included studies, wherein one study reported a significant increase in HOMA-IR,[195] while another reported a 31% decrease in HOMA-IR.[196]

The studies included in both the reviews are limited by their sample size, shorter follow-up, and unaccounted confounders such as BMI, and are methodologically heterogeneous in terms of the PCOS exposure ascertainment criteria and the COCP preparation used.

6.5.3 Biological mechanism

When analysing the effect of COCP stratified by the type of progestin component in the preparation, women with PCOS and COCP use had a similarly reduced risk of developing IGR when exposed to COCPs with and without antiandrogenic progestin components, suggesting that the oestrogen-induced increase in SHBG may be the primary driver of the risk-mitigating

effect. However, this finding is potentially limited by the lower number of women receiving antiandrogenic COCPs. Cyproterone acetate and drospirenone are progestins with antiandrogenic properties, as opposed to progestins such as desogestrel or levonorgestrel, which have neutral or proandrogenic effects.[197] While cyproterone acetate and drospirenone exert antiandrogen activity via androgen receptor blockade, their antiandrogen activity is considerably less than that of recently approved novel antiandrogens mainly used in the treatment of prostate cancer.[198]

6.5.4 Strengths and limitations

This study has many strengths including novelty, large sample size, and appropriate study design suitable for pharmacepidemiological studies using real world data. However, this study also has several limitations such as potential prescription-by-indication and misclassification bias.

Women using metformin and single-agent antiandrogen therapy had an increased risk of incident IGR in this study. This is very likely reflective of a confounding-by-indication bias. Accordingly, the women with PCOS at highest risk of IGR based on metabolic or androgen phenotype may have been systematically prescribed metformin and single agent antiandrogen therapy. It is possible that the observation of reduced IGR risk in women with PCOS on COCPs may also reflect a prescription-by-indication bias, whereby those women with cardiovascular risk factors such as obesity, dyslipidaemia, and hypertension were contraindicated for COCP prescription.

In defining the base cohort, Read code records of PCO was considered in addition to PCOS, which may have included a subset of women who do not have PCOS, and thereby misclassifying their exposure status.

6.5.5 Conclusion

This study suggests that COCP has a protective effect among women with PCOS against development of IGR. Similarity in the effect estimates for prescription of COCPs with and without anti-androgenic progestin component suggests a biological mechanism underpinned by the oestrogen component of COCP, increasing SHBG and alleviating androgen excess. Large scale randomised controlled trials with sufficient follow-up are needed to evaluate the long-term protective effect of COCP among women with various PCOS phenotypes.

Chapter 7 - Discussion

In this doctoral thesis, I aimed to (1) estimate time trends in the prevalence and incidence rate of polycystic ovary syndrome (PCOS) based on both diagnosed and probable cases of PCOS supported by relevant symptoms recorded within primary care; (2) estimate time trends in the prevalence and incidence rate of type 2 diabetes and impaired glucose regulation (IGR) among women with PCOS, and time trends in the incidence of gestational diabetes mellitus (GDM) among pregnant women with PCOS using primary care data and its derived pregnancy register data respectively; (3) estimate the odds of obstetric outcomes including preterm birth, high and low birthweight, stillbirth and mode of delivery among women with PCOS compared to an age matched cohort of women without PCOS using primary and secondary care linked data; (4) estimate the incidence of confirmed or suspected COVID-19 diagnosis during the first wave of the pandemic among women with PCOS compared to an age matched cohort of women without PCOS and; (5) estimate the odds of developing IGR among women who were prescribed combined oral contraceptive pills (COCPs) compared to women who were not prescribed COCPs in a base cohort of women with PCOS using a nested case control study designed within primary care data. In this discussion chapter, I aim to summarize the findings, highlight the strengths and limitations of the research methods and databases employed to capture the findings, discuss the implications of the findings, and propose future research recommendations.

7.1 Summary of findings

Prevalence of PCOS has been rising over the last two and a half decades (between 1995 and 2020). Between 1995 and 2015, there was a widening gap between prevalence estimates from primary care records based on diagnostic codes alone, and prevalence estimates based on diagnostic codes and a combination of symptom codes, indicating an increasing level of missed PCOS diagnosis within primary care. Since 2015, the gap in prevalence estimates persisted, although it remained stable rather than continuing to widen. The incidence rate of PCOS rose

particularly sharply in 2004, coinciding with the year of implementation of the Rotterdam criteria for PCOS diagnosis, and introduction of the Quality and Outcomes Framework which led to improved recording of chronic conditions within UK primary care. Since 2014, there has been a decline in the incidence rate of PCOS, also reflected by the stabilization of prevalence estimates of PCOS from 2014 onwards.

When stratifying the overall prevalence and incidence rate estimates of PCOS by ethnicity, women from the South Asian community, followed by Black Afro-Caribbean women was found to have the highest prevalence and incidence of PCOS. Over the years of the study, the ethnic differences in the incidence and prevalence estimates of PCOS have widened across the ethnic subgroups.

In addition to the increasing prevalence of PCOS in primary care, there has been an increase in the incidence rate of type 2 diabetes and impaired glucose regulation (IGR), and in the incidence of gestational diabetes mellitus among women and pregnant women with PCOS, respectively. The overall incidence estimates of the three glucose intolerance outcomes are higher among women with a diagnostic code for PCOS in comparison to women with a combination of symptom codes indicating PCOS diagnosis, suggesting that diagnosis of PCOS is likely to be more frequently coded in women with a more severe metabolic phenotype. Also, a sharp rise was found in the incidence of type 2 diabetes and impaired glucose regulation in the years 2012 and 2017, coinciding with the publication of a public health guideline (PH28) recommending active screening of patients at high risk of type 2 diabetes, including women with a pre-existing diagnosis of PCOS, and the nationwide launch of diabetes prevention week raising awareness for the uptake of free NHS health checks, respectively. This suggests a positive impact of the guideline and diabetes prevention programme on potentially early diagnosis of impaired glucose regulation and type 2 diabetes, especially among women with PCOS.

Considering the intergenerational effects of PCOS, a cohort study linking primary and secondary care data was conducted to observe the obstetric outcomes of pregnant women with PCOS in comparison to an age matched cohort of pregnant women without PCOS. In this study, pregnant women with PCOS were observed to be at a higher risk of delivering preterm and of having a caesarean section even after accounting for several confounders. Notably, pregnant women with PCOS were observed to be at a higher risk of delivering babies weighing less than 2.5 kgs in the unadjusted analysis, however after adjustment for gestational age, the effect estimate attenuated close to null [adjusted OR: 1.03 (0.95-1.13)] and became statistically insignificant. Similarly, pregnant women with PCOS were observed to be at a higher risk of delivering babies that were large for gestational age in the unadjusted analysis, however after adjustment for maternal pre-gravid BMI, the effect estimate attenuated to null [adjusted OR: 1.00 (0.97-1.04)], suggesting a mediating and confounding effect by gestational age at delivery and maternal pre-gravid BMI, respectively, on the association between maternal PCOS and baby's birthweight.

In the next chapter of this doctoral thesis, considering the increased incidence and prevalence of type 2 diabetes, and many other chronic health conditions such as non-alcoholic fatty liver disease (NAFLD) and obstructive sleep apnoea (OSA) among women with PCOS, the global outbreak of SARS CoV-2 created concern regarding susceptibility to infection among women with PCOS, especially due to the frequent concurrence of the aforementioned chronic conditions and other risk factors for PCOS such as being from ethnic minority, overlapping with the risk factors of COVID-19 severity. In line with this, a retrospective study was conducted during the first wave of the pandemic. In this study, a 52% increased risk of confirmed or suspected COVID-19 was observed among women with PCOS in comparison to an age matched cohort of women without PCOS, which attenuated to a 28% increase in risk after adjusting for available confounding variables such as age, body mass index, impaired

glucose regulation, androgen excess, anovulation, vitamin D deficiency, hypertension, and cardiovascular diseases. This emphasised the significance of timely prioritization of women with PCOS for vaccine uptake and the need to encourage women with PCOS to adhere to infection control measures such as social distancing and wearing masks.

Finally, considering PCOS as a significant contributor to the overall burden of type 2 diabetes development among women, and combined oral contraceptive pills (COCP) as a routinely prescribed medication among women with PCOS, I aimed to explore the effect of COCP on the development of impaired glucose regulation among women with PCOS using real world data. In a nested case control study, it was observed that COCP conferred a protective effect against development of type 2 diabetes among women with PCOS. A dose response relationship of increasing COCP prescription counts with that of decreasing odds of developing IGR was also observed. Finally, I aimed to see if variation in the progestin component of the COCP, in terms of their androgenic effect, conferred varying protective effect on the odds of developing IGR. However, a similar protective effect of COCPs with and without anti-androgenic progestin component was observed, suggesting that the protective effect may be attributable to the oestrogen component of COCP. This study was not a randomised clinical trial, hence a possible influence of a prescription by indication bias cannot be excluded, in this case by prescription of an OCP with anti-androgenic progestin component in particular to women with severe PCOS.

7.2 Comparison with literature

Previous studies have already established the apparent under-recording and delayed diagnosis of PCOS within primary care[9,199] similar to the observations made in Chapter 2 of this doctoral thesis. There are also similarities between the trend in the incidence of PCOS observed in this thesis and the findings by Ding et. al.,[9] reporting a slightly increasing incidence of PCOS between the years 2004 and 2014. However, this thesis (Chapter 2) explores, for the very first

time, the trend in incidence and prevalence estimates of PCOS both before and after the establishment of the Rotterdam criteria in the year 2004, highlighting the significant impact of this diagnostic criteria on PCOS recording within primary care in the UK.

Notably, several studies have also assessed ethnic differences in the severity and symptomatic presentation of PCOS and have reported a more severe presentation and an earlier diagnosis of PCOS among the South Asian community.[80–82] Systematic reviews have also pooled together prevalence estimates of PCOS from various countries with distinguished ethnic predominance and have identified higher prevalence estimates in the region of Latin America and South East Asia wherein the population is predominantly of Hispanic and South Asian ethnicity.[83,200] However, methodological differences across the studies included in the review limits comparison of prevalence estimates between ethnic groups obtained from a comparable background population. In this thesis (Chapter 2), the widening difference in the prevalence of PCOS by ethnic subgroups over time is highlighted. In the year 2019, while South Asian women and Black Afro-Caribbean women represent roughly 4% and 8% of the population with a record of ethnicity within the UK primary care database, nearly 7% and 10% of the PCOS diagnoses made within the database are in women from South Asian and Black Afro-Caribbean ethnicity, respectively.

The increased incident risk of type 2 diabetes, IGR and GDM among women with PCOS compared to women without PCOS have been established previously.[28–30] However, this thesis (Chapter 3) explores, for the very first time, time trends in the incident diagnosis of type 2 diabetes, IGR and GDM among women with PCOS.

Anovulatory infertility is one of the defining features of PCOS among women with a certain PCOS phenotype, representing up to 80% of women with PCOS.[201] In the UK, assisted reproduction is offered to women with PCOS if they are aged < 40 and have been trying to conceive for two years or more or if they are aged < 42, have been trying to conceive for two

years or more, show no evidence of low ovarian reserve, and they have never before had in-vitro fertilization (IVF) treatment.[202] With assisted reproduction, it has been reported, albeit using outdated data and diagnostic criteria for PCOS, that pregnancy can be achieved among 84% of the women who have PCOS related anovulatory infertility.[203] More recently, using Swedish data, it has been reported that 13.7% of singleton births with maternal PCOS underwent assisted reproduction. [60]

With treatment advances to improve fertility,[55–57,112,204] several research groups have sought to explore the risk of pregnancy outcomes among women with PCOS, and several systematic reviews have been conducted to pool together the findings.[60] While the pooled estimates across all these reviews are in agreement and suggest a significant increase in the risk of adverse obstetric outcomes such as preterm birth and caesarean section, there is still contradicting evidence on the association between maternal PCOS and birthweight. While the review by Yu et. al meta-analysed effect estimates from 11 and 10 studies and reported a non-significant 14% [RR: 1.14 (95% CI, 0.93,1.39)] and 45% [RR: 1.45 (0.96,2.20)] increase in the risk of LGA and SGA, respectively, among women with PCOS,[112] the review by Kjerulff et. al. did not pool together findings from the primary studies they included owing to methodological heterogeneity among them.[55] The review by Qin et. al. pooled together findings from 17 primary studies and suggested a 11g lower birthweight among babies born to mothers with PCOS [Weighted mean difference: -0.11 (95% CI, -0.19, -0.03)].[57] In this doctoral thesis (Chapter 4), compelling evidence on the association between maternal PCOS and obstetric outcomes including low and high birthweight, LGA and SGA was obtained by performing a methodologically rigorous epidemiological study using contemporary, large and representative data from the UK.

During my time as a doctoral candidate, the SARS CoV-2 outbreak originating in the city of Wuhan in China (in November 2019) evolved into a global pandemic creating a considerable

death toll, anxiety, and isolation, including within the UK. In July 2020, Kyrou et. al., published a review highlighting the overlapping features of PCOS and risk factors for COVID-19 severity such as obesity, hypertension, type 2 diabetes, metabolic syndrome, belonging to an ethnic minority, high cytokine and androgen levels, and low vitamin D levels.[70] The review highlighted the importance of assessing and informing women with PCOS of their potential risks during the pandemic. Following on from the review, this research question was addressed for the first time in this doctoral thesis (Chapter 5) and women with PCOS were identified as a cohort of women with higher susceptibility to infection using real world data. Concurrent literature has also highlighted the increased risk of COVID-19 related morbidity among women with PCOS, as well as the isolation, anxiety and uncertainty experienced by women with PCOS related to limited access to healthcare services during the pandemic to support and manage their condition.[205] Study by Kite et. al., has also explored the adverse impact of the pandemic on sleep, quality of life and overall mental well-being of women with PCOS. [206]

Anovulatory infertility being the defining feature of PCOS for predominant of women with PCOS, several treatment advances have been made to treat anovulatory infertility. However, there has been little to no advances to improve the adverse metabolic outcomes of women with PCOS. NICE guidance recommends combined oral contraceptive pills as standard treatment for women with PCOS to ameliorate hirsutism and acne, and to attain menstrual regularity.[93] Several studies have suggested an increased risk of endometrial cancer among women with PCOS,[207] and up to 30% reduction in the risk of endometrial cancer has been reported with COCP use.[208] In line with the multifactorial benefits of COCP for women with PCOS, and the considerable apprehension for COCP use due to stigma and misinformation, several investigators have highlighted the need to study the effect of COCP on glucose regulation among women with PCOS.[209] While the study by Adenjii et. al. estimated Homeostasis Model Assessment – Insulin Sensitivity Index (HOMA-ISI), fasting insulin and fasting glucose before

and after COCP use among women with and without PCOS and reported worsening of glucose tolerance after COCP use among women with PCOS compared to women without PCOS,[177] there has been no study that has assessed the effectiveness of COCP use in a base cohort of women with PCOS with an appropriate comparator group of women without COCP use. In this doctoral thesis (Chapter 6), for the first time, the protective effect of COCP use on the incidence of impaired glucose regulation was assessed among women with PCOS.

7.3 Implications to practice

The magnitude of missed PCOS diagnosis within primary care relates to the magnitude of the missed opportunistic period for enacting early interventions to reduce the burden of PCOS. Notably, the period of reducing incidence of PCOS post-2014 coincides with a period of decreasing face-to-face consultations by GPs within primary care.[210] In the coming years post pandemic, remote consultations are likely to become the norm, and according to the report "Getting the best out of remote consulting in general practice: Practical challenges and policy opportunities" by Nuffield Trust, GPs have expressed concerns over delayed or missed diagnoses from remote consultations owing to their inability to pick up on visual clues from face-to-face physical examination and 'doorknob' concerns.[211] Remote consultations can further widen inequalities due to differential access to care, especially reduced access among the vulnerable communities, such as ethnic minority communities, among whom we see a higher incidence of PCOS despite their lower engagement with healthcare services. Considering the foreseeable shift in the way primary care consultations are likely to be in the near future, continuous evaluations of the effect of remote consultations is essential to avoid incidental exclusion of vulnerable patients such as ethnic minority women who are at a higher risk of PCOS diagnosis. Patients including women with PCOS who are end users of remote digital services should be consulted in co-designing and developing the existing remote

healthcare technology to enhance patient experience. Remote healthcare services should complement rather than replace face-to-face services.

Considering the increased odds of adverse obstetric outcomes among women with PCOS, and the potential confounding by pre-gravid BMI for some of the outcomes such as babies born LGA, weight loss interventions such as lifestyle modifications during the preconception care period may be beneficial to reduce the burden of adverse obstetric outcomes. Obstetric risks should be clearly conveyed alongside the benefits of dietary interventions by healthcare professionals prior to and during pregnancy of women with PCOS.

Finally, considering the stigma associated with the use of combined oral contraceptive pills, and claims of their adverse effects without factual evidence, global outreach programmes are essential to provide clear information on advice around the safety profile of prescribing oral contraceptive pills for women with PCOS.

7.4 Strengths and limitations

Data used for the analyses performed in this doctoral thesis have several strengths including the large sample size and its generalizability to the UK population. In Chapter 2 of this thesis, data was obtained for a period of two and half decades between 1995 and 2020, to observe the long-term changing patterns of PCOS burden within primary care based on its incidence and prevalence trends during this period. The large sample size also aided stratified observation and provided meaningful estimates of the incidence and prevalence trends of PCOS by ethnic subgroups. While the database provided strength in terms of size, there were also limitations associated with the data such as the retrospective, routine, and passive nature of data collection and thereby under-recording of PCOS and the inability to provide an outlook on the true incidence and prevalence estimates for PCOS.

Missing data on ethnicity within primary care, a variable by which the incidence estimates were stratified, and missing data on outcomes (baby's birthweight, baby's sex, and gestational age

within secondary care examined in Chapter 4 of this thesis and suspected or confirmed COVID-19 within primary care examined in Chapter 5 of this thesis) were some of the major limitations inherent to the databases utilized for this thesis.

Furthermore, due to the observational nature of the studies looking at the association between PCOS and incident diagnosis of adverse obstetric outcomes (in Chapter 4) and incident diagnosis of suspected/confirmed COVID-19 (in Chapter 5), residual and unmeasured confounding may have biased the observed incidence estimates. Furthermore, bias due to systematic difference in surveillance between women with and without PCOS, wherein women with PCOS make more frequent visits to primary care, may have resulted in an increased voluntary reporting of outcomes such as suspected COVID-19 to general practitioners by women with PCOS. Confounding in this study was mitigated by the step-by-step adjustment for available covariates.

Chapter 6 of this thesis employed a pharmacoepidemiological study design to estimate the effectiveness of COCP on the incidence of impaired glucose regulation among women with PCOS. However, this study employed real world data and therefore may have been affected by confounding-by-indication bias (i.e., lower prescription rates of COCP among women with high BMI, who are contra-indicated for COCP prescription yet at a higher risk of developing glucose intolerance). However, confounding was mitigated by the adjustment for crucial variables such as high BMI.

Despite the limitations, a major strength of several aspects of this doctoral thesis, especially Chapters 2, 5 and 6, is the novelty of the research questions examined in these chapters. Future research is needed to confirm the observations made in these chapters.

7.5 Future investigations

7.5.1 Ethnic barriers to care and management of PCOS

Women from ethnic minority communities such as the South Asian and Black Afro-Caribbean communities are at a higher risk of developing PCOS, as established within Chapter 2 of this doctoral thesis. With the recognised disparity in healthcare access and health inequalities between White Caucasian and ethnic minority communities in the UK and the now established increased risk of PCOS diagnosis, it is imperative to both quantitatively and qualitatively explore the ethnic variation in the management of women with PCOS. A previous study by Hillman et. al. has already established that women from ethnic minority communities are less likely to discuss their PCOS symptoms including infertility and associated mental illness with their GP. Further investigations exploring barriers of care and management of PCOS among women from ethnic minority communities is necessary to curb the widening of health inequalities. Furthermore, an increasing time trend in the incidence of type 2 diabetes among women with PCOS is established within Chapter 3 of this doctoral thesis. Primary care-based cohort studies within the UK exploring the ethnic difference in the incidence of type 2 diabetes and other adverse metabolic outcomes such as OSA, hypertension and NAFLD may provide a valuable evidence base that can drive policy implementations to improve barriers of care for ethnic minority women with PCOS in the UK.

7.5.2 Exploring the prognostic value of comorbidity clusters and diagnostic phenotype on the development of adverse obstetric outcomes among women with PCOS

Within Chapter 4, the increased risk of obstetric complications among women with PCOS compared to women without PCOS was established. Within this, using step-by-step inclusion of covariates, the effect of confounding conferred by a range of risk factors including demographic features, hallmark features of PCOS, a limited number of concurrent

comorbidities such as hypertension and thyroid disorders, and pregnancy related variables was explored. While the study was limited to common comorbid conditions observed among women with PCOS as potential confounders, the wide range of prognostic comorbidities and their implications on adverse obstetric outcomes among women with PCOS remains controversial and yet to be explored.[212]

Recent population-based studies have reported a high morbidity burden among women with PCOS,[43] although a systematic understanding of what morbidities are predominantly presented among women with PCOS and how they cluster together is not well understood. It is possible that women with different PCOS phenotypes prognostically develop different co-morbidities and present with different obstetric risks. By employing clustering algorithms such as K-mode analysis, latent class analysis and Gaussian mixture models, clusters of women with PCOS presenting with different comorbidity patterns can be established and their obstetric risks can be explored.

Furthermore, PCOS is a heterogenous endocrine condition and previous studies have established 4 different sub-phenotypes of PCOS based on their symptom presentation – type A being the complete PCOS phenotype wherein a woman presents with hyperandrogenism, ovulatory dysfunction, and polycystic ovarian morphology on ultrasound (HA + OD + PCOM); type B being the classic PCOS phenotype wherein a woman presents with hyperandrogenism and ovulatory dysfunction (HA + OD); type C being the ovulatory PCOS phenotype wherein a woman presents with hyperandrogenism and polycystic ovarian morphology (HA + PCOM) and type D being a non-hyperandrogenic PCOS phenotype wherein a woman presents with ovulatory dysfunction and polycystic ovarian morphology on ultrasound (OD + PCOM).[213] The diagnostic phenotype may be a prognostic parameter for the development of obstetric outcomes, and this remains to be explored.

Finally, triangulation of the comorbidity clusters and the PCOS phenotypes with that of androgens and androgen precursors and metabolites may provide crucial understanding of the pathophysiology of this heterogeneous transgenerational condition.

7.5.3 Exploring the risk of long COVID among women with PCOS

Results from Chapter 5 of this doctoral thesis suggests that women with PCOS are more susceptible to a diagnosis of confirmed or suspected COVID-19 compared to an age matched cohort of women without PCOS. With recent evidence regarding the longer-term effects of COVID-19, referred to as 'post-acute COVID-19 syndrome (PASC)', or 'long COVID', it is of value to explore the association between PCOS and long COVID. Establishing an association between PCOS and long COVID maybe valuable in providing clues to the unknown underlying pathophysiology of long COVID and to create awareness of the condition within the at-risk community of women with PCOS in the, as of now, ongoing COVID-19 epidemic.

7.5.4 Randomized controlled trial exploring the efficacy of COCP among women with PCOS for prevention of impaired glucose regulation

Evidence from Chapter 6 of this doctoral thesis suggests a potential protective effect of COCP among women with PCOS against development of impaired glucose regulation. However, due to the observational nature of the design, the study is limited by potential confounding. A pilot trial exploring the efficacy of COCP with a randomized controlled trial environment will provide valuable and conclusive evidence for this research question.

7.5.5 Real world data-based signal detection and drug re-purposing analysis

Currently, limited pharmacological therapies are available for women with PCOS; those in use include metformin, oral contraceptives, and high-dose anti-androgenic medications to manage symptoms related to hyperandrogenism such as hirsutism. Considering the limited therapeutic advancements for the management of PCOS, and for the prevention of long-term metabolic outcomes among women with PCOS, future research may employ real world data to detect

protective signals of drugs against adverse outcomes such as development of type 2 diabetes. Considering the high morbidity burden of women with PCOS, it is likely that women with PCOS are prescribed with drugs for the management of these conditions, and the effect of these drugs on PCOS-related outcomes may provide insights for therapeutic advancements.

Furthermore, the effectiveness of pharmacological agents might differ across the PCOS sub-phenotypes. Therefore, a subgroup analysis of this exploratory signal detection exercise may provide further valuable information.

7.6 Conclusion

PCOS is a heterogeneous endocrine condition more commonly diagnosed among the ethnic minority communities in the UK, conferring an increased risk of metabolic conditions, infections, and obstetric complications. Treatment for PCOS is limited, and COCP may confer a protective effect against development of type 2 diabetes. Future investigations exploring (1) barriers of care among ethnic minority communities, (2) prognostic drivers of obstetric outcomes, (3) association with the incidence of long COVID and (4) signal detection for repurposing drugs will benefit the care of women with PCOS.

References

1. Hillman SC, Dale J. Polycystic ovarian syndrome: an under-recognised problem? 2018;68(670). doi:10.3399/BJGP18X696101
2. Deswal R, Narwal V, Dang A, Pundir CS. The Prevalence of Polycystic Ovary Syndrome: A Brief Systematic Review. J Hum Reprod Sci. 2020;13(4):261. doi:10.4103/JHRS.JHRS_95_18
3. Teede HJ, Misso ML, Costello MF, Dokras A, Laven J, Moran L, Piltonen T, Norman RJ, Andersen M, Azziz R, Balen A, Baye E, Boyle J, Brennan L, Broekmans F, Dabadghao P, Devoto L, Dewailly D, Downes L, et al. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. Published online 2018. Accessed May 27, 2021. https://pubmed.ncbi.nlm.nih.gov/30033227/
4. Asunción M, Calvo RM, San Millán JL, Sancho J, Avila S, Escobar-Morreale HF. A prospective study of the prevalence of the polycystic ovary syndrome in unselected Caucasian women from Spain. J Clin Endocrinol Metab. 2000;85(7):2434-2438. doi:10.1210/JCEM.85.7.6682
5. Knochenhauer ES, Key TJ, Kahsar-Miller M, Waggoner W, Boots LR, Azziz R. Prevalence of the Polycystic Ovary Syndrome in Unselected Black and White Women of the Southeastern United States: A Prospective Study 1 . J Clin Endocrinol Metab. 1998;83(9):3078-3082. doi:10.1210/jcem.83.9.5090
6. Bozdag G, Mumusoglu S, Zengin D, Karabulut E, Yildiz BO. The prevalence and phenotypic features of polycystic ovary syndrome: A systematic review and meta-Analysis. Hum Reprod. 2016;31(12):2841-2855. doi:10.1093/humrep/dew218
7. Azziz R, Carmina E, Chen Z, Dunaif A, Laven JSE, Legro RS, Lizneva D, Natterson-Horowtiz B, Teede HJ, Yildiz BO. Polycystic ovary syndrome. Nat Rev Dis Prim.

2016;2. doi:10.1038/nrdp.2016.57

8. Fernandez RC, Moore VM, Rumbold AR, Whitrow MJ, Avery JC, Davies MJ. Diagnosis delayed: health profile differences between women with undiagnosed polycystic ovary syndrome and those with a clinical diagnosis by age 35 years. Hum Reprod. 2021;36(8):2275-2284. doi:10.1093/HUMREP/DEAB101
9. Ding T, Baio G, Hardiman PJ, Petersen I, Sammon C. Diagnosis and management of polycystic ovary syndrome in the UK (2004-2014): a retrospective cohort study. BMJ Open. 2016;6(7):e012461. doi:10.1136/bmjopen-2016-012461
10. Lujan ME, Chizen DR, Pierson RA. Diagnostic Criteria for Polycystic Ovary Syndrome: Pitfalls and Controversies. J Obstet Gynaecol Can. 2008;30(8):671. doi:10.1016/S1701-2163(16)32915-2
11. Dunaif A, Given JR, Haseltine FP. Diagnostic criteria for Polycystic ovary syndrome: towards a rational apporoach. polycystic ovary Syndr Bost. Published online 1992:377-384.
12. Cussen L, Mcdonnell T, Bennett G, Christopher |, Thompson J, Sherlock M, O'reilly MW. Approach to androgen excess in women: Clinical and biochemical insights. Clin Endocrinol (Oxf). Published online March 29, 2022. doi:10.1111/CEN.14710
13. Eilertsen TB, Vanky E, Carlsen SM. Anti-Mullerian hormone in the diagnosis of polycystic ovary syndrome: can morphologic description be replaced? Hum Reprod. 2012;27(8):2494-2502. doi:10.1093/HUMREP/DES213
14. Fauser BCJM. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. Fertil Steril. 2004;81(1):19-25. doi:10.1016/J.FERTNSTERT.2003.10.004
15. Teede H, Deeks A, Moran L. Polycystic Ovary Syndrome: A Complex Condition with Psychological, Reproductive and Metabolic Manifestations That Impacts on Health

across the Lifespan. Vol 8. BMC Med; 2010:1-10. doi:10.1186/1741-7015-8-41

16. Azziz R, Woods KS, Reyna R, Key TJ, Knochenhauer ES, Yildiz BO. The Prevalence and Features of the Polycystic Ovary Syndrome in an Unselected Population. J Clin Endocrinol Metab. 2004;89(6):2745-2749. doi:10.1210/JC.2003-032046

17. Lauritsen MP, Bentzen JG, Pinborg A, Loft A, Forman JL, Thuesen LL, Cohen A, Hougaard DM, Nyboe Andersen A. The prevalence of polycystic ovary syndrome in a normal population according to the Rotterdam criteria versus revised criteria including anti-Müllerian hormone. Hum Reprod. 2014;29(4):791-801. doi:10.1093/HUMREP/DET469

18. March WA, Moore VM, Willson KJ, Phillips DIW, Norman RJ, Davies MJ. The prevalence of polycystic ovary syndrome in a community sample assessed under contrasting diagnostic criteria. Hum Reprod. 2010;25(2):544-551. doi:10.1093/HUMREP/DEP399

19. Wilcock J, Taylor D. Polycystic ovarian syndrome: overdiagnosed and overtreated? Br J Gen Pract. 2018;68(670):243-243. doi:10.3399/BJGP18X696113

20. Copp T, Doust J, McCaffery K, Hersch J, Jansen J. Polycystic ovary syndrome: why widening the diagnostic criteria may be harming women. BMJ. 2021;373. doi:10.1136/BMJ.N700

21. Wang R, Mol BWJ. The Rotterdam criteria for polycystic ovary syndrome: evidence-based criteria? Hum Reprod. 2017;32(2):261-264. doi:10.1093/HUMREP/DEW287

22. Azziz R, Carmina E, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Futterweit W, Janssen OE, Legro RS, Norman RJ, Taylor AE, Witchel SF. Positions statement: criteria for defining polycystic ovary syndrome as a predominantly hyperandrogenic syndrome: an Androgen Excess Society guideline. J Clin Endocrinol Metab. 2006;91(11):4237-4245. doi:10.1210/JC.2006-0178

23. Rosenfield RL, Ehrmann DA. The Pathogenesis of Polycystic Ovary Syndrome (PCOS): The Hypothesis of PCOS as Functional Ovarian Hyperandrogenism Revisited. Endocr Rev. 2016;37(5):467. doi:10.1210/ER.2015-1104

24. ROBINSON S, RODIN DA, DEACON A, WHEELER MJ, CLAYTON RN. Which hormone tests for the diagnosis of polycystic ovary syndrome? Br J Obstet Gynaecol. 1992;99(3):232-238. doi:10.1111/J.1471-0528.1992.TB14505.X

25. Rodin A, Thakkar H, Taylor N, Clayton R. Hyperandrogenism in Polycystic Ovary Syndrome -- Evidence of Dysregulation of 11β-Hydroxysteroid Dehydrogenase. https://doi.org/101056/NEJM199402173300703. 1994;330(7):460-465. doi:10.1056/NEJM199402173300703

26. Kumarendran B, O'Reilly MWMW, Manolopoulos KN, Toulis KAKA, Gokhale KMKM, Sitch AJAJ, Wijeyaratne CNCN, Coomarasamy A, Arlt W, Nirantharakumar K. Polycystic ovary syndrome, androgen excess, and the risk of nonalcoholic fatty liver disease in women: A longitudinal study based on a United Kingdom primary care database. PLoS Med. 2018;15(3). doi:10.1371/journal.pmed.1002542

27. O'Reilly MW, Glisic M, Kumarendran B, Subramanian A, Manolopoulos KN, Tahrani AA, Keerthy D, Muka T, Toulis KA, Hanif W, Thomas GN, Franco OH, Arlt W, Nirantharakumar K. Serum testosterone, sex hormone-binding globulin and sex-specific risk of incident type 2 diabetes in a retrospective primary care cohort. Clin Endocrinol (Oxf). 2019;90(1):145-154. doi:10.1111/cen.13862

28. Kakoly NS, Khomami MB, Joham AE, Cooray SD, Misso ML, Norman RJ, Harrison CL, Ranasinha S, Teede HJ, Moran LJ. Ethnicity, obesity and the prevalence of impaired glucose tolerance and type 2 diabetes in PCOS: a systematic review and meta-regression. Hum Reprod Update. 2018;24(4):455-467. doi:10.1093/HUMUPD/DMY007

29. Rubin KH, Glintborg D, Nybo M, Abrahamsen B, Andersen M. Development and risk factors of type 2 diabetes in a nationwide population of women with polycystic ovary syndrome. J Clin Endocrinol Metab. 2017;102(10):3848-3857. doi:10.1210/jc.2017-01354

30. Kumarendran B, O'reilly MW, Subramanian A, Šsumilo D, Toulis K, Gokhale KM, Wijeratne CN, Coomarasamy A, Tahrani AA, Azoulay L, Arlt W, Nirantharakumar K, O'Reilly MW, Subramanian A, Šumilo D, Toulis K, Gokhale KM, Wijeratne CN, Coomarasamy A, et al. Polycystic Ovary Syndrome, Combined Oral Contraceptives, and the Risk of Dysglycemia: A Population-Based Cohort Study With a Nested Pharmacoepidemiological Case-Control Study. Diabetes Care. 2021;44(12):2758-2766. doi:10.2337/DC21-0437

31. Marshall JC, Dunaif A. All Women With PCOS Should Be Treated For Insulin Resistance. Fertil Steril. 2012;97(1):18. doi:10.1016/J.FERTNSTERT.2011.11.036

32. Azziz R, Marin C, Hoq L, Badamgarav E, Song P. Health care-related economic burden of the polycystic ovary syndrome during the reproductive life span. J Clin Endocrinol Metab. 2005;90(8):4650-4658. doi:10.1210/JC.2005-0628

33. Ding T, Hardiman PJ, Petersen I, Baio G. Incidence and prevalence of diabetes and cost of illness analysis of polycystic ovary syndrome: a Bayesian modelling study. Hum Reprod. 2018;33(7):1299-1306. doi:10.1093/HUMREP/DEY093

34. Toulis KA, Goulis DG, Kolibianakis EM, Venetis CA, Tarlatzis BC, Papadimas I. Risk of gestational diabetes mellitus in women with polycystic ovary syndrome: a systematic review and a meta-analysis. Fertil Steril. 2009;92(2):667-677. doi:10.1016/J.FERTNSTERT.2008.06.045

35. Scenario: Management - adults | Management | Polycystic ovary syndrome | CKS | NICE. Accessed March 25, 2022. https://cks.nice.org.uk/topics/polycystic-ovary-

syndrome/management/management-adults/#screening-for-type-2-diabetes-insulin-resistance

36. Barber TM, Hanson P, Weickert MO, Franks S. Obesity and Polycystic Ovary Syndrome: Implications for Pathogenesis and Novel Management Strategies. Clin Med Insights Reprod Heal. 2019;13:117955811987404. doi:10.1177/1179558119874042

37. Chan JL, Kar S, Vanky E, Morin-Papunen L, Piltonen T, Puurunen J, Tapanainen JS, Maciel GAR, Hayashida SAY, Soares JM, Baracat EC, Mellembakken JR, Dokras A. Racial and ethnic differences in the prevalence of metabolic syndrome and its components of metabolic syndrome in women with polycystic ovary syndrome: a regional cross-sectional study. Am J Obstet Gynecol. 2017;217(2):189.e1-189.e8. doi:10.1016/J.AJOG.2017.04.007

38. Livadas S, Anagnostis P, Bosdou JK, Bantouna D, Paparodis R. Polycystic ovary syndrome and type 2 diabetes mellitus: A state-of-the-art review. World J Diabetes. 2022;13(1):5. doi:10.4239/WJD.V13.I1.5

39. Hart R. Generational Health Impact of PCOS on Women and their Children. Med Sci. 2019;7(3):49. doi:10.3390/medsci7030049

40. Sirmans SM, Pate KA. Epidemiology, diagnosis, and management of polycystic ovary syndrome. 2013;6(1):1-13. Accessed March 24, 2021. https://pubmed.ncbi.nlm.nih.gov/24379699/

41. Palomba S, De Wilde MA, Falbo A, Koster MPH, La Sala GB, Fauser BCJM. Pregnancy complications in women with polycystic ovary syndrome. Hum Reprod Update. 2015;21(5):575-592. doi:10.1093/HUMUPD/DMV029

42. Palomba S, Falbo A, Chiossi G, Orio F, Tolino A, Colao A, La Sala GB, Zullo F. Low-grade chronic inflammation in pregnant women with polycystic ovary syndrome: a prospective controlled clinical study. J Clin Endocrinol Metab. 2014;99(8):2942-

2951. doi:10.1210/JC.2014-1214

43. Glintborg D, Rubin KH, Nybo M, Abrahamsen B, Andersen M. Morbidity and medicine prescriptions in a nationwide Danish population of patients diagnosed with polycystic ovary syndrome. 2015;172(5):627-638. doi:10.1530/EJE-14-1108

44. Kumarendran B, Sumilo D, O'Reilly MW, Toulis KA, Gokhale KM, Wijeyaratne CN, Coomarasamy A, Arlt W, Tahrani AA, Nirantharakumar K. Increased risk of obstructive sleep apnoea in women with polycystic ovary syndrome: a population-based cohort study. Eur J Endocrinol. 2019;180(4):265-272. doi:10.1530/EJE-18-0693

45. Zierau L, Cortes R, Thomsen SF, Jimenez-Solem E, Lindenberg S, Backer V. Asthma severity and fertility outcome in women with polycystic ovary syndrome: a registry-based study. ERJ Open Res. 2018;4(4). doi:10.1183/23120541.00138-2017

46. Ince N, Serdar DM, Keskin M, Dirikoc A, Topaloglu O, Ersoy R, Cakir B. Headache in patients with polycystic ovary syndrome. Endocr Abstr. 2020;70. doi:10.1530/ENDOABS.70.AEP852

47. Romitti M, Fabris VC, Ziegelmann PK, Maia AL, Spritzer PM. Association between PCOS and autoimmune thyroid disease: a systematic review and meta-analysis. Endocr Connect. 2018;7(11):1158-1167. doi:10.1530/EC-18-0309

48. Yin W, Falconer H, Yin L, Xu L, Ye W. Association Between Polycystic Ovary Syndrome and Cancer Risk. JAMA Oncol. 2019;5(1):106-107. doi:10.1001/JAMAONCOL.2018.5188

49. Yin X, Ji Y, Chan CLW, Chan CHY. The mental health of women with polycystic ovary syndrome: a systematic review and meta-analysis. Arch Womens Ment Health. 2021;24(1):11-27. doi:10.1007/S00737-020-01043-X/FIGURES/9

50. Khalil H, Malhotra K, Sheikh J, Hebbar M, Zia N, Jayaprakash S, Wicks S, Subramanian A, Chapman R, Gleeson H, Robinson L, Chu J, Lathia T, Selvan C,

O’Reilly MW, Manolopoulos K, Arlt W, Kempegowda P. A systematic review and meta-analysis assessing psychosexual wellbeing in people with polycystic ovary syndrome. Endocr Abstr. 2022;81. doi:10.1530/ENDOABS.81.P457

51. Ruth KS, Day FR, Tyrrell J, Thompson DJ, Wood AR, Mahajan A, Beaumont RN, Wittemans L, Martin S, Busch AS, Erzurumluoglu AM, Hollis B, O’Mara TA, McCarthy MI, Langenberg C, Easton DF, Wareham NJ, Burgess S, Murray A, et al. Using human genetics to understand the disease impacts of testosterone in men and women. Nat Med 2020 262. 2020;26(2):252-258. doi:10.1038/s41591-020-0751-5

52. Riaz H, Khan MS, Siddiqi TJ, Usman MS, Shah N, Goyal A, Khan SS, Mookadam F, Krasuski RA, Ahmed H. Association Between Obesity and Cardiovascular Outcomes: A Systematic Review and Meta-analysis of Mendelian Randomization Studies. JAMA Netw Open. 2018;1(7):e183788-e183788. doi:10.1001/JAMANETWORKOPEN.2018.3788

53. Zhu T, Cui J, Goodarzi MO. Polycystic Ovary Syndrome and Risk of Type 2 Diabetes, Coronary Heart Disease, and Stroke. Diabetes. 2021;70(2):627-637. doi:10.2337/DB20-0800

54. Zhu T, Cui J, Goodarzi MO. Polycystic ovary syndrome and breast cancer subtypes: a Mendelian randomization study. Am J Obstet Gynecol. 2021;225(1):99-101. doi:10.1016/J.AJOG.2021.03.020

55. Kjerulff LE, Sanchez-Ramos L, Duffy D. Pregnancy outcomes in women with polycystic ovary syndrome: a metaanalysis. 2011;204(6):558.e1-558.e6. doi:10.1016/J.AJOG.2011.03.021

56. Boomsma CM, Eijkemans MJC, Hughes EG, Visser GHA, Fauser BCJM, Macklon NS. A meta-analysis of pregnancy outcomes in women with polycystic ovary syndrome. Hum Reprod Update. 2006;12(6):673-683.

doi:10.1093/HUMUPD/DML036

57. Qin JZ, Pang LH, Li MJ, Fan XJ, Huang RD, Chen HY. Obstetric complications in women with polycystic ovary syndrome: A systematic review and meta-analysis. Reprod Biol Endocrinol. 2013;11(1):1-14. doi:10.1186/1477-7827-11-56/FIGURES/8

58. Maliqueo M, Echiburú B, Crisosto N, Amigo P, Aranda P, Sánchez F, Sir-Petermann T. Metabolic parameters in cord blood of newborns of women with polycystic ovary syndrome. Fertil Steril. 2009;92(1):277-282. doi:10.1016/J.FERTNSTERT.2008.04.022

59. Hu S, Leonard A, Seifalian A, Hardiman P. Vascular dysfunction during pregnancy in women with polycystic ovary syndrome. Hum Reprod. 2007;22(6):1532-1539. doi:10.1093/HUMREP/DEM028

60. Roos N, Kieler H, Sahlin L, Ekman-Ordeberg G, Falconer H, Stephansson O. Risk of adverse pregnancy outcomes in women with polycystic ovary syndrome: population based cohort study. BMJ. 2011;343(7828). doi:10.1136/BMJ.D6309

61. Han AR, Kim HO, Cha SW, Park CW, Kim JY, Yang KM, Song IO, Koong MK, Kang IS. Adverse pregnancy outcomes with assisted reproductive technology in non-obese women with polycystic ovary syndrome: a case-control study. Clin Exp Reprod Med. 2011;38(2):103-108. doi:10.5653/CERM.2011.38.2.103

62. Bell GA, Sundaram R, Mumford SL, Park H, Mills J, Bell EM, Broadney M, Yeung EH. Maternal polycystic ovarian syndrome and early offspring development. Hum Reprod. 2018;33(7):1307. doi:10.1093/HUMREP/DEY087

63. PCOS Increases Risk of Psychiatric, Neurodevelopment Disorders in Offspring. Accessed June 4, 2022. https://www.hcplive.com/view/pcos-increases-risk-psychiatric-neurodevelopment-disorders-offspring

64. Chen X, Koivuaho E, Piltonen TT, Gissler M, Lavebratt C. Association of maternal

polycystic ovary syndrome or anovulatory infertility with obesity and diabetes in offspring: a population-based cohort study. Hum Reprod. 2021;36(8):2345-2357. doi:10.1093/HUMREP/DEAB112

65. Risal S, Pei Y, Lu H, Manti M, Fornes R, Pui HP, Zhao Z, Massart J, Ohlsson C, Lindgren E, Crisosto N, Maliqueo M, Echiburú B, Ladrón de Guevara A, Sir-Petermann T, Larsson H, Rosenqvist MA, Cesta CE, Benrick A, et al. Prenatal androgen exposure and transgenerational susceptibility to polycystic ovary syndrome. Nat Med 2019 2512. 2019;25(12):1894-1904. doi:10.1038/s41591-019-0666-1

66. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. Accessed May 27, 2021. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

67. Deng G, Yin M, Chen X, Zeng F. Clinical determinants for fatality of 44,672 patients with COVID-19. Crit Care. 2020;24(1):1-3. doi:10.1186/s13054-020-02902-w

68. Docherty AB, Harrison EM, Green CA, Hardwick HE, Pius R, Norman L, Holden KA, Read JM, Dondelinger F, Carson G, Merson L, Lee J, Plotkin D, Sigfrid L, Halpin S, Jackson C, Gamble C, Horby PW, Nguyen-Van-Tam JS, et al. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. BMJ. 2020;369.

69. Booth A, Reed AB, Ponzo S, Yassaee A, Aral M, Plans D, Labrique A, Mohan D. Population risk factors for severe disease and mortality in COVID-19: A global systematic review and meta-analysis. PLoS One. 2021;16(3). doi:10.1371/JOURNAL.PONE.0247461

70. Kyrou I, Karteris E, Robbins T, Chatha K, Drenos F, Randeva HS. Polycystic ovary syndrome (PCOS) and COVID-19: An overlooked female patient population at

potentially higher risk during the COVID-19 pandemic. BMC Med. 2020;18(1):1-10. doi:10.1186/s12916-020-01697-5

71. Escobar-Morreale HF, Luque-Ramírez M, González F. Circulating inflammatory markers in polycystic ovary syndrome: a systematic review and metaanalysis. Fertil Steril. 2011;95(3):1048-1058.e2. doi:10.1016/J.FERTNSTERT.2010.11.036

72. Kalyan S, Goshtesabi A, Sarray S, Joannou A, Almawi WY. Assessing C reactive protein/albumin ratio as a new biomarker for polycystic ovary syndrome: a case–control study of women from Bahraini medical clinics. BMJ Open. 2018;8(10):e021860. doi:10.1136/BMJOPEN-2018-021860

73. Duleba AJ, Dokras A. Is PCOS an inflammatory process? Fertil Steril. 2012;97(1):7-12. doi:10.1016/J.FERTNSTERT.2011.11.023

74. Hillman SC, Bryce C, Caleyachetty R, Dale J. Women's experiences of diagnosis and management of polycystic ovary syndrome: a mixed-methods study in general practice. Br J Gen Pract. 2020;70(694):e322-e329. doi:10.3399/BJGP20X708881

75. de Medeiros SF. Risks, benefits size and clinical implications of combined oral contraceptive use in women with polycystic ovary syndrome. Reprod Biol Endocrinol 2017 151. 2017;15(1):1-17. doi:10.1186/S12958-017-0313-Y

76. Misso M, Costello M, Dokras A, Laven J, Moran L, Piltonen T, Norman R. under section 14A of the National Health and Medical Research Council Act. Natl Heal Med Res Counc. 2018;2.

77. Melo AS de, Reis RM dos, Ferriani RA, Vieira CS. Hormonal contraception in women with polycystic ovary syndrome: choices, challenges, and noncontraceptive benefits. Open Access J Contracept. 2017;8:13. doi:10.2147/OAJC.S85543

78. Słopień R, Milewska E, Rynio P, Męczekalski B. Use of oral contraceptives for management of acne vulgaris and hirsutism in women of reproductive and late

reproductive age. Przegląd Menopauzalny = Menopause Rev. 2018;17(1):1. doi:10.5114/PM.2018.74895

79. Hoeger KM, Dokras A, Piltonen T. Update on PCOS: Consequences, Challenges, and Guiding Treatment. J Clin Endocrinol Metab. 2021;106(3):e1071-e1083. doi:10.1210/CLINEM/DGAA839

80. Sendur SN, Yildiz BO. Influence of ethnicity on different aspects of polycystic ovary syndrome: a systematic review. Reprod Biomed Online. 2021;42(4):799-818. doi:10.1016/J.RBMO.2020.12.006

81. Williamson K, Gunn AJ, Johnson N, Milsom SR. The impact of ethnicity on the presentation of polycystic ovarian syndrome. Aust New Zeal J Obstet Gynaecol. 2001;41(2):202-206. doi:10.1111/J.1479-828X.2001.TB01210.X

82. Engmann L, Jin S, Sun F, Legro RS, Polotsky AJ, Hansen KR, Coutifaris C, Diamond MP, Eisenberg E, Zhang H, Santoro N, Bartlebaugh C, Dodson W, Estes S, Gnatuk C, Ober J, Brzyski R, Easton C, Hernandez A, et al. Racial and Ethnic Differences in the Polycystic Ovary Syndrome (PCOS) Metabolic Phenotype. Am J Obstet Gynecol. 2017;216(5):493.e1. doi:10.1016/J.AJOG.2017.01.003

83. Wolf WM, Wattick RA, Kinkade ON, Olfert MD. Geographical Prevalence of Polycystic Ovary Syndrome as Determined by Region and Race/Ethnicity. Int J Environ Res Public Health. 2018;15(11). doi:10.3390/IJERPH15112589

84. Dayo OM, Gordon NP, Chandra M, Weintraub ML, Lo JC. RACIAL AND ETHNIC DIFFERENCES IN THE PREVALENCE OF DIAGNOSED POLYCYSTIC OVARY SYNDROME IN WOMEN RECEIVING AMBULATORY CARE. Fertil Steril. 2020;114(3):e402-e403. doi:10.1016/J.FERTNSTERT.2020.08.1180

85. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, Staa T van, Smeeth L. Data Resource Profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol.

2015;44(3):827-836. doi:10.1093/IJE/DYV098

86. Okoth K, Crowe F, Marshall T, Thomas GN, Nirantharakumar K, Adderley NJ. Sex-specific temporal trends in the incidence and prevalence of cardiovascular disease in young adults: a population-based study using UK primary care data. Eur J Prev Cardiol. Published online February 9, 2022. doi:10.1093/EURJPC/ZWAC024

87. Adderley NJ, Subramanian A, Nirantharakumar K, Yiangou A, Gokhale KM, Mollan SP, Sinclair AJ. Association Between Idiopathic Intracranial Hypertension and Risk of Cardiovascular Diseases in Women in the United Kingdom. JAMA Neurol. 2019;76(9):1088-1098. doi:10.1001/JAMANEUROL.2019.1812

88. Harvey PR, Thomas T, Chandan JS, Mytton J, Coupland B, Bhala N, Evison F, Patel P, Nirantharakumar K, Trudgill NJ. Incidence, morbidity and mortality of patients with achalasia in England: findings from a study of nationwide hospital and primary care data. Gut. 2019;68(5). doi:10.1136/GUTJNL-2018-316089

89. Chandan JS, Gokhale KM, Bradbury-Jones C, Nirantharakumar K, Bandyopadhyay S, Taylor J. Exploration of trends in the incidence and prevalence of childhood maltreatment and domestic abuse recording in UK primary care: a retrospective cohort study using ‘the health improvement network’ database. BMJ Open. 2020;10(6):e036949. doi:10.1136/BMJOPEN-2020-036949

90. Crowe FL, Jolly K, Macarthur C, Manaseki-Holland S, Gittoes N, Hewison M, Scragg R, Nirantharakumar K. Trends in the incidence of testing for vitamin D deficiency in primary care in the UK: a retrospective analysis of The Health Improvement Network (THIN), 2005–2015. BMJ Open. 2019;9(6):e028355. doi:10.1136/BMJOPEN-2018-028355

91. Subramanian A, Adderley NJ, Gkoutos G V., Gokhale KM, Nirantharakumar K, Krishna MT. Ethnicity-based differences in the incident risk of allergic diseases and

autoimmune disorders: A UK-based retrospective cohort study of 4.4 million participants. Clin Exp Allergy. 2021;51(1):144-147. doi:10.1111/CEA.13741

92. Sachdeva G, Gainder S, Suri V, Sachdeva N, Chopra S. Comparison of the Different PCOS Phenotypes Based on Clinical Metabolic, and Hormonal Profile, and their Response to Clomiphene. Indian J Endocrinol Metab. 2019;23(3):326. doi:10.4103/IJEM.IJEM_30_19

93. Diagnosis | Polycystic ovary syndrome | CKS | NICE. Accessed March 21, 2022. https://cks.nice.org.uk/topics/polycystic-ovary-syndrome/diagnosis/

94. Long-term Consequences of Polycystic Ovary Syndrome. Published online 2014.

95. Taggar JS, Coleman T, Lewis S, Szatkowski L. The impact of the Quality and Outcomes Framework (QOF) on the recording of smoking targets in primary care medical records: Cross-sectional analyses from the Health Improvement Network (THIN) database. BMC Public Health. 2012;12(1):1-11. doi:10.1186/1471-2458-12-329/TABLES/3

96. PCOS UK guideline | Independent professional body guideline | Guidelines. Accessed March 21, 2022. https://www.guidelines.co.uk/womens-health/pcos-uk-guideline/236071.article

97. Wild RA, Carmina E, Diamanti-Kandarakis E, Dokras A, Escobar-Morreale HF, Futterweit W, Lobo R, Norman RJ, Talbott E, Dumesic DA. Assessment of Cardiovascular Risk and Prevention of Cardiovascular Disease in Women with the Polycystic Ovary Syndrome: A Consensus Statement by the Androgen Excess and Polycystic Ovary Syndrome (AE-PCOS) Society. J Clin Endocrinol Metab. 2010;95(5):2038-2049. doi:10.1210/JC.2009-2724

98. Mustaniemi S, Vääräsmäki M, Eriksson JG, Gissler M, Laivuori H, Ijäs H, Bloigu A, Kajantie E, Morin-Papunen L. Polycystic ovary syndrome and risk factors for

gestational diabetes. Endocr Connect. 2018;7(7):859-869. doi:10.1530/EC-18-0076

99. Palm CVB, Glintborg D, Kyhl HB, McIntyre HD, Jensen RC, Jensen TK, Jensen DM, Andersen M. Polycystic ovary syndrome and hyperglycaemia in pregnancy. A narrative review and results from a prospective Danish cohort study. Diabetes Res Clin Pract. 2018;145:167-177. doi:10.1016/J.DIABRES.2018.04.030

100. Minassian C, Williams R, Meeraus WH, Smeeth L, Campbell OMR, Thomas SL. Methods to generate and validate a Pregnancy Register in the UK Clinical Practice Research Datalink primary care database. Pharmacoepidemiol Drug Saf. 2019;28(7):923-933. Accessed March 24, 2021. https://onlinelibrary.wiley.com/doi/full/10.1002/pds.4811

101. Glossary | Type 2 diabetes: prevention in people at high risk | Guidance | NICE. Accessed June 2, 2022. https://www.nice.org.uk/guidance/ph38/chapter/glossary

102. Gokhale KM, Chandan JS, Toulis K, Gkoutos G, Tino P, Nirantharakumar K. Data extraction for epidemiological research (DExtER): a novel tool for automated clinical epidemiology studies. Eur J Epidemiol. 2021;36(2):165-178. doi:10.1007/S10654-020-00677-6/TABLES/6

103. Recommendations | Type 2 diabetes: prevention in people at high risk | Guidance | NICE. Accessed June 2, 2022. https://www.nice.org.uk/guidance/ph38/chapter/Recommendations#risk-assessment

104. Diabetes Prevention Week 2018-Preliminary Briefing Document for Those Involved in Delivery Background.; 2018. Accessed June 3, 2022. https://www.guywoodpractice.co.uk/website/P88607/files/Diabetes_Prevention_Week_2018_-_Briefing_Document.pdf

105. Pal K, Horsfall L, Sharma M, Nazareth I, Petersen I. Time trends in the incidence of clinically diagnosed type 2 diabetes and pre-diabetes in the UK 2009-2018: a

retrospective cohort study. BMJ Open Diab Res Care. 2021;9:1989. doi:10.1136/bmjdrc-2020-001989

106. Holden SH, Barnett AH, Peters JR, Jenkins-Jones S, Poole CD, Morgan CL, Currie CJ. The incidence of type 2 diabetes in the United Kingdom from 1991 to 2010. Diabetes, Obes Metab. 2013;15(9):844-852. doi:10.1111/DOM.12123

107. McManus E, Meacock R, Parkinson B, Sutton M. Population level impact of the NHS Diabetes Prevention Programme on incidence of type 2 diabetes in England: An observational study. Lancet Reg Heal - Eur. 2022;19:100420. doi:10.1016/J.LANEPE.2022.100420/ATTACHMENT/CF4F2DB1-6661-4D21-B055-BADA9AF0C0D5/MMC2.DOCX

108. Dabelea D, Snell-Bergeon JK, Hartsfield CL, Bischoff KJ, Hamman RF, McDuffie RS. Increasing Prevalence of Gestational Diabetes Mellitus (GDM) Over Time and by Birth CohortKaiser Permanente of Colorado GDM Screening Program. Diabetes Care. 2005;28(3):579-584. doi:10.2337/DIACARE.28.3.579

109. Recommendations | Diabetes in pregnancy: management from preconception to the postnatal period | Guidance | NICE. Accessed July 26, 2022. https://www.nice.org.uk/guidance/ng3/chapter/Recommendations#gestational-diabetes

110. Zeng X, Xie Y jie, Liu Y ting, Long S lian, Mo Z cheng. Polycystic ovarian syndrome: Correlation between hyperandrogenism, insulin resistance and obesity. Clin Chim Acta. 2020;502:214-221. doi:10.1016/J.CCA.2019.11.003

111. Glintborg D, Andersen M. MANAGEMENT OF ENDOCRINE DISEASE: Morbidity in polycystic ovary syndrome. Eur J Endocrinol. 2017;176(2):R53-R65. doi:10.1530/EJE-16-0373

112. Yu HF, Chen HS, Rao DP, Gong J. Association between polycystic ovary syndrome and the risk of pregnancy complications A PRISMA-compliant systematic review and

meta-analysis. Med (United States). 2016;95(51):e4863. doi:10.1097/MD.0000000000004863

113. Morris A. Transgenerational effects of polycystic ovary syndrome identified. Nat Rev Endocrinol 2019 162. 2019;16(2):67-67. doi:10.1038/s41574-019-0314-3

114. Palomba S, Santagni S, Falbo A, La Sala GB. Complications and challenges associated with polycystic ovary syndrome: current perspectives. Int J Womens Health. 2015;7:745. doi:10.2147/IJWH.S70314

115. Wu CH, Chiu LT, Chang YJ, Lee CI, Lee MS, Lee TH, Wei JCC. Hypertension Risk in Young Women With Polycystic Ovary Syndrome: A Nationwide Population-Based Cohort Study. Front Med. 2020;7:585. doi:10.3389/FMED.2020.574651/BIBTEX

116. Gaberšček S, Zaletel K, Schwetz V, Pieber T, Obermayer-Pietsch B, Lerchbaum E. MECHANISMS IN ENDOCRINOLOGY: Thyroid and polycystic ovary syndrome. Eur J Endocrinol. 2015;172(1):R9-R21. doi:10.1530/EJE-14-0295

117. Villar J, Ismail LC, Victora CG, Ohuma EO, Bertino E, Altman DG, Lambert A, Papageorghiou AT, Carvalho M, Jaffer YA, Gravett MG, Purwar M, Frederick IO, Noble AJ, Pang R, Barros FC, Chumlea C, Bhutta ZA, Kennedy SH, et al. International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. Lancet. 2014;384(9946):857-868. doi:10.1016/S0140-6736(14)60932-6

118. The English Indices of Deprivation 2019 (IoD2019).; 2019. Accessed February 27, 2022. https://www.gov.uk/government/publications/english-indices-of-deprivation-2019-technical-report

119. Sullivan TR, Lee KJ, Ryan P, Salter AB. Multiple imputation for handling missing outcome data when estimating the relative risk. BMC Med Res Methodol. 2017;17(1):1-10. doi:10.1186/S12874-017-0414-5/TABLES/4

120. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - A practical guide with flowcharts. BMC Med Res Methodol. 2017;17(1):1-10. doi:10.1186/S12874-017-0442-1/TABLES/2

121. Kontopantelis E, White IR, Sperrin M, Buchan I. Outcome-sensitive multiple imputation: A simulation study. BMC Med Res Methodol. 2017;17(1):1-13. doi:10.1186/S12874-016-0281-5/FIGURES/5

122. Thirumala Krishna M, Subramanian A, Adderley NJ, Zemedikun DT, Gkoutos G V, Nirantharakumar K. Allergic diseases and long-term risk of autoimmune disorders: longitudinal cohort study and cluster analysis. doi:10.1183/13993003.00476-2019

123. Hospital Episode Statistics (HES) - NHS Digital. Accessed May 19, 2022. https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics

124. Mills G, Badeghiesh A, Suarthana E, Baghlaf H, Dahan MH. Polycystic ovary syndrome as an independent risk factor for gestational diabetes and hypertensive disorders of pregnancy: a population-based study on 9.1 million pregnancies. Hum Reprod. 2020;35(7):1666-1674. doi:10.1093/HUMREP/DEAA099

125. Mulualem G, Wondim A, Woretaw A. The effect of pregnancy induced hypertension and multiple pregnancies on preterm birth in Ethiopia: A systematic review and meta-analysis. BMC Res Notes. 2019;12(1):1-7. doi:10.1186/S13104-019-4128-0/FIGURES/2

126. Hedderson MM, Ferrara A, Sacks DA. Gestational diabetes mellitus and lesser degrees of pregnancy hyperglycemia: association with increased risk of spontaneous preterm birth. Obstet Gynecol. 2003;102(4):850-856. doi:10.1016/S0029-7844(03)00661-6

127. Recommendations | Preterm labour and birth | Guidance | NICE. Accessed March 6,

2022. https://www.nice.org.uk/guidance/ng25/chapter/Recommendations#mode-of-birth

128. d'Alterio MN, Sigilli M, Succu AG, Ghisu V, Laganà AS, Sorrentino F, Nappi L, Tinelli R, Angioni S. Pregnancy outcomes in women with polycystic ovarian syndrome. Minerva Obstet Gynecol. 2022;74(1):45-59. doi:10.23736/S2724-606X.21.04758-4

129. Valgeirsdottir H, Poromaa IS, Kallak TK, Vanky E, Akhter T, Roos N, Stephansson O, Wikström AK. Polycystic ovary syndrome and extremely preterm birth: A nationwide register-based study. PLoS One. 2021;16(2):e0246743. doi:10.1371/JOURNAL.PONE.0246743

130. Zhang G, Feenstra B, Bacelis J, Liu X, Muglia LM, Juodakis J, Miller DE, Litterman N, Jiang PP, Russell L, Hinds DA, Hu Y, Weirauch MT, Chen X, Chavan AR, Wagner GP, Pavličev M, Nnamani MC, Maziarz J, et al. Genetic Associations with Gestational Duration and Spontaneous Preterm Birth. N Engl J Med. 2017;377(12):1156-1167. doi:10.1056/NEJMOA1612665/SUPPL_FILE/NEJMOA1612665_DISCLOSURES.PDF

131. Devarbhavi P, Telang L, Vastrad B, Tengli A, Vastrad C, Kottursheti I. Identification of key pathways and genes in polycystic ovary syndrome via integrated bioinformatics analysis and prediction of small therapeutic molecules. Reprod Biol Endocrinol. 2021;19(1):31. doi:10.1186/S12958-021-00706-3

132. Kempegowda P, Melson E, Manolopoulos KN, Arlt W, O'Reilly MW. Implicating androgen excess in propagating metabolic disease in polycystic ovary syndrome. Ther Adv Endocrinol Metab. 2020;11. doi:10.1177/2042018820934319

133. Dole N, Savitz DA, Hertz-Picciotto I, Siega-Riz AM, McMahon MJ, Buekens P. Maternal stress and preterm birth. Am J Epidemiol. 2003;157(1):14-24.

doi:10.1093/AJE/KWF176

134. Valgeirsdottir H, Kunovac Kallak T, Sundström Poromaa I, Jonsson M, Roos N, Lindström L, Wikström AK. Polycystic ovary syndrome and risk of stillbirth: a nationwide register-based study. BJOG An Int J Obstet Gynaecol. 2021;128(13):2073-2082. doi:10.1111/1471-0528.16890

135. Holton S, Hammarberg K, Johnson L. Fertility concerns and related information needs and preferences of women with PCOS. Hum Reprod Open. 2018;2018(4):1-7. doi:10.1093/HROPEN/HOY019

136. Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, Curtis HJ, Mehrkar A, Evans D, Inglesby P, Cockburn J, McDonald HI, MacKenna B, Tomlinson L, Douglas IJ, Rentsch CT, Mathur R, Wong AYS, Grieve R, et al. Factors associated with COVID-19-related death using OpenSAFELY. Nature. 2020;584(7821):430-436. doi:10.1038/s41586-020-2521-4

137. Cho LW, Randeva HS, Atkin SL. Cardiometabolic aspects of polycystic ovarian syndrome. Vasc Health Risk Manag. 2007;3(1):55-63. https://pubmed.ncbi.nlm.nih.gov/17583175

138. Moran LJ, Misso ML, Wild RA, Norman RJ. Impaired glucose tolerance, type 2 diabetes and metabolic syndrome in polycystic ovary syndrome: a systematic review and meta-analysis. Hum Reprod Update. 2010;16(4):347-363. doi:10.1093/humupd/dmq001

139. Diamanti-Kandarakis E, Dunaif A. Insulin Resistance and the Polycystic Ovary Syndrome Revisited: An Update on Mechanisms and Implications. Endocr Rev. 2012;33(6):981-1030. doi:10.1210/er.2011-1034

140. Mani H, Levy MJ, Davies MJ, Morris DH, Gray LJ, Bankart J, Blackledge H, Khunti K, Howlett TA. Diabetes and cardiovascular events in women with polycystic ovary

syndrome: A 20-year retrospective cohort study. Clin Endocrinol (Oxf). 2013;78(6):926-934. doi:10.1111/cen.12068

141. Ding T, Hardiman PJ, Petersen I, Wang FF, Qu F, Baio G. The prevalence of polycystic ovary syndrome in reproductiveaged women of different ethnicity: A systematic review and meta-analysis. Oncotarget. 2017;8(56):96351-96358. doi:10.18632/oncotarget.19180

142. Niedzwiedz CL, O'Donnell CA, Jani BD, Demou E, Ho FK, Celis-Morales C, Nicholl BI, Mair FS, Welsh P, Sattar N, Pell JP, Katikireddi SV. Ethnic and socioeconomic differences in SARS-CoV-2 infection: prospective cohort study using UK Biobank. BMC Med. 2020;18(1):160. doi:10.1186/s12916-020-01640-8

143. DiMaggio C, Klein M, Berry C, Frangos S. Black/African American Communities are at highest risk of COVID-19: spatial modeling of New York City ZIP Code–level testing results. Ann Epidemiol. 2020;51:7-13. doi:https://doi.org/10.1016/j.annepidem.2020.08.012

144. Martin CA, Jenkins DR, Minhas JS, Gray LJ, Tang J, Williams C, Sze S, Pan D, Jones W, Verma R, Knapp S, Major R, Davies M, Brunskill N, Wiselka M, Brightling C, Khunti K, Haldar P, Pareek M. Socio-demographic heterogeneity in the prevalence of COVID-19 during lockdown is associated with ethnicity and household size: Results from an observational cohort study. EClinicalMedicine. 2020;25. doi:10.1016/j.eclinm.2020.100466

145. Sapey E, Gallier S, Mainey C, Nightingale P, McNulty D, Crothers H, Evison F, Reeves K, Pagano D, Denniston AK, Nirantharakumar K, Diggle P, Ball S. Ethnicity and risk of death in patients hospitalised for COVID-19 infection in the UK: an observational cohort study in an urban catchment area. BMJ Open Respir Res. 2020;7(1):e000644. doi:10.1136/bmjresp-2020-000644

146. Hull SA, Williams C, Ashworth M, Carvalho C, Boomla K. Prevalence of suspected COVID-19 infection in patients from ethnic minority populations: a cross-sectional study in primary care. Br J Gen Pract. 2020;70(699):e696-e704. doi:10.3399/bjgp20X712601

147. Springate DA, Kontopantelis E, Ashcroft DM, Olier I, Parisi R, Chamapiwa E, Reeves D. ClinicalCodes: an online clinical codes repository to improve the validity and reproducibility of research using electronic medical records. PLoS One. 2014;9(6):e99825-e99825. doi:10.1371/journal.pone.0099825

148. Vintzileos WS, Muscat J, Hoffmann E, John NS, Vertichio R, Vintzileos AM, Vo D. Screening all pregnant women admitted to labor and delivery for the virus responsible for coronavirus disease 2019. Am J Obstet Gynecol. 2020;223(2):284-286. doi:10.1016/j.ajog.2020.04.024

149. COVID-19 Next steps in primary care surveillance. Accessed May 27, 2021. https://www.rcgp.org.uk/clinical-and-research/about/clinical-news/2020/march/covid-19-next-steps-in-primary-care-surveillance.aspx

150. Krishna MT, Subramanian A, Adderley NJ, Zemedikun DT, Gkoutos G V, Nirantharakumar K. Allergic diseases and long-term risk of autoimmune disorders: longitudinal cohort study and cluster analysis. Eur Respir J. 2019;54(5):1900476. doi:10.1183/13993003.00476-2019

151. Holman N, Knighton P, Kar P, O'Keefe J, Curley M, Weaver A, Barron E, Bakhai C, Khunti K, Wareham NJ, Sattar N, Young B, Valabhji J. Risk factors for COVID-19-related mortality in people with type 1 and type 2 diabetes in England: a population-based cohort study. Lancet Diabetes Endocrinol. 2020;8(10):823-833. doi:10.1016/S2213-8587(20)30271-0

152. Gibson-Helm ME, Lucas IM, Boyle JA, Teede HJ. Women's experiences of polycystic

ovary syndrome diagnosis. Fam Pr. 2014;31(5):545-549. doi:10.1093/fampra/cmu028

153. Gibson-Helm M, Teede H, Dunaif A, Dokras A. Delayed Diagnosis and a Lack of Information Associated With Dissatisfaction in Women With Polycystic Ovary Syndrome. J Clin Endocrinol Metab. 2017;102(2):604-612. doi:10.1210/jc.2016-2963

154. Liu M, Gao J, Zhang Y, Li P, Wang H, Ren X, Li C. Serum levels of TSP-1, NF-kappaB and TGF-beta1 in polycystic ovarian syndrome (PCOS) patients in northern China suggest PCOS is associated with chronic inflammation. Clin Endocrinol. 2015;83(6):913-922. doi:10.1111/cen.12951

155. Puder JJ, Varga S, Kraenzlin M, De Geyter C, Keller U, Muller B. Central fat excess in polycystic ovary syndrome: relation to low-grade inflammation and insulin resistance. J Clin Endocrinol Metab. 2005;90(11):6014-6021. doi:10.1210/jc.2005-1002

156. Escobar-Morreale HF, Botella-Carretero JI, Villuendas G, Sancho J, San Millan JL. Serum interleukin-18 concentrations are increased in the polycystic ovary syndrome: relationship to insulin resistance and to obesity. J Clin Endocrinol Metab. 2004;89(2):806-811. doi:10.1210/jc.2003-031365

157. Han MS, White A, Perry RJ, Camporez JP, Hidalgo J, Shulman GI, Davis RJ. Regulation of adipose tissue inflammation by interleukin 6. Proc Natl Acad Sci U S A. 2020;117(6):2751-2760. doi:10.1073/pnas.1920004117

158. Sbierski-Kind J, Mai K, Kath J, Jurisch A, Streitz M, Kuchenbecker L, Babel N, Nienen M, Jurchott K, Spranger L, Jumpertz von Schwartzenberg R, Decker AM, Kruger U, Volk HD, Spranger J. Association between Subcutaneous Adipose Tissue Inflammation, Insulin Resistance, and Calorie Restriction in Obese Females. J Immunol. 2020;205(1):45-55. doi:10.4049/jimmunol.2000108

159. Ortolan A, Lorenzin M, Felicetti M, Doria A, Ramonda R. Does gender influence

clinical expression and disease outcomes in COVID-19? A systematic review and meta-analysis. Int J Infect Dis. Published online 2020. doi:10.1016/j.ijid.2020.07.076

160. Montopoli M, Zumerle S, Vettor R, Rugge M, Zorzi M, Catapano C V, Carbone GM, Cavalli A, Pagano F, Ragazzi E, Prayer-Galetti T, Alimonti A. Androgen-deprivation therapies for prostate cancer and risk of infection by SARS-CoV-2: a population-based study (N = 4532). Ann Oncol. 2020;31(8):1040-1045. doi:10.1016/j.annonc.2020.04.479

161. Wambier CG, Vano-Galvan S, McCoy J, Pai S, Dhurat R, Goren A. Androgenetic alopecia in COVID-19: compared to age-matched epidemiologic studies and hospital outcomes with or without the Gabrin sign. J Am Acad Dermatol. Published online 2020. doi:10.1016/j.jaad.2020.07.099

162. Rastrelli G, Di Stasi V, Inglese F, Beccaria M, Garuti M, Di Costanzo D, Spreafico F, Greco GF, Cervi G, Pecoriello A, Magini A, Todisco T, Cipriani S, Maseroli E, Corona G, Salonia A, Lenzi A, Maggi M, De Donno G, et al. Low testosterone levels predict clinical adverse outcomes in SARS-CoV-2 pneumonia patients. Androl. Published online 2020. doi:10.1111/andr.12821

163. Çayan S, Uğuz M, Saylam B, Akbay E. Effect of serum total testosterone and its relationship with other laboratory parameters on the prognosis of coronavirus disease 2019 (COVID-19) in SARS-CoV-2 infected male patients: a cohort study. aging male. 2020;ahead-of-p(ahead-of-print):1-11. doi:10.1080/13685538.2020.1807930

164. Schiffer L, Kempegowda P, Arlt W, O'Reilly MW. MECHANISMS IN ENDOCRINOLOGY: The sexually dimorphic role of androgens in human metabolic disease. Eur J Endocrinol. 2017;177(3):R125-R143. doi:10.1530/EJE-17-0124

165. Bhowmick NA, Oft J, Dorff T, Pal S, Agarwal N, Figlin RA, Posadas EM, Freedland SJ, Gong J. COVID-19 and androgen-targeted therapy for prostate cancer patients.

Endocr Relat Cancer. 2020;27(9):R281-R292. doi:10.1530/ERC-20-0165

166. O'Reilly MW, Kempegowda P, Jenkinson C, Taylor AE, Quanson JL, Storbeck KH, Arlt W. 11-Oxygenated C19 Steroids Are the Predominant Androgens in Polycystic Ovary Syndrome. J Clin Endocrinol Metab. 2017;102(3):840-848. doi:10.1210/jc.2016-3285

167. Gilliver SC. Sex steroids as inflammatory regulators. J Steroid Biochem Mol Biol. 2010;120(2-3):105-115. doi:10.1016/j.jsbmb.2009.12.015

168. Schiffer L, Bossey A, Taylor AE, Akerman I, Scheel-Toellner D, Storbeck KH, Arlt W. Peripheral blood mononuclear cells preferentially activate 11-oxygenated androgens. bioRxiv. Published online 2020.

169. O'Reilly MW, Taylor AE, Crabtree NJ, Hughes BA, Capper F, Crowley RK, Stewart PM, Tomlinson JW, Arlt W. Hyperandrogenemia Predicts Metabolic Phenotype in Polycystic Ovary Syndrome: The Utility of Serum Androstenedione. J Clin Endocrinol Metab. 2014;99(3):1027-1036. doi:10.1210/jc.2013-3399

170. Dahlgren E, Landin K, Krotkiewski M, Holm G, Janson PO. Effects of two antiandrogen treatments on hirsutism and insulin sensitivity in women with polycystic ovary syndrome. Hum Reprod. 1998;13(10):2706-2711.

171. Vrbíková J, Hill M, Dvořáková K, Stanická S, Vondra K, Stárka L. Flutamide suppresses adrenal steroidogenesis but has no effect on insulin resistance and secretion and lipid levels in overweight women with polycystic ovary syndrome. Gynecol Obstet Invest. 2004;58(1):36-41. doi:10.1159/000077827

172. Firman N, Palmer MJ, Timæus IM, Wellings K. Contraceptive method use among women and its association with age, relationship status and duration: findings from the third British National Survey of Sexual Attitudes and Lifestyles (Natsal-3). BMJ Sex Reprod Heal. 2018;44(3):165-174. doi:10.1136/BMJSRH-2017-200037

173. Rivera R, Yacobson I, Grimes D. The mechanism of action of hormonal contraceptives and intrauterine contraceptive devices. Am J Obstet Gynecol. 1999;181(5 I):1263-1269. doi:10.1016/S0002-9378(99)70120-1

174. The effect of combined oral contraception on testosterone levels in healthy women: a systematic review and meta-analysis - Database of Abstracts of Reviews of Effects (DARE): Quality-assessed Reviews - NCBI Bookshelf. Accessed June 9, 2022. https://www.ncbi.nlm.nih.gov/books/NBK164632/

175. Nader S, Diamanti-Kandarakis E. Polycystic ovary syndrome, oral contraceptives and metabolic issues: new perspectives and a unifying hypothesis. Hum Reprod. 2007;22(2):317-322. doi:10.1093/HUMREP/DEL407

176. Cheang KI, Essah PA, Sharma S, Wickham EP, Nestler JE. Divergent effects of a combined hormonal oral contraceptive on insulin sensitivity in lean versus obese women. Fertil Steril. 2011;96(2):353-359.e1. doi:10.1016/j.fertnstert.2011.05.039

177. Adeniji AA, Essah PA, Nestler JE, Cheang KI. Metabolic Effects of a Commonly Used Combined Hormonal Oral Contraceptive in Women With and Without Polycystic Ovary Syndrome. J Women's Heal. 2016;25(6):638. doi:10.1089/JWH.2015.5418

178. Molloy GJ, Sweeney LA, Byrne M, Hughes CM, Ingham R, Morgan K, Murphy AW. Prescription contraception use: a cross-sectional population study of psychosocial determinants. BMJ Open. 2015;5(8):e007794. doi:10.1136/BMJOPEN-2015-007794

179. PCOS & The Pill – a discussion marking International Women's Day. Accessed June 9, 2022. https://www.birmingham.ac.uk/events/international-womens-day/2022/pcos-the-pill.aspx

180. Blak BT, Thompson M, Dattani H, Bourke A. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence

and mortality rates. Inform Prim Care. 2011;19(4):251-255. doi:10.14236/JHI.V19I4.820

181. Gokhale KM, Adderley NJ, Subramanian A, Lee WH, Han D, Coker J, Braithwaite T, Denniston AK, Keane PA, Nirantharakumar K, Pearse D, Keane A. Metformin and risk of age-related macular degeneration in individuals with type 2 diabetes: a retrospective cohort study. Br J Ophthalmol. 2022;0:bjophthalmol-2021-319641. doi:10.1136/BJOPHTHALMOL-2021-319641

182. Kontopantelis E, Reeves D, Valderas JM, Campbell S, Doran T. Recorded quality of primary care for patients with diabetes in England before and after the introduction of a financial incentive scheme: a longitudinal observational study. BMJ Qual Saf. 2013;22(1):53-64. doi:10.1136/BMJQS-2012-001033

183. Maguire A, Blak BT, Thompson M. The importance of defining periods of complete mortality reporting for research using automated data from primary care. Pharmacoepidemiol Drug Saf. 2009;18(1):76-83. doi:10.1002/pds.1688

184. Subramanian A, Anand A, Adderley NJ, Okoth K, Toulis KA, Gokhale K, Sainsbury C, O'Reilly MW, Arlt W, Nirantharakumar K. Increased COVID-19 infections in women with polycystic ovary syndrome: A population-based study. Eur J Endocrinol. 2021;184(5):637-645. doi:10.1530/EJE-20-1163

185. Suissa S, Dell'Aniello S, Vahey S, Renoux C. Time-window bias in case-control studies: statins and lung cancer. Epidemiology. 2011;22(2):228-231. doi:10.1097/EDE.0B013E3182093A0F

186. Kim SW, Jeon JH, Lee WK, Lee S, Kim JG, Lee IK, Park KG. Long-term effects of oral contraceptives on the prevalence of diabetes in post-menopausal women: 2007-2012 KNHANES. Endocrine. 2016;53(3):816-822. doi:10.1007/S12020-016-0972-Z

187. Yao W, Dong X, Yu X, Luo J, Zhang D. The use of oral contraceptive is inversely

associated with the risk of type 2 diabetes mellitus among middle-aged women. Gynecol Endocrinol. 2021;37(8):758-763. doi:10.1080/09513590.2021.1932802

188. Rimm EB, Manson JE, Stampfer MJ, Colditz GA, Willett WC, Rosner B, Hennekens CH, Speizer FE. Oral contraceptive use and the risk of type 2 (non-insulin-dependent) diabetes mellitus in a large prospective study of women. Diabetologia. 1992;35(10):967-972. doi:10.1007/BF00401427

189. Halperin IJ, Sujana Kumar S, Stroup DF, Laredo SE. The association between the combined oral contraceptive pill and insulin resistance, dysglycemia and dyslipidemia in women with polycystic ovary syndrome: a systematic review and meta-analysis of observational studies. Hum Reprod. 2011;26(1):191-201. doi:10.1093/HUMREP/DEQ301

190. Kahraman K, Şükür YE, Atabekoğlu CS, Ateş C, Taşkin S, Çetinkaya ŞE, Tolunay HE, Özmen B, Sönmezer M, Berker B. Comparison of two oral contraceptive forms containing cyproterone acetate and drospirenone in the treatment of patients with polycystic ovary syndrome: a randomized clinical trial. Arch Gynecol Obstet. 2014;290(2):321-328. doi:10.1007/S00404-014-3217-5

191. Chen MJ, Yang WS, Chen HF, Kuo JJ, Ho HN, Yang YS, Chen SU. Increased follistatin levels after oral contraceptive treatment in obese and non-obese women with polycystic ovary syndrome. Hum Reprod. 2010;25(3):779-785. doi:10.1093/HUMREP/DEP459

192. Bhattacharya SM, Jha A. Comparative study of the therapeutic effects of oral contraceptive pills containing desogestrel, cyproterone acetate, and drospirenone in patients with polycystic ovary syndrome. Fertil Steril. 2012;98(4):1053-1059. doi:10.1016/J.FERTNSTERT.2012.06.035

193. Cagnacci A, Tirelli A, Renzi A, Paoletti AM, Volpe A. Effects of two different oral

contraceptives on homocysteine metabolism in women with polycystic ovary syndrome. Contraception. 2006;73(4):348-351. doi:10.1016/J.CONTRACEPTION.2005.09.011

194. Elter K, Imir G, Durmusoglu F. Clinical, endocrine and metabolic effects of metformin added to ethinyl estradiol-cyproterone acetate in non-obese women with polycystic ovarian syndrome: a randomized controlled study. Hum Reprod. 2002;17(7):1729-1737. doi:10.1093/HUMREP/17.7.1729

195. Battaglia C, Mancini F, Fabbri R, Persico N, Busacchi P, Facchinetti F, Venturoli S. Polycystic ovary syndrome and cardiovascular risk in young patients treated with drospirenone-ethinylestradiol or contraceptive vaginal ring. A prospective, randomized, pilot study. Fertil Steril. 2010;94(4):1417-1425. doi:10.1016/J.FERTNSTERT.2009.05.044

196. Wang QY, Song Y, Huang W, Xiao L, Wang QS, Feng GM. Comparison of Drospirenone- with Cyproterone Acetate-Containing Oral Contraceptives, Combined with Metformin and Lifestyle Modifications in Women with Polycystic Ovary Syndrome and Metabolic Disorders: A Prospective Randomized Control Trial. Chin Med J (Engl). 2016;129(8):883-890. doi:10.4103/0366-6999.179783

197. Mathur R, Levin O, Azziz R. Use of ethinylestradiol/drospirenone combination in patients with the polycystic ovary syndrome. Ther Clin Risk Manag. 2008;4(2):487. doi:10.2147/TCRM.S6864

198. Rice MA, Malhotra S V., Stoyanova T. Second-Generation Antiandrogens: From Discovery to Standard of Care in Castration Resistant Prostate Cancer. Front Oncol. 2019;9(AUG):801. doi:10.3389/FONC.2019.00801

199. Mani H, Levy M, Howlett T, Fray L, Webb D, Srinivasan B, Khunti K, Davies M. Apparent under-reporting of polycystic ovary syndrome in primary care |

SFEBES2009 | Society for Endocrinology BES 2010 | Endocrine Abstracts. In: Endocrine Abstracts. ; 2010. Accessed July 18, 2022. https://www.endocrine-abstracts.org/ea/0021/ea0021P324

200. Liu J, Wu Q, Hao Y, Jiao M, Wang X, Jiang S, Han L. Measuring the global disease burden of polycystic ovary syndrome in 194 countries: Global Burden of Disease Study 2017. Hum Reprod. 2021;36(4):1108-1119. doi:10.1093/HUMREP/DEAA371

201. Gorry A, White DM, Franks S. Infertility in polycystic ovary syndrome: focus on low-dose gonadotropin treatment. Endocrine. 2006;30(1):27-33. doi:10.1385/ENDO:30:1:27

202. Recommendations | Fertility problems: assessment and treatment | Guidance | NICE. Accessed July 18, 2022. https://www.nice.org.uk/guidance/cg156/chapter/Recommendations#people-with-cancer-who-wish-to-preserve-fertility

203. Gjønnaess H. Ovarian electrocautery in the treatment of women with polycystic ovary syndrome (PCOS). Factors affecting the results. Acta Obstet Gynecol Scand. 1994;73(5):407-412. doi:10.3109/00016349409006253

204. Sawant S, Bhide P. Fertility Treatment Options for Women With Polycystic Ovary Syndrome. Clin Med Insights Reprod Heal. 2019;13:117955811989086. doi:10.1177/1179558119890867

205. Atkinson L, Kite C, McGregor G, James T, Clark CCT, Randeva HS, Kyrou I. Uncertainty, anxiety and isolation: Experiencing the covid-19 pandemic and lockdown as a woman with polycystic ovary syndrome (pcos). J Pers Med. 2021;11(10):952. doi:10.3390/JPM11100952/S1

206. Kite C, Atkinson L, McGregor G, Clark CCT, Brown JE, Kyrou I, Randeva HS. Sleep Disruption and Depression, Stress and Anxiety Levels in Women With Polycystic

Ovary Syndrome (PCOS) During the Lockdown Measures for COVID-19 in the UK. Front Glob Women's Heal. 2021;0:35. doi:10.3389/FGWH.2021.649104

207. Barry JA, Azizia MM, Hardiman PJ. Risk of endometrial, ovarian and breast cancer in women with polycystic ovary syndrome: a systematic review and meta-analysis. Hum Reprod Update. 2014;20(5):748-758. doi:10.1093/HUMUPD/DMU012

208. Michels KA, Pfeiffer RM, Brinton LA, Trabert B. Modification of the Associations Between Duration of Oral Contraceptive Use and Ovarian, Endometrial, Breast, and Colorectal Cancers. JAMA Oncol. 2018;4(4):516-521. doi:10.1001/JAMAONCOL.2017.4942

209. Diamanti-Kandarakis E, Baillargeon JP, Iuorno MJ, Jakubowicz DJ, Nestler JE. A Modern Medical Quandary: Polycystic Ovary Syndrome, Insulin Resistance, and Oral Contraceptive Pills. J Clin Endocrinol Metab. 2003;88(5):1927-1932. doi:10.1210/JC.2002-021528

210. Kontopantelis E, Panagioti M, Farragher T, Munford LA, Parisi R, Planner C, Spooner S, Tse A, Ashcroft DM, Esmail A. Consultation patterns and frequent attenders in UK primary care from 2000 to 2019: a retrospective cohort analysis of consultation events across 845 general practices. BMJ Open. 2021;11(12):e054666. doi:10.1136/BMJOPEN-2021-054666

211. Getting the best out of remote consulting in general practice – practical challenges and policy opportunities | The Nuffield Trust. Accessed July 18, 2022. https://www.nuffieldtrust.org.uk/research/getting-the-best-out-of-remote-consulting-in-general-practice-practical-challenges-and-policy-opportunities

212. Hardy T, De Sousa S, Norman RJ. Polycystic ovary syndrome: Prognosis and risk of comorbidity. Nov Insights into Pathophysiol Treat PCOS. Published online February 1, 2013:123-137. doi:10.2217/EBO.12.214

213. Mumusoglu S, Yildiz BO. Polycystic ovary syndrome phenotypes and prevalence: Differential impact of diagnostic criteria and clinical versus unselected population. Curr Opin Endocr Metab Res. 2020;12:66-71. doi:10.1016/J.COEMR.2020.03.004

Appendix

Supplementary 1: Papers published during the period of postgraduate study

1	**Subramanian, A.**, Gokhale, K., Sainsbury, C., Nirantharakumar, K., Toulis, KA. (2022). Sodium-glucosecotransporter-2 inhibitors and the risk of gout in patients with type 2 diabetes mellitus: A propensity-score-matched, new-user design study with an active comparator using the IQVIA Medical Research Data UK database. Diabetes Obes Metab, 1-10. https://dom-pubs.onlinelibrary.wiley.com/doi/10.1111/dom.14858
2	Jaime Adderley, N., **Subramanian, A.,** Perrins, M., Nirantharakumar, K., Mph Mfph Mrcp, M., Mollan, S. P., & Jean Sinclair, A. (2022). Headache, Opiate Use, and Prescribing Trends in Women With Idiopathic Intracranial Hypertension: A Population-Based Matched Cohort Study. Neurology, 2, 10.1212/WNL.0000000000201064. https://doi.org/10.1212/WNL.0000000000201064
3	Alshehri, Z., **Subramanian, A.,** Adderley, N. J., Gokhale, K. M., Karamat, M. A., Ray, C. J., Kumar, P., Nirantharakumar, K., & Tahrani, A. A. (2022). Risk of incident obstructive sleep apnoea in patients with type 1 diabetes: a population-based retrospective cohort study. Diabetologia, 65(8), 1353–1363. https://doi.org/10.1007/s00125-022-05714-5
4	Hughes, S. E., Haroon, S., **Subramanian, A.,** McMullan, C., Aiyegbusi, O. L., Turner, G. M., Jackson, L., Davies, E. H., Frost, C., McNamara, G., Price, G., Matthews, K., Camaradou, J., Ormerod, J., Walker, A., & Calvert, M. J. (2022). Development and validation of the symptom burden questionnaire for long covid (SBQ-LC): Rasch analysis. BMJ (Clinical research ed.), 377, e070230. https://doi.org/10.1136/bmj-2022-070230
5	**Subramanian, A**., Han, D., Braithwaite, T., Thayakaran, R., Zemedikun, D. T., Gokhale, K. M., Lee, W. H., Coker, J., Keane, P. A., Denniston, A. K., Nirantharakumar, K., Azoulay, L., & Adderley, N. J. (2022). Angiotensin-converting enzyme inhibitors and risk of age-related macular degeneration in individuals with hypertension. British journal of clinical pharmacology, 10.1111/bcp.15366. Advance online publication. https://doi.org/10.1111/bcp.15366
6	Haroon, S., Nirantharakumar, K., Hughes, S. E., **Subramanian, A.,** Aiyegbusi, O. L., Davies, E. H., Myles, P., Williams, T., Turner, G., Chandan, J. S., McMullan, C., Lord, J., Wraith, D. C., McGee, K., Denniston, A. K., Taverner, T., Jackson, L. J., Sapey, E., Gkoutos, G., Gokhale, K., … Calvert, M. (2022). Therapies for Long COVID in non-hospitalised individuals: from symptoms, patient-reported outcomes and immunology to targeted therapies (The TLC Study). BMJ open, 12(4), e060413. https://doi.org/10.1136/bmjopen-2021-060413
7	Lee, S. I., Azcoaga-Lorenzo, A., Agrawal, U., Kennedy, J. I., Fagbamigbe, A. F., Hope, H., **Subramanian, A**., Anand, A., Taylor, B., Nelson-Piercy, C., Damase-Michel, C., Yau, C., Crowe, F., Santorelli, G., Eastwood, K. A., Vowles, Z., Loane, M., Moss, N., Brocklehurst, P., Plachcinski, R., … MuM-PreDiCT Group (2022). Epidemiology of pre-existing multimorbidity in pregnant women in the UK in 2018: a population-based cross-sectional study. BMC pregnancy and childbirth, 22(1), 120. https://doi.org/10.1186/s12884-022-04442-3

8	Gokhale, K. M., Adderley, N. J., **Subramanian, A.,** Lee, W. H., Han, D., Coker, J., Braithwaite, T., Denniston, A. K., Keane, P. A., & Nirantharakumar, K. (2022). Metformin and risk of age-related macular degeneration in individuals with type 2 diabetes: a retrospective cohort study. The British journal of ophthalmology, bjophthalmol-2021-319641. Advance online publication. https://doi.org/10.1136/bjophthalmol-2021-319641
9	Okoth, K., **Subramanian, A.,** Chandan, J. S., Adderley, N. J., Thomas, G. N., Nirantharakumar, K., & Antza, C. (2022). Long term miscarriage-related hypertension and diabetes mellitus. Evidence from a United Kingdom population-based cohort study. PloS one, 17(1), e0261769. https://doi.org/10.1371/journal.pone.0261769
10	Prete, A., **Subramanian, A.,** Bancos, I., Chortis, V., Tsagarakis, S., Lang, K., Macech, M., Delivanis, D. A., Pupovac, I. D., Reimondo, G., Marina, L. V., Deutschbein, T., Balomenaki, M., O'Reilly, M. W., Gilligan, L. C., Jenkinson, C., Bednarczuk, T., Zhang, C. D., Dusek, T., Diamantopoulos, A., … ENSAT EURINE-ACT Investigators (2022). Cardiometabolic Disease Burden and Steroid Excretion in Benign Adrenal Tumors : A Cross-Sectional Multicenter Study. Annals of internal medicine, 175(3), 325–334. https://doi.org/10.7326/M21-1737
11	Lee, S. I., Cooper, J., Fenton, A., **Subramanian, A.,** Taverner, T., Gokhale, K. M., Phillips, K., Patel, M., Harper, L., Thomas, G. N., & Nirantharakumar, K. (2022). Decreased renal function is associated with incident dementia: An IMRD-THIN retrospective cohort study in the UK. Alzheimer's & dementia : the journal of the Alzheimer's Association, 10.1002/alz.12539. Advance online publication. https://doi.org/10.1002/alz.12539
12	Lee, S. I., Eastwood, K. A., Moss, N., Azcoaga-Lorenzo, A., **Subramanian, A.,** Anand, A., Taylor, B., Nelson-Piercy, C., Yau, C., McCowan, C., O'Reilly, D., Hope, H., Kennedy, J. I., Abel, K. M., Locock, L., Brocklehurst, P., Plachcinski, R., Brophy, S., Agrawal, U., Thangaratinam, S., … Black, M. (2021). Protocol for the development of a core outcome set for studies of pregnant women with pre-existing multimorbidity. BMJ open, 11(10), e044919. https://doi.org/10.1136/bmjopen-2020-044919
13	Chandan, J. S., **Subramanian, A.,** Chandan, J. K., Gokhale, K. M., Vitoc, A., Taylor, J., Bradbury-Jones, C., Bandyopadhyay, S., & Nirantharakumar, K. (2021). The risk of COVID-19 in survivors of domestic violence and abuse. BMC medicine, 19(1), 246. https://doi.org/10.1186/s12916-021-02119-w
14	Bhachu, H. K., Cockwell, P., **Subramanian, A**., Adderley, N. J., Gokhale, K., Fenton, A., Kyte, D., Nirantharakumar, K., & Calvert, M. (2021). Impact of Using Risk-Based Stratification on Referral of Patients With Chronic Kidney Disease From Primary Care to Specialist Care in the United Kingdom. Kidney international reports, 6(8), 2189–2199. https://doi.org/10.1016/j.ekir.2021.05.031
15	Phillips, K., **Subramanian, A.,** Thomas, G. N., Khan, N., Chandan, J. S., Brady, P., Marshall, T., Nirantharakumar, K., Fabritz, L., & Adderley, N. J. (2022). Trends in the pharmacological management of atrial fibrillation in UK general practice 2008-2018.

	Heart (British Cardiac Society), 108(7), 517–522. https://doi.org/10.1136/heartjnl-2021-319338
16	Parry-Smith, W., Šumilo, D., **Subramanian, A.,** Gokhale, K., Okoth, K., Gallos, I., Coomarasamy, A., & Nirantharakumar, K. (2021). Postpartum haemorrhage and risk of long-term hypertension and cardiovascular disease: an English population-based longitudinal study using linked primary and secondary care databases. BMJ open, 11(5), e041566. https://doi.org/10.1136/bmjopen-2020-041566
17	Parretti, H. M., **Subramanian, A.,** Adderley, N. J., Abbott, S., Tahrani, A. A., & Nirantharakumar, K. (2021). Post-bariatric surgery nutritional follow-up in primary care: a population-based cohort study. The British journal of general practice : the journal of the Royal College of General Practitioners, 71(707), e441–e449. https://doi.org/10.3399/bjgp20X714161
18	Parry-Smith, W., Okoth, K., **Subramanian, A.,** Gokhale, K. M., Chandan, J. S., Humpston, C., Coomarasamy, A., Nirantharakumar, K., & Šumilo, D. (2021). Postpartum haemorrhage and risk of mental ill health: A population-based longitudinal study using linked primary and secondary care databases. Journal of psychiatric research, 137, 419–425. https://doi.org/10.1016/j.jpsychires.2021.03.022
19	Braithwaite, T., Adderley, N. J., **Subramanian, A.,** Galloway, J., Kempen, J. H., Gokhale, K., Cope, A. P., Dick, A. D., Nirantharakumar, K., & Denniston, A. K. (2021). Epidemiology of Scleritis in the United Kingdom From 1997 to 2018: Population-Based Analysis of 11 Million Patients and Association Between Scleritis and Infectious and Immune-Mediated Inflammatory Disease. Arthritis & rheumatology (Hoboken, N.J.), 73(7), 1267–1276. https://doi.org/10.1002/art.41709
20	Haroon, S., **Subramanian, A.,** Cooper, J., Anand, A., Gokhale, K., Byne, N., Dhalla, S., Acosta-Mena, D., Taverner, T., Okoth, K., Wang, J., Chandan, J. S., Sainsbury, C., Zemedikun, D. T., Thomas, G. N., Parekh, D., Marshall, T., Sapey, E., Adderley, N. J., & Nirantharakumar, K. (2021). Renin-angiotensin system inhibitors and susceptibility to COVID-19 in patients with hypertension: a propensity score-matched cohort study in primary care. BMC infectious diseases, 21(1), 262. https://doi.org/10.1186/s12879-021-05951-w
21	Wang, J., Cooper, J. M., Gokhale, K., Acosta-Mena, D., Dhalla, S., Byne, N., Chandan, J. S., Anand, A., Okoth, K., **Subramanian, A.,** Bangash, M. N., Jackson, T., Zemedikun, D., Taverner, T., Hanif, W., Ghosh, S., Narendran, P., Toulis, K. A., Tahrani, A. A., Surenthirakumaran, R., … Nirantharakumar, K. (2021). Association of Metformin with Susceptibility to COVID-19 in People with Type 2 Diabetes. The Journal of clinical endocrinology and metabolism, 106(5), 1255–1268. https://doi.org/10.1210/clinem/dgab067
22	Riley, J., Antza, C., Kempegowda, P., **Subramanian, A.,** Chandan, J. S., Gokhale, K., Thomas, N., Sainsbury, C., Tahrani, A. A., & Nirantharakumar, K. (2021). Social Deprivation and Incident Diabetes-Related Foot Disease in Patients With Type 2 Diabetes: A Population-Based Cohort Study. Diabetes care, 44(3), 731–739. https://doi.org/10.2337/dc20-1027

23	Singh, P., Adderley, N., **Subramanian, A.,** Gokhale, K., Singhal, R., Toulis, K. A., Bellary, S., Nirantharakumar, K., & Tahrani, A. A. (2021). The Impact of Bariatric Surgery on Incident Microvascular Complications in Patients With Type 2 Diabetes: A Matched Controlled Population-Based Retrospective Cohort Study. Diabetes care, 44(1), 116–124. https://doi.org/10.2337/dc20-0571
24	Chandan, J. S., Zemedikun, D. T., Thayakaran, R., Byne, N., Dhalla, S., Acosta-Mena, D., Gokhale, K. M., Thomas, T., Sainsbury, C., **Subramanian, A.,** Cooper, J., Anand, A., Okoth, K. O., Wang, J., Adderley, N. J., Taverner, T., Denniston, A. K., Lord, J., Thomas, G. N., Buckley, C. D., … Haroon, S. (2021). Nonsteroidal Antiinflammatory Drugs and Susceptibility to COVID-19. Arthritis & rheumatology (Hoboken, N.J.), 73(5), 731–739. https://doi.org/10.1002/art.41593
25	**Subramanian, A.,** Idkowiak, J., Toulis, K. A., Thangaratinam, S., Arlt, W., & Nirantharakumar, K. (2021). Pubertal timing in boys and girls born to mothers with gestational diabetes mellitus: a systematic review. European journal of endocrinology, 184(1), 51–64. https://doi.org/10.1530/EJE-20-0296
26	Braithwaite, T., **Subramanian, A**., Petzold, A., Galloway, J., Adderley, N. J., Mollan, S. P., Plant, G. T., Nirantharakumar, K., & Denniston, A. K. (2020). Trends in Optic Neuritis Incidence and Prevalence in the UK and Association With Systemic and Neurologic Disease. JAMA neurology, 77(12), 1514–1523. https://doi.org/10.1001/jamaneurol.2020.3502
27	Sainsbury, C., Wang, J., Gokhale, K., Acosta-Mena, D., Dhalla, S., Byne, N., Chandan, J. S., Anand, A., Cooper, J., Okoth, K., **Subramanian, A.,** Bangash, M. N., Taverner, T., Hanif, W., Ghosh, S., Narendran, P., Cheng, K. K., Marshall, T., Gkoutos, G., Toulis, K., … Nirantharakumar, K. (2021). Sodium-glucose co-transporter-2 inhibitors and susceptibility to COVID-19: A population-based retrospective cohort study. Diabetes, obesity & metabolism, 23(1), 263–269. https://doi.org/10.1111/dom.14203
28	**Subramanian, A.,** Adderley, N. J., Gkoutos, G. V., Gokhale, K. M., Nirantharakumar, K., & Krishna, M. T. (2021). Ethnicity-based differences in the incident risk of allergic diseases and autoimmune disorders: A UK-based retrospective cohort study of 4.4 million participants. Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology, 51(1), 144–147. https://doi.org/10.1111/cea.13741
29	Adderley, N. J., **Subramanian, A.,** Toulis, K., Gokhale, K., Taverner, T., Hanif, W., Haroon, S., Thomas, G. N., Sainsbury, C., Tahrani, A. A., & Nirantharakumar, K. (2020). Obstructive Sleep Apnea, a Risk Factor for Cardiovascular and Microvascular Disease in Patients With Type 2 Diabetes: Findings From a Population-Based Cohort Study. Diabetes care, 43(8), 1868–1877. https://doi.org/10.2337/dc19-2116
30	King, D., Reulen, R. C., Thomas, T., Chandan, J. S., Thayakaran, R., **Subramanian, A.,** Gokhale, K., Bhala, N., Nirantharakumar, K., Adderley, N. J., & Trudgill, N. (2020). Changing patterns in the epidemiology and outcomes of inflammatory bowel disease in the United Kingdom: 2000-2018. Alimentary pharmacology & therapeutics, 51(10), 922–934. https://doi.org/10.1111/apt.15701

31	Thomas, T., Chandan, J. S., **Subramanian, A**., Gokhale, K., Gkoutos, G., Harper, L., Buckley, C., Chandratre, P., Raza, K., Situnayake, D., & Nirantharakumar, K. (2020). Epidemiology, morbidity and mortality in Behçet's disease: a cohort study using The Health Improvement Network (THIN). Rheumatology (Oxford, England), 59(10), 2785–2795. https://doi.org/10.1093/rheumatology/keaa010
32	Krishna, M. T., **Subramanian, A.,** Adderley, N. J., Zemedikun, D. T., Gkoutos, G. V., & Nirantharakumar, K. (2019). Allergic diseases and long-term risk of autoimmune disorders: longitudinal cohort study and cluster analysis. The European respiratory journal, 54(5), 1900476. https://doi.org/10.1183/13993003.00476-2019
33	Lavrentaki, A., Thomas, T., **Subramanian, A.,** Valsamakis, G., Thomas, N., Toulis, K. A., Wang, J., Daly, B., Saravanan, P., Sumilo, D., Mastorakos, G., Tahrani, A. A., & Nirantharakumar, K. (2019). Increased risk of non-alcoholic fatty liver disease in women with gestational diabetes mellitus: A population-based cohort study, systematic review and meta-analysis. Journal of diabetes and its complications, 33(10), 107401. https://doi.org/10.1016/j.jdiacomp.2019.06.006
34	Adderley, N. J., **Subramanian, A.,** Nirantharakumar, K., Yiangou, A., Gokhale, K. M., Mollan, S. P., & Sinclair, A. J. (2019). Association Between Idiopathic Intracranial Hypertension and Risk of Cardiovascular Diseases in Women in the United Kingdom. JAMA neurology, 76(9), 1088–1098. https://doi.org/10.1001/jamaneurol.2019.1812
35	Bhachu, H. K., Cockwell, P., **Subramanian, A.,** Nirantharakumar, K., Kyte, D., & Calvert, M. (2019). Cross-sectional observation study to investigate the impact of risk-based stratification on care pathways for patients with chronic kidney disease: protocol paper. BMJ open, 9(6), e027315. https://doi.org/10.1136/bmjopen-2018-027315
36	**Subramanian, A.,** Adderley, N. J., Tracy, A., Taverner, T., Hanif, W., Toulis, K. A., Thomas, G. N., Tahrani, A. A., & Nirantharakumar, K. (2019). Risk of Incident Obstructive Sleep Apnea Among Patients With Type 2 Diabetes. Diabetes care, 42(5), 954–963. https://doi.org/10.2337/dc18-2004
37	Tracy, A., **Subramanian, A.,** Adderley, N. J., Harper, L., & Nirantharakumar, K. (2020). Response to 'IgA vasculitis in adults: few certainties and many uncertainties' by Hočevar et al. Annals of the rheumatic diseases, 79(4), e48. https://doi.org/10.1136/annrheumdis-2018-214951
38	Fountas, A., Lavrentaki, A., **Subramanian, A.,** Toulis, K. A., Nirantharakumar, K., & Karavitaki, N. (2018). Recurrence in silent corticotroph adenomas after primary treatment: A systematic review and meta-analysis. The Journal of clinical endocrinology and metabolism, 10.1210/jc.2018-01956. Advance online publication. https://doi.org/10.1210/jc.2018-01956
39	Tracy, A., **Subramanian, A.,** Adderley, N. J., Cockwell, P., Ferro, C., Ball, S., Harper, L., & Nirantharakumar, K. (2019). Cardiovascular, thromboembolic and renal outcomes in IgA vasculitis (Henoch-Schönlein purpura): a retrospective cohort study using routinely collected primary care data. Annals of the rheumatic diseases, 78(2), 261–269. https://doi.org/10.1136/annrheumdis-2018-214142

40	O'Reilly, M. W., Glisic, M., Kumarendran, B., **Subramanian, A.,** Manolopoulos, K. N., Tahrani, A. A., Keerthy, D., Muka, T., Toulis, K. A., Hanif, W., Thomas, G. N., Franco, O. H., Arlt, W., & Nirantharakumar, K. (2019). Serum testosterone, sex hormone-binding globulin and sex-specific risk of incident type 2 diabetes in a retrospective primary care cohort. Clinical endocrinology, 90(1), 145–154. https://doi.org/10.1111/cen.13862

Supplementary 2: Read code lists for diagnosis of polycystic ovary syndrome (PCOS)

READ CODE	DESCRIPTION
C165.00	Polycystic ovarian syndrome
C164.12	Stein - Leventhal syndrome

Supplementary 3: Read code lists for hair loss

READ CODE	DESCRIPTION
M240000	Alopecia unspecified
M240.00	Alopecia
M240012	Hair loss
22D7.11	O/E - alopecia
M240z00	Alopecia NOS
M240200	Male pattern alopecia
1N02.00	C/O: hair loss
M240300	Frontal alopecia of women
22D4.00	O/E - loss of hair
M240H00	Alopecia seborrhoeica
Myu6300	[X]Other androgenic alopecia
M240D00	Marginal alopecia

Supplementary 4: Read code lists for hirsutism

READ CODE	DESCRIPTION
M241.00	Hirsutism - hypertrichosis
22D8000	O/E - facial hair
22D8.00	O/E - hirsutism

Supplementary 5: Read code lists for acne

READ CODE	DESCRIPTION
M261000	Acne vulgaris
M261600	Cystic acne
M261.00	Other acne
M261A00	Pustular acne
M261100	Acne conglobata
M261H00	Acne keloid
M261E00	Acne excoriee des jeunes filles
M261z00	Other acne NOS
M25y600	Acne keloid
2FG5.00	Acne scar
M261X00	Acne, unspecified

Myu6F00	[X]Acne, unspecified
M261J00	Acne necrotica
Myu6800	[X]Other acne
M261F00	Acne fulminans
M261K00	Acne keloidalis
679g000	Acne management education

Supplementary 6: Read code lists for anovulation

READ CODE	DESCRIPTION
K590.11	Amenorrhoea
K591100	Oligomenorrhoea
K590100	Secondary amenorrhoea
K5B0.00	Female infertility of anovulatory origin
K590.00	Absence of menstruation
K591300	Secondary oligomenorrhoea
K591000	Hypomenorrhoea
K5B0.11	Anovular cycle
1571.00	H/O: amenorrhoea
K591.00	Scanty or infrequent menstruation
K591.11	Infrequent menstruation
K591z00	Scanty or infrequent menstruation NOS
K590z00	Amenorrhoea NOS
K5B0100	Secondary anovulatory infertility
K5B0z00	Female infertility of anovulatory origin NOS

Supplementary 7: Read code lists for Polycystic ovaries (PCO)

READ CODE	DESCRIPTION
C164.00	Polycystic ovaries
7E25211	Laparoscopic drainage ovarian cyst
7E25200	Endoscopic drainage of cyst of ovary
C164.13	Multicystic ovaries
7E25300	Endoscopic drilling of ovary
C164.11	Isosexual virilisation
ZV13G00	[V]Personal history of ovarian cyst

Supplementary 8: Read code lists for white ethnicity

READ CODE	DESCRIPTION
134B.00	RACE: Caucasian
9i0..00	British or mixed British - ethnic category 2001 census
9i20.00	English - ethnic category 2001 census
9i27.00	Greek - ethnic category 2001 census
9i2R.00	Oth White European/European unsp/Mixed European 2001 census
9i2B.00	Italian - ethnic category 2001 census
9i2..00	Other White background - ethnic category 2001 census
9i2G.00	Baltic Estonian/Latvian/Lithuanian - ethn categ 2001 census
9i21.00	Scottish - ethnic category 2001 census
9S12.00	Other white ethnic group
9S10.00	White British
9i2F.00	Polish - ethnic category 2001 census
9i1..00	Irish - ethnic category 2001 census
9SAC.00	Other European (NMO)
9i22.00	Welsh - ethnic category 2001 census
9i29.00	Turkish - ethnic category 2001 census
9i28.00	Greek Cypriot - ethnic category 2001 census
9S1..00	White
9SA9.00	Irish (NMO)
9S11.00	White Irish
2261	O/E - Europeanoid
9i2K.00	Albanian - ethnic category 2001 census
134N.00	RACE: White
1341	European origin
9S14.00	Other white British ethnic group
9i2J.00	Kosovan - ethnic category 2001 census
9i2Q.00	Mixed Irish and other White - ethnic category 2001 census
9S13.00	White Scottish
9i2M.00	Croatian - ethnic category 2001 census
9i23.00	Cornish - ethnic category 2001 census
9i2P.00	Other republics former Yugoslavia - ethnic categ 2001 census
9i2H.00	Commonwealth (Russian) Indep States - ethn categ 2001 census
9SAB.00	Turkish/Turkish Cypriot (NMO)
9SAB.12	Turkish Cypriot (NMO)
9SAB.11	Turkish (NMO)
9i2A.00	Turkish Cypriot - ethnic category 2001 census
9i26.00	Cypriot (part not stated) - ethnic category 2001 census
9T2..00	Traveller - gypsy
9i25.00	Ulster Scots - ethnic category 2001 census
9i2E.00	Gypsy/Romany - ethnic category 2001 census
9i24.00	Northern Irish - ethnic category 2001 census
9SAA.00	Greek/Greek Cypriot (NMO)
9SAA.11	Greek (NMO)

9i2L.00	Bosnian - ethnic category 2001 census
9i2N.00	Serbian - ethnic category 2001 census
9SI..00	Irish traveller
9SAA.12	Greek Cypriot (NMO)
9i2D.00	Traveller - ethnic category 2001 census
9i2C.00	Irish Traveller - ethnic category 2001 census
9i00.00	White British - ethnic category 2001 census
9i10.00	White Irish - ethnic category 2001 census
9T5..00	Bulgarian
9T4..00	Romanian
9T6..00	Czech
9T8..00	Portuguese
1X61.00	NSCTSP Scandinavia, Switzerland family origin
9T7..00	Slovak
1X60.00	NSCTSP Austr, Belg, Ire, Fr, Germ, Netherland family origin
1X41.00	NSCTSP Greece, Turkey, Cyprus family origin
1X6..00	NHS Sic Cel Thal Scr Prog fam orig Northern European (white)
1X42.00	NSCTSP Italy, Portugal, Spain family origin
9t03.00	White: other White backgrd- Eng+Wales ethnic cat 2011 census
9t00.00	White:Eng/Welsh/Scot/NI/Brit - England and Wales 2011 census
9t20.00	White: Scottish - Scotland ethnic category 2011 census
9t24.00	White: Polish - Scotland ethnic category 2011 census
9t01.00	White: Irish - England and Wales ethnic category 2011 census
9t22.00	White: Irish - Scotland ethnic category 2011 census
9t21.00	White: other British - Scotland ethnic category 2011 census
9t25.00	White: other White ethnic grp- Scotland ethnic cat 2011 cens
9t02.00	White: Gypsy/Irish Traveller - Eng+Wales eth cat 2011 census
9t10.00	White - Northern Ireland ethnic category 2011 census
1X40.00	NHS Sickle Cell Thal Scr Prog Sardinia family origin
1X5..00	NHS Sick Cel Thal Scr Prog fam origin United Kingdom (white)
9t23.00	White: Gypsy/Irish Traveller - Scotland ethnic cat 2011 cens

Supplementary 9: Read code lists for South Asian ethnicity

READ CODE	DESCRIPTION
9i7..00	Indian or British Indian - ethnic category 2001 census
9i8..00	Pakistani or British Pakistani - ethnic category 2001 census
9S6..00	Indian
9iA4.00	Sri Lankan - ethnic category 2001 census
9iA8.00	British Asian - ethnic category 2001 census
1358.00	Hindu
9iA5.00	Tamil - ethnic category 2001 census
9iA6.00	Sinhalese - ethnic category 2001 census
13Z6500	Language Punjabi
13Z6300	Language Hindi
135B.00	Sikh

9S7..00	Pakistani
13Z6600	Language Urdu
13Z6200	Language Gujurati
9S8..00	Bangladeshi
13Z6100	Language Bengali
13lp.00	Main spoken language Malayalam
13l1.00	Main spoken language Bengali
13lE.00	Main spoken language Punjabi
9T1D.00	Indian
13eW.00	Born in Pakistan
134M.00	RACE: Pakistani
13Z6400	Language Pashtu
134I.00	RACE: Bangladeshi
13lL.00	Main spoken language Urdu
9iA1.00	Punjabi - ethnic category 2001 census
9i9..00	Bangladeshi or British Bangladeshi - ethn categ 2001 census
9iAA.00	Other Asian or Asian unspecified ethnic category 2001 census
13e0.00	Born in Afghanistan
13eD.00	Born in India
13e3.00	Born in Bangladesh
13eT.00	Born in Nepal
13ef.00	Born in Sri Lanka
9iFJ.00	Mauritian/Seychellois/Maldivian/St Helena eth cat 2001census
9iA7.00	Caribbean Asian - ethnic category 2001 census
9SA6.00	E Afric Asian/Indo-Carib (NMO)
9SA7.00	Indian sub-continent (NMO)
1347.00	Indian origin
13l8.00	Main spoken language Hindi
9iA9.00	Mixed Asian - ethnic category 2001 census
9T1B.00	South East Asian
9SA6.11	East African Asian (NMO)
13lK.00	Main spoken language Tamil
13lO.00	Main spoken language Farsi
9iA3.00	East African Asian - ethnic category 2001 census
13n2.00	Reads Punjabi
13n7.00	Reads Urdu
13lJ.00	Main spoken language Sylheti
13lu.00	Main spoken language Sinhala
135V.00	Jainism
13l6.00	Main spoken language Gujerati
9iF8.00	Sikh - ethnic category 2001 census
13b2.00	Sylhety
13b4.00	Mirpuri language
13nD.00	Reads Hindi

13n8.00	Reads Bengali
9iF5.00	Hindu - ethnic category 2001 census
9iA2.00	Kashmiri - ethnic category 2001 census
13n6.00	Reads Tamil
13lA.00	Main spoken language Kutchi
13eQ.00	Born in Maldives
13nc.00	Reads Pashto
13nK.00	Reads Gujarati
13lu.11	Main spoken language Sinhalese
13w1.00	Main spoken language Nepali
13lE.11	Main spoken language Panjabi
13e4.00	Born in Bhutan
9SA6.12	Indo-Caribbean (NMO)
9T9..00	Nepali
135v.00	Radha Soami
13za.00	Sanatana Dharma
13zd.00	Shakti Hindu
1X1..00	Sick Cell Thalas Scr Prog fam orig South Asia (Asian)
1X2..00	NHS Sick Cell Thal Scr Prog fam orig South East Asia (Asian)
13ze.00	Smarta Hindu
1X12.00	NHS Sickle Cell Thal Screening Prog Bangladesh family origin
13zg.00	Arya Samaj Hindu
13zc.00	Shiva Hindu
13l6.11	Main spoken language Gujarati
9t16.00	Asian or Asian British: Indian - NI ethnic cat 2011 census
9t1A.00	Asian/Asian British: other Asian - NI ethnic cat 2011 census
9t27.00	Asian: Pakistani/Pakistani Scot/Pakistani Brit- Scot 2011
9t09.00	Asian/Asian British:Pakistani- Eng+Wales eth cat 2011 census
9t08.00	Asian/Asian Brit: Indian - Eng+Wales ethnic cat 2011 census
9t17.00	Asian/Asian British: Pakistani - NI ethnic cat 2011 census
9t0A.00	Asian/Asian Brit: Bangladeshi- Eng+Wales eth cat 2011 census
9t18.00	Asian/Asian British: Bangladeshi - NI ethnic cat 2011 census
9t0B.00	Asian/Asian Brit: Chinese - Eng+Wales ethnic cat 2011 census
9t28.00	Asian: Indian, Indian Scot/Indian Brit- Scotland 2011 census
13zf.00	Advaitin Hindu
9t0C.00	Asian/Asian Brit: other Asian- Eng+Wales eth cat 2011 census
13ne.00	Reads Sinhala
9t29.00	Bangladeshi, Bangladeshi Scot or Bangladeshi Brit- Scot 2011
9t19.00	Asian/Asian British: Chinese - NI ethnic cat 2011 census
135g.00	Jain

Supplementary 10: Read code lists for Black Afro-Caribbean ethnicity

READ CODE	DESCRIPTION
9iC..00	African - ethnic category 2001 census
9iB..00	Caribbean - ethnic category 2001 census

9iD0.00	Somali - ethnic category 2001 census
9S41.00	Black British
9S2..00	Black Caribbean
9S3..00	Black African
9S4..00	Black, other, non-mixed origin
2262.00	O/E - Negroid
13go.00	Born in Zimbabwe
13gi.00	Born in Tanzania
13gC.00	Born in Congo
9S5..00	Black - other, mixed
134H.00	RACE: Afro-caribbean
13lP.00	Main spoken language Shona
9S48.00	Black Black - other
13gY.00	Born in Niger
13jC.00	Born in Trinidad and Tobago
13gJ.00	Born in Ghana
13gl.00	Born in Uganda
13j6.00	Born in Jamaica
13j2.00	Born in Barbados
134K.00	RACE: West indian
13gM.00	Born in Ivory Coast
9S42.13	Black Guyana
13gZ.00	Born in Nigeria
13gN.00	Born in Kenya
13gS.00	Born in Malawi
13j9.00	Born in St. Lucia
9SG..00	Other black ethnic group
13gG.00	Born in Ethiopia
13gW.00	Born in Mozambique
13g5.00	Born in Burundi
13gn.00	Born in Zambia
13ge.00	Born in Somalia
13gd.00	Born in Sierra Leone
13gI.00	Born in Gambia
13gX.00	Born in Namibia
13gP.00	Born in Liberia
13g7.00	Born in Cameroon
13j0.00	Born in Antigua and Barbuda
13jB.00	Born in Togo
13gc.00	Born in Senegal
13ga.00	Born in Rwanda
13f4.00	Born in British Guyana
13gL.00	Born in Guinea Republic
9iD..00	Other Black background - ethnic category 2001 census

13lG.00	Main spoken language Somali
9iD1.00	Nigerian - ethnic category 2001 census
9S47.00	Black - other Asian
9S44.00	Black - other African country
13gg.00	Born in Sudan
13gR.00	Born in Madagascar
13gA.00	Born in Chad
9iD3.00	Mixed Black - ethnic category 2001 census
9iD2.00	Black British - ethnic category 2001 census
13gm.00	Born in Zaire
13g3.00	Born in Botswana
13g2.00	Born in Benin
9S43.00	Black N African/Arab/Iranian
13gh.00	Born in Swaziland
13gK.00	Born in Guinea Bissau
1342.00	African origin
134A.00	West Indian origin
9iD4.00	Other Black or Black unspecified ethnic category 2001 census
9S43.11	Black North African
13ln.00	Main spoken language Lingala
13ld.00	Main spoken language Amharic
13lI.00	Main spoken language Swahili
13ly.00	Main spoken language Tigrinya
13lm.00	Main spoken language Igbo
9S42.11	Black Caribbean
9S45.00	Black E Afric Asia/Indo-Caribb
9SA5.00	Other African countries (NMO)
9S42.12	Black West Indian
9S46.00	Black Indian sub-continent
9S43.13	Black Iranian
13lM.00	Main spoken language Yoruba
13lc.00	Main spoken language Akan
9SA3.00	Caribbean I./W.I./Guyana (NMO)
9SA3.12	West Indian (NMO)
9SA3.11	Caribbean Island (NMO)
9S42.00	Black Caribbean/W.I./Guyana
9S43.12	Black Arab
9S45.11	Black East African Asian
9S45.12	Black Indo-Caribbean
13j5.00	Born in Haiti
13l7.00	Main spoken language Hausa
13g4.00	Born in Burkina Faso
13n4.00	Reads Somali
13j1.00	Born in Bahamas

13gH.00	Born in Gabon
13gj.00	Born in The Gambia
13ls.00	Main spoken language Patois
13gp.00	Born in Eritrea
13eR.00	Born in Mali
13jA.00	Born in St. Vincent
13j8.00	Born in St. Kitts and Nevis
13gD.00	Born in Djibouti
13lg.00	Main spoken language Ethiopian
13lo.00	Main spoken language Luganda
13fG.00	Born in Nicaragua
9SA3.13	Guyana (NMO)
13nf.00	Reads Tigrinya
13gO.00	Born in Lesotho
13gF.00	Born in Equatorial Guinea
13d1.00	Born in Andorra
13u5.00	Main spoken language Afrikaans
13na.00	Reads Amharic
13le.00	Main spoken language Brawa
13wN.00	Main spoken language Tongan
13u2.00	Main spoken language Oromo
13u4.00	Main spoken language Afar
13wa.00	Main spoken language Zulu
13no.00	Reads Ndebele
13wB.00	Main spoken language Southern Sotho
13t1.00	Born in Bermuda
13um.00	Main spoken language Kinyarwanda
13nN.00	Reads Swahili
13nP.00	Reads Yoruba
13nb.00	Reads Lingala
13g9.00	Born in Central African Republic
13jD.00	Born in Dominica
13gB.00	Born in Comoros Islands
13gb.00	Born in Sao Tome and Principe
1X01.00	NHS Sick Cell Thalassaemia Screen Prog African family origin
13wC.00	Main spoken language Tswana
13v0.00	Born in Martinique
1X0..00	Sick Cell Thalas Scr Prog fam orig African or African-Carib
13t2.00	Born in Anguilla
13nk.00	Reads Kinyarwanda
13t0.00	Born in Montserrat
13jE.00	Born in Aruba
1X00.00	NHS Sick Cell Thal Scr Prog Caribbean Islands family origin
13t5.00	Born in Saint Helena, Ascension and Tristan da Cunha

13jG.00	Born in Saint Vincent and the Grenadines
13t3.00	Born in British Virgin Islands
13gq.00	Born in Democratic Republic of Congo
9t0E.00	Black/African/Caribbn/Black Brit: Caribbean - Eng+Wales 2011
9t0D.00	Black/African/Carib/Black Brit: African- Eng+Wales 2011 cens
9t0F.00	Black/Afr/Carib/Black Brit: other Black- Eng+Wales 2011 cens
9t1B.00	Black/Afri/Carib/Black Brit: African- NI eth cat 2011 census
9t2D.00	African: any other African - Scotland ethnic cat 2011 census
9t1C.00	Black/Afri/Carib/Black Brit: Caribbean- NI eth cat 2011 cens
9t2C.00	African: African/African Scot/African Brit - Scotland 2011
13jH.00	Born in Sint Maarten
9t1D.00	Black/Afri/Carib/Black Brit: other - NI eth cat 2011 census
9t2G.00	Carib/Black: any other Black/Caribbean grp - Scotland 2011
9t2F.00	Carib/Black: Black/Black Scot/Black Brit- Scotland 2011 cens
13nL.00	Reads Hausa
13jF.00	Born in United States Virgin Islands
9t2E.00	Carib/Black: Caribbean/Carib Scot/Carib Brit- Scotland 2011

Supplementary 11: Read code lists for mixed ethnicity

READ CODE	DESCRIPTION
9i4..00	White and Black African - ethnic category 2001 census
9i5..00	White and Asian - ethnic category 2001 census
9SB..00	Other ethnic, mixed origin
9i63.00	Chinese and White - ethnic category 2001 census
9i3..00	White and Black Caribbean - ethnic category 2001 census
9i60.00	Black and Asian - ethnic category 2001 census
9i6..00	Other Mixed background - ethnic category 2001 census
9S51.00	Other Black - Black/White orig
134J.00	RACE: Mixed
9S52.00	Other Black - Black/Asian orig
9SB2.00	Other ethnic, Asian/White orig
9i65.00	Other Mixed or Mixed unspecified ethnic category 2001 census
9SB4.00	Other ethnic, other mixed orig
9SB5.00	Black Caribbean and White
9SB6.00	Black African and White
9SB3.00	Other ethnic, mixed white orig
9i62.00	Black and White - ethnic category 2001 census
134L.00	RACE: Afro-caucasian
9i64.00	Asian and Chinese - ethnic category 2001 census
9SB1.00	Other ethnic, Black/White orig
9i61.00	Black and Chinese - ethnic category 2001 census
9t05.00	Mixed: White+Black African - Eng+Wales eth cat 2011 census
9t04.00	Mixed: White+Black Caribbean - Eng+Wales eth cat 2011 census
9t14.00	Mixed: White and Asian - NI ethnic category 2011 census
9t15.00	Mixed: other Mixed/multiple ethnic backgrd - NI 2011 census

9t13.00	Mixed: White and Black African - NI ethnic cat 2011 census
9t06.00	Mixed: White+Asian - Eng+Wales ethnic category 2011 census
9t07.00	Mixed: other Mixed/multiple backgrd - Eng+Wales 2011 census
9t12.00	Mixed: White and Black Caribbean - NI ethnic cat 2011 census
9t26.00	Mixed/multiple ethnic grps: any- Scot ethnic cat 2011 census

Supplementary 12: Read code lists for other ethnic minority groups

READ_CODE	DESCRIPTION
9iF2.00	Filipino - ethnic category 2001 census
9iF..00	Other - ethnic category 2001 census
9iE..00	Chinese - ethnic category 2001 census
9iF1.00	Japanese - ethnic category 2001 census
9iA..00	Other Asian background - ethnic category 2001 census
13eG.00	Born in Iraq
9T1C.00	Chinese
9iF0.00	Vietnamese - ethnic category 2001 census
9iF3.00	Malaysian - ethnic category 2001 census
9iFH.00	South and Central American - ethnic category 2001 census
134C.00	RACE: Arab
9S9..00	Chinese
13eH.00	Born in Israel
9SA4.12	Iranian (NMO)
13eY.00	Born in Philippines
9T18.00	Tokelauan
13b0.00	Vietnamese language
9SC..00	Vietnamese
13l2.00	Main spoken language Cantonese
13lx.00	Main spoken language Thai
9iFD.00	Iranian - ethnic category 2001 census
9iFG.00	Latin American - ethnic category 2001 census
13eo.00	Born in Vietnam
13lb.00	Main spoken language Vietnamese
13eF.00	Born in Iran
13eg.00	Born in Syria
9iFB.00	Mid East (excl Israeli, Iranian & Arab) - eth cat 2001 cens
13e8.00	Born in China
13eI.00	Born in Japan
13f3.00	Born in Brazil
13g0.00	Born in Algeria
13fN.00	Born in Venezuela
1345	South American origin
13gV.00	Born in Morocco
13fB.00	Born in Grenada
13eM.00	Born in Kyrgyzstan
13ej.00	Born in Thailand
13ec.00	Born in Saudi Arabia

13ed.00	Born in Singapore
13fJ.00	Born in Peru
13eP.00	Born in Malaysia
13fF.00	Born in Mexico
13eC.00	Born in Hong Kong
134D.00	RACE: Chinese
13e6.00	Born in Burma
13g1.00	Born in Angola
13e2.00	Born in Bahrain
13k4.00	Born in Seychelles
13f7.00	Born in Columbia
13f9.00	Born in Ecuador
13f0.00	Born in Argentina
13gE.00	Born in Egypt
13e1.00	Born in Armenia
13j4.00	Born in Dominican Republic
13eh.00	Born in Taiwan
9T13.00	New Zealand Maori
13gk.00	Born in Tunisia
13lB.00	Main spoken language Mandarin
13lW.00	Main spoken language Japanese
13fD.00	Born in Guyana
134G.00	RACE: Oriental
13eO.00	Born in Lebanon
13eL.00	Born in Kuwait
13ee.00	Born in South Korea
13f2.00	Born in Bolivia
13fE.00	Born in Honduras
13eE.00	Born in Indonesia
13j3.00	Born in Cuba
13f6.00	Born in Chile
13eK.00	Born in Kazakhstan
13eX.00	Born in Palestine
13el.00	Born in Turkmenistan
13eJ.00	Born in Jordan
13d3.00	Born in Azerbaijan
13gQ.00	Born in Libya
13ep.00	Born in Yemen
13eS.00	Born in Mongolia
13e9.00	Born in Democratic People's Republic of Korea
13ea.00	Born in Republic of Korea
13fM.00	Born in Uruguay
13em.00	Born in United Arab Emirates
134F.00	RACE: Korean
13fI.00	Born in Paraguay
13en.00	Born in Uzbekistan

13k0.00	Born in Fiji
13eZ.00	Born in Qatar
9T1..00	New Zealand ethnic groups
134E.00	RACE: Japanese
9iFE.00	Kurdish - ethnic category 2001 census
9iF9.00	Arab - ethnic category 2001 census
9T1A.00	Other Pacific ethnic group
9iFC.00	Israeli - ethnic category 2001 census
9SA4.11	North African Arab (NMO)
13k6.00	Born in Tonga
13lw.00	Main spoken language Tagalog
1348.00	Middle Eastern origin
1349.00	Far Eastern origin
13k5.00	Born in Solomon Islands
13eV.00	Born in Oman
13lX.00	Main spoken language Korean
9T17.00	Niuean
13eA.00	Born in East Timor
13k7.00	Born in Tuvalu
13b1.00	Cantonese Chinese dialect
9T16.00	Tongan
13e5.00	Born in Brunei
9iF4.00	Buddhist - ethnic category 2001 census
13fH.00	Born in Panama
13fA.00	Born in El Salvador
13k3.00	Born in Papua New Guinea
9T19.00	Fijian
9T14.00	Samoan
13f8.00	Born in Costa Rica
2263.00	O/E - Mongoloid origin
13j7.00	Born in Puerto Rico
13fK.00	Born in Suriname
13eN.00	Born in Laos
13fC.00	Born in Guatemala
9T1Z.00	New Zealand ethnic group NOS
9T11.11	Pakeha
9T15.00	Cook Island Maori
13lk.00	Main spoken language Hakka
13f1.00	Born in Belize
9T3..00	Yemeni
13wR.00	Main spoken language Twi
13u6.00	Main spoken language Armenian
13w5.00	Main spoken language Quechua
13uv.00	Main spoken language Maltese
13uu.00	Main spoken language Malay
13uG.00	Main spoken language Burmese

13g6.00	Born in Cambodia
13wT.00	Main spoken language Uzbek
9T1Y.00	Other New Zealand ethnic group
13uz.00	Main spoken language Mongolian
13wQ.00	Main spoken language Turkmen
13u9.00	Main spoken language Azerbaijani
13wH.00	Main spoken language Sundanese
13uc.00	Main spoken language Indonesian
13w6.00	Main spoken language Romansh
13uB.00	Main spoken language Basque
13ul.00	Main spoken language Kazakh
13wP.00	Main spoken language Tsonga
13uP.00	Main spoken language Esperanto
13up.00	Main spoken language Lao
13uw.00	Main spoken language Maori
13w4.00	Main spoken language Filipino
13uY.00	Main spoken language Kalaallisut
13u8.00	Main spoken language Aymara
13wS.00	Main spoken language Uighur
13ei.00	Born in Tajikistan
13ui.00	Main spoken language Javanese
13uS.00	Main spoken language Fijian
13ug.00	Main spoken language Inuktitut
13uW.00	Main spoken language Galician
13uJ.00	Main spoken language Central Khmer
13wA.00	Main spoken language Dari
1X22.00	NSCTSP Malaysia, Vietnam, Philippin, Cambodia, Laos fam orig
13ut.00	Main spoken language Malagasy
13uA.00	Main spoken language Bashkir
1X31.00	NHS Sickle Cell Thal Scr Program South America family origin
1X3..00	NHS Sickle Cell Thal Scr Prog fam origin other non-European
13w7.00	Main spoken language Samoan
13k9.00	Born in Western Samoa
13uq.00	Main spoken language Bamun
13wV.00	Main spoken language Tetum
13wJ.00	Main spoken language Tajik
1X21.00	N Sic Cell Th Scr P Thailand, Indonesia, Burma family origin
13kB.00	Born in American Samoa
9t0G.00	Other ethnic group: Arab - Eng+Wales ethnic cat 2011 census
9t1F.00	Other ethnic group: any other grp- NI ethnic cat 2011 census
9t0H.00	Other ethnic: any other grp - Eng+Wales eth cat 2011 census
9t1E.00	Other ethnic group: Arab - NI ethnic category 2011 census
9t2B.00	Asian: other Asian group - Scotland ethnic cat 2011 census
13v4.00	Born in French Guiana
9t2A.00	Asian: Chinese - Scotland ethnic category 2011 census
13k8.00	Born in Vanuatu

13wZ.00	Main spoken language Zhuang
9t2J.00	Other ethnic grp: any other ethnic grp- Scotland 2011 census
13w8.00	Main spoken language Sango
9t2H.00	Other ethnic grp: Arab/Arab Scot/Arab British- Scotland 2011
13wK.00	Main spoken language Tatar
13uq.11	Main spoken language Bamoun
13uE.00	Main spoken language Bislama

Supplementary 13: Incidence trend of PCOS based on diagnostic code and a combination of symptom codes fulfilling each of the diagnostic criteria

Year	Incidence per 100,000 person-years based on each of the PCOS Definition											
	Diagnostic code			National Institute of Child Health and Human Development criteria			Rotterdam criteria			Androgen Excess Society criteria		
	Nr	Dr	Incidence (95% CI)	Nr	Dr	Incidence (95% CI)	Nr	Dr	Incidence (95% CI)	Nr	Dr	Incidence (95% CI)
1995	7	332147	2.11 (1.95-2.26)	713	329538	216.36 (214.96-217.77)	855	328933	259.93 (258.43-261.43)	772	329300	234.44 (232.99-235.88)
1996	13	355013	3.66 (3.46-3.86)	838	351683	238.28 (236.88-239.69)	1034	350878	294.69 (293.18-296.2)	906	351373	257.85 (256.4-259.29)
1997	11	400457	2.75 (2.58-2.91)	950	396116	239.83 (238.5-241.16)	1129	395037	285.8 (284.39-287.2)	1026	395694	259.29 (257.93-260.66)
1998	12	473298	2.54 (2.39-2.68)	1192	467562	254.94 (253.69-256.19)	1409	466135	302.27 (300.96-303.59)	1287	466955	275.62 (274.33-276.9)
1999	26	545985	4.76 (4.58-4.94)	1252	538800	232.37 (231.24-233.5)	1548	536944	288.3 (287.09-289.51)	1400	537996	260.23 (259.05-261.4)
2000	39	674184	5.78 (5.6-5.97)	1511	665170	227.16 (226.15-228.17)	1920	662715	289.72 (288.63-290.81)	1735	664061	261.27 (260.21-262.33)
2001	74	843424	8.77 (8.57-8.97)	1937	832380	232.71 (231.8-233.61)	2496	828991	301.09 (300.1-302.08)	2238	830749	269.4 (268.44-270.35)
2002	91	959638	9.48 (9.29-9.68)	2146	947027	226.6 (225.76-227.45)	2945	942715	312.4 (311.46-313.33)	2603	944873	275.49 (274.59-276.39)
2003	149	1118581	13.32 (13.11-13.53)	2656	1103972	240.59 (239.79-241.38)	3586	1098481	326.45 (325.57-327.33)	3243	1101139	294.51 (293.66-295.36)
2004	964	1275000	75.61 (75.15-76.07)	3069	1258751	243.81 (243.06-244.56)	3939	1252100	314.59 (313.78-315.4)	3595	1255211	286.41 (285.62-287.2)
2005	1152	1405188	81.98 (81.53-82.44)	3413	1388011	245.89 (245.18-246.61)	4286	1380316	310.51 (309.74-311.28)	3939	1383835	284.64 (283.89-285.4)

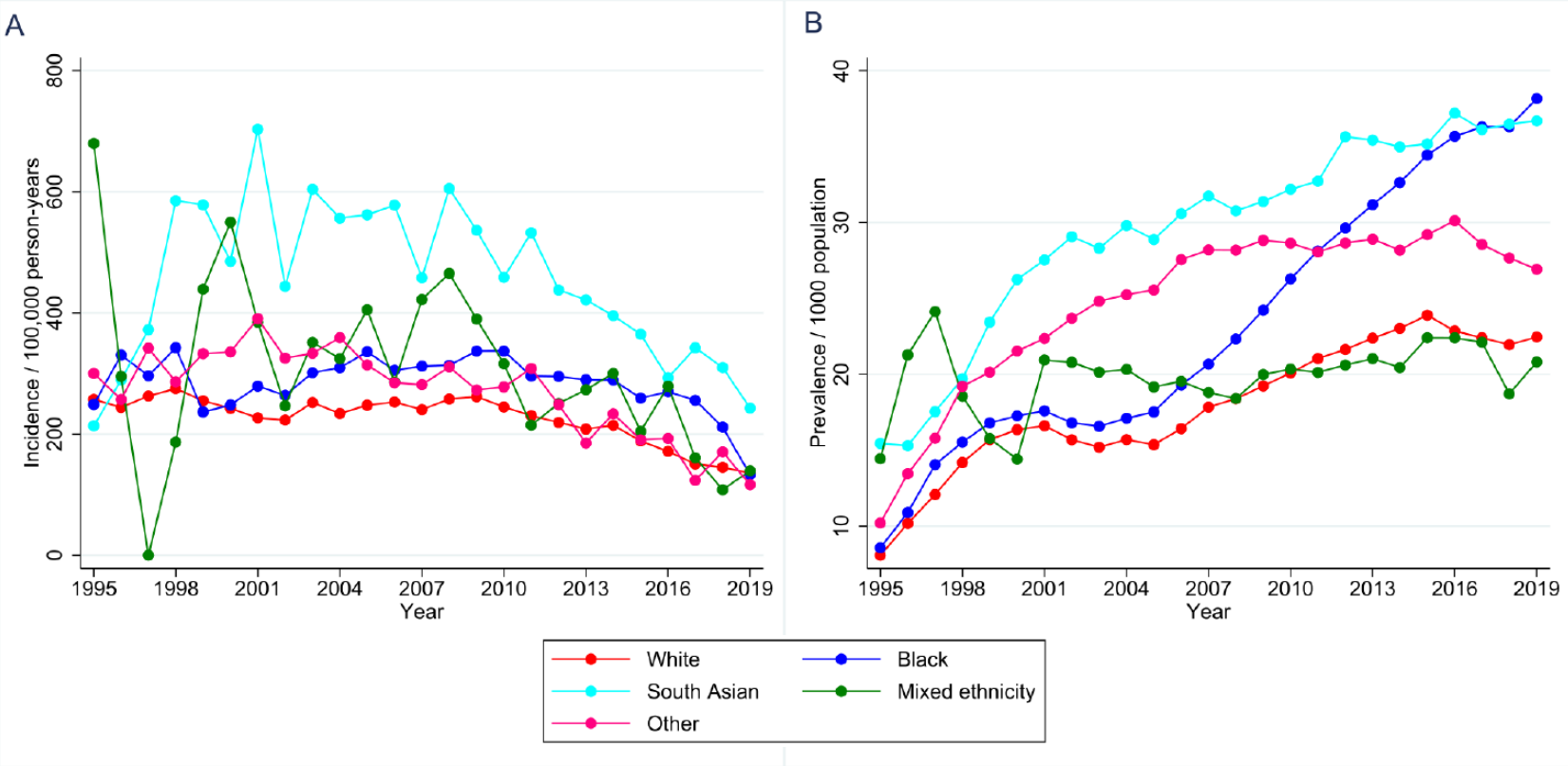
2006	1194	1437968	83.03 (82.58-83.48)	3423	1420186	241.02 (240.32-241.73)	4331	1411766	306.78 (306.02-307.54)	4013	1415521	283.5 (282.76-284.24)
2007	1038	1451159	71.53 (71.11-71.95)	3334	1432911	232.67 (231.98-233.37)	4219	1423833	296.31 (295.56-297.06)	3886	1427805	272.17 (271.44-272.9)
2008	1080	1471566	73.39 (72.97-73.81)	3615	1452510	248.88 (248.18-249.58)	4602	1442654	319 (318.24-319.76)	4265	1446866	294.78 (294.03-295.52)
2009	1071	1469365	72.89 (72.47-73.31)	3604	1449626	248.62 (247.91-249.32)	4712	1439046	327.44 (326.67-328.21)	4383	1443402	303.66 (302.91-304.41)
2010	1131	1449661	78.02 (77.58-78.45)	3398	1429639	237.68 (236.98-238.38)	4495	1418389	316.91 (316.14-317.67)	4123	1422902	289.76 (289.01-290.51)
2011	1147	1413818	81.13 (80.68-81.58)	3224	1393942	231.29 (230.59-231.99)	4373	1382060	316.41 (315.64-317.19)	3989	1386754	287.65 (286.9-288.4)
2012	1050	1391214	75.47 (75.03-75.91)	3063	1371519	223.33 (222.63-224.03)	4060	1359012	298.75 (297.98-299.52)	3709	1363882	271.94 (271.2-272.69)
2013	1036	1321510	78.4 (77.94-78.85)	2770	1302845	212.61 (211.91-213.31)	3726	1290178	288.8 (288.02-289.58)	3454	1295022	266.71 (265.95-267.48)
2014	906	1205547	75.15 (74.68-75.62)	2538	1188827	213.49 (212.75-214.22)	3471	1176692	294.98 (294.16-295.8)	3210	1181214	271.75 (270.95-272.56)
2015	810	1069132	75.76 (75.26-76.26)	1973	1054697	187.07 (186.32-187.81)	2717	1043844	260.29 (259.45-261.13)	2494	1047855	238.01 (237.19-238.83)
2016	666	912077	73.02 (72.49-73.55)	1623	900792	180.17 (179.38-180.97)	2274	891573	255.05 (254.15-255.96)	2107	894923	235.44 (234.56-236.32)
2017	612	820062	74.63 (74.06-75.2)	1294	810515	159.65 (158.85-160.45)	1861	802144	232 (231.08-232.93)	1708	805138	212.14 (211.24-213.03)
2018	511	765003	66.8 (66.24-67.36)	1147	756529	151.61 (150.81-152.42)	1622	748645	216.66 (215.73-217.59)	1514	751416	201.49 (200.58-202.39)
2019	537	708420	75.8 (75.19-76.42)	974	700811	138.98 (138.17-139.79)	1390	693417	200.46 (199.51-201.4)	1287	695978	184.92 (184.01-185.83)

Supplementary 14: Prevalence trend of PCOS based on diagnostic code and a combination of symptom codes fulfilling each of the diagnostic criteria

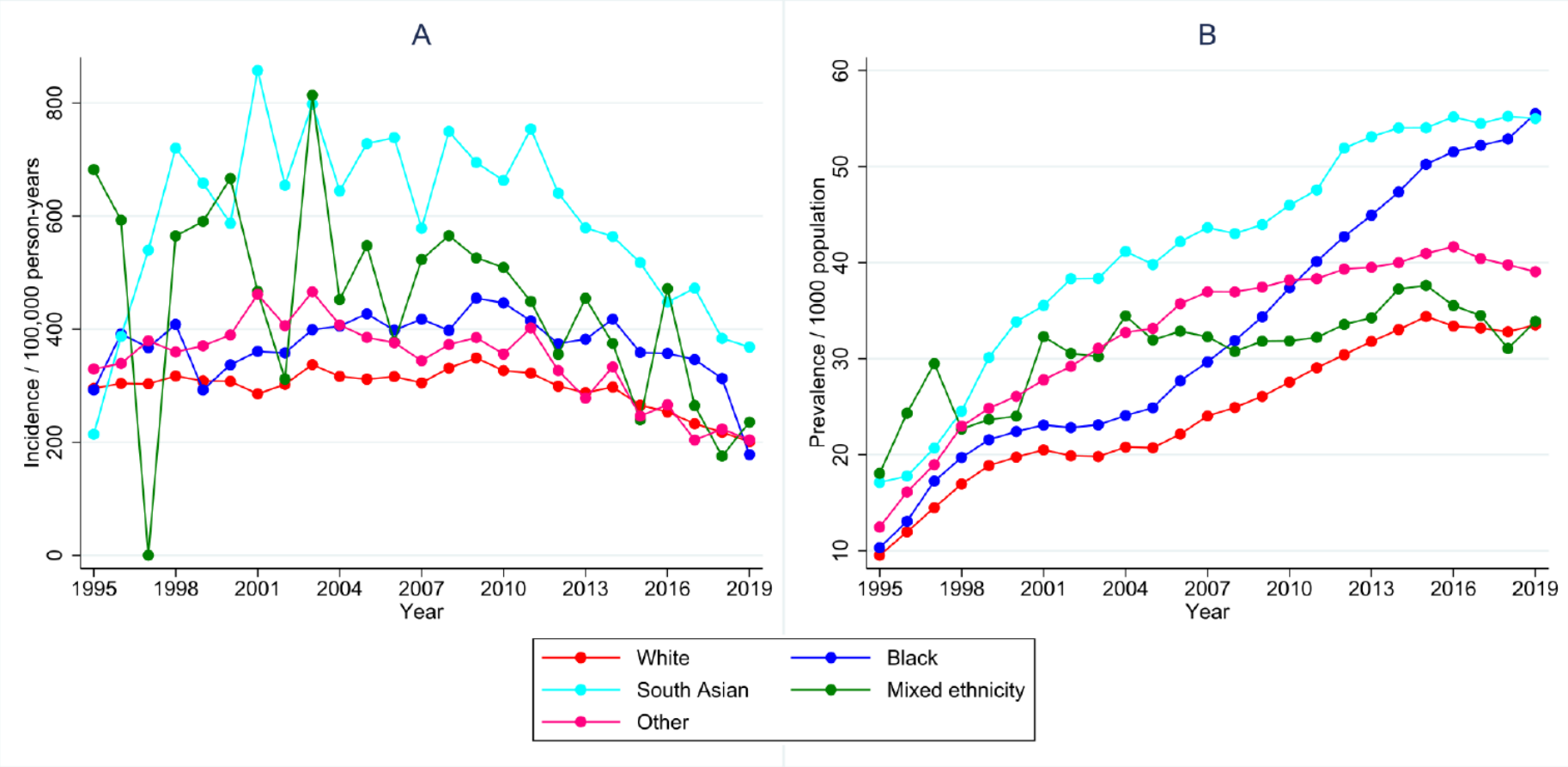
Year	Prevalence per 1,000 women based on each of the PCOS Definition											
	Diagnostic code			National Institute of Child Health and Human Development criteria			Rotterdam criteria			Androgen Excess Society criteria		
	Nr	Dr	Prevalence (95% CI)	Nr	Dr	Prevalence (95% CI)	Nr	Dr	Prevalence (95% CI)	Nr	Dr	Prevalence (95% CI)
1995	117	317088	0.37 (0.3-0.44)	2392	317088	7.54 (7.24-7.84)	2920	317088	9.21 (8.88-9.54)	2596	317088	8.19 (7.87-8.5)
1996	134	342364	0.39 (0.33-0.46)	3071	342364	8.97 (8.65-9.29)	3766	342364	11 (10.65-11.35)	3340	342364	9.76 (9.43-10.08)
1997	157	364329	0.43 (0.36-0.5)	3860	364329	10.59 (10.26-10.93)	4781	364329	13.12 (12.75-13.49)	4215	364329	11.57 (11.22-11.92)
1998	197	430425	0.46 (0.39-0.52)	5194	430425	12.07 (11.74-12.39)	6441	430425	14.96 (14.6-15.33)	5708	430425	13.26 (12.92-13.6)
1999	243	508545	0.48 (0.42-0.54)	6691	508545	13.16 (12.84-13.47)	8321	508545	16.36 (16.01-16.71)	7393	508545	14.54 (14.21-14.87)
2000	358	618005	0.58 (0.52-0.64)	8431	618005	13.64 (13.35-13.93)	10593	618005	17.14 (16.82-17.46)	9381	618005	15.18 (14.87-15.48)
2001	500	764180	0.65 (0.6-0.71)	10656	764180	13.94 (13.68-14.21)	13584	764180	17.78 (17.48-18.07)	12011	764180	15.72 (15.44-16)
2002	704	905155	0.78 (0.72-0.84)	12469	905155	13.78 (13.54-14.02)	16278	905155	17.98 (17.71-18.26)	14337	905155	15.84 (15.58-16.1)
2003	1062	1045736	1.02 (0.95-1.08)	14596	1045736	13.96 (13.73-14.18)	19465	1045736	18.61 (18.35-18.87)	17054	1045736	16.31 (16.07-16.55)
2004	1453	1177090	1.23 (1.17-1.3)	17003	1177090	14.44 (14.23-14.66)	23065	1177090	19.59 (19.34-19.85)	20180	1177090	17.14 (16.91-17.38)
2005	3009	1354587	2.22 (2.14-2.3)	19625	1354587	14.49 (14.29-14.69)	26809	1354587	19.79 (19.56-20.03)	23483	1354587	17.34 (17.12-17.56)

2006	4753	1441473	3.3 (3.2-3.39)	22416	1441473	15.55 (15.35-15.75)	30541	1441473	21.19 (20.95-21.42)	26874	1441473	18.64 (18.42-18.86)
2007	6353	1454824	4.37 (4.26-4.47)	24254	1454824	16.67 (16.46-16.88)	33031	1454824	22.7 (22.46-22.95)	29162	1454824	20.05 (19.82-20.27)
2008	7699	1463111	5.26 (5.14-5.38)	25943	1463111	17.73 (17.52-17.95)	35336	1463111	24.15 (23.9-24.4)	31277	1463111	21.38 (21.14-21.61)
2009	9058	1481499	6.11 (5.99-6.24)	28376	1481499	19.15 (18.93-19.37)	38531	1481499	26.01 (25.75-26.26)	34255	1481499	23.12 (22.88-23.36)
2010	10176	1470903	6.92 (6.78-7.05)	29944	1470903	20.36 (20.13-20.59)	40880	1470903	27.79 (27.53-28.06)	36453	1470903	24.78 (24.53-25.03)
2011	11181	1440611	7.76 (7.62-7.9)	30907	1440611	21.45 (21.22-21.69)	42393	1440611	29.43 (29.15-29.7)	37828	1440611	26.26 (26-26.52)
2012	12123	1405505	8.63 (8.47-8.78)	31551	1405505	22.45 (22.2-22.69)	43648	1405505	31.06 (30.77-31.34)	38944	1405505	27.71 (27.44-27.98)
2013	12874	1370668	9.39 (9.23-9.55)	31978	1370668	23.33 (23.08-23.58)	44570	1370668	32.52 (32.22-32.81)	39691	1370668	28.96 (28.68-29.24)
2014	13101	1275902	10.27 (10.09-10.44)	30492	1275902	23.9 (23.63-24.16)	42896	1275902	33.62 (33.31-33.93)	38218	1275902	29.95 (29.66-30.25)
2015	12500	1146863	10.9 (10.71-11.09)	28016	1146863	24.43 (24.15-24.71)	39543	1146863	34.48 (34.15-34.81)	35277	1146863	30.76 (30.44-31.08)
2016	10885	970411	11.22 (11.01-11.43)	23189	970411	23.9 (23.59-24.2)	32801	970411	33.8 (33.44-34.16)	29259	970411	30.15 (29.81-30.49)
2017	10003	864130	11.58 (11.35-11.8)	20176	864130	23.35 (23.03-23.67)	28820	864130	33.35 (32.97-33.73)	25714	864130	29.76 (29.4-30.12)
2018	9556	792143	12.06 (11.82-12.3)	18195	792143	22.97 (22.64-23.3)	26179	792143	33.05 (32.65-33.44)	23356	792143	29.48 (29.11-29.86)
2019	9414	738629	12.75 (12.49-13)	17335	738629	23.47 (23.12-23.81)	24936	738629	33.76 (33.35-34.17)	22286	738629	30.17 (29.78-30.56)

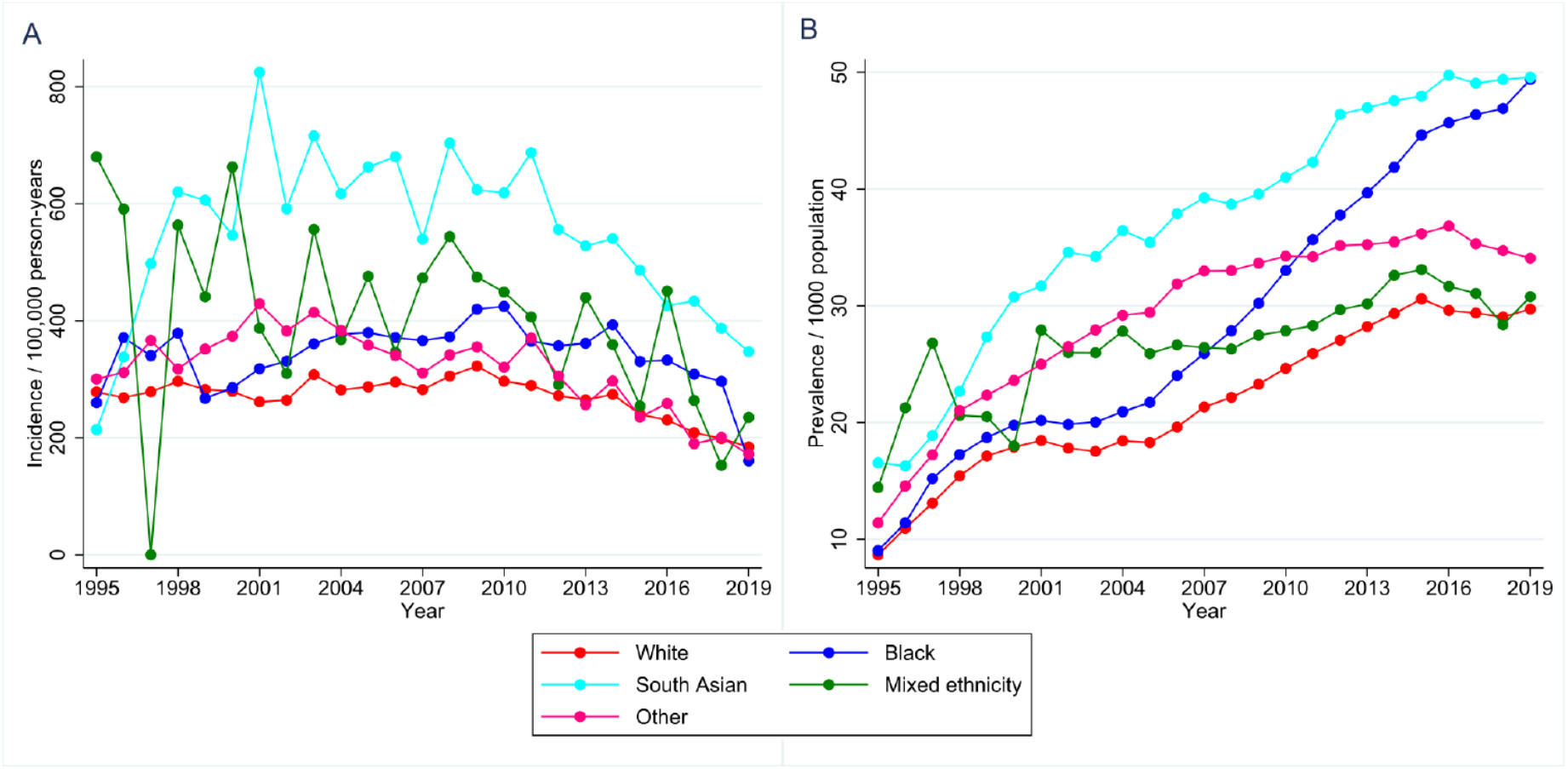
Supplementary 15: (A) Incidence and (B) prevalence trends of PCOS based on a combination of symptom codes fulfilling National Institute of Child Health and Human Development (NICHD) criteria stratified by ethnicity status



Supplementary 16: (A) Incidence and (B) prevalence trends of PCOS based on a combination of symptom codes fulfilling Rotterdam criteria stratified by ethnicity status



Supplementary 17: (A) Incidence and (B) prevalence trends of PCOS based on a combination of symptom codes fulfilling Androgen Excess Society (AES) criteria stratified by ethnicity status



Supplementary 18: Incidence trend of PCOS based on diagnostic codes stratified by ethnicity

Year	Incidence of PCOS per 100,000 person-years based on diagnostic code stratified by ethnicity														
	Caucasian			Black Afro-Caribbean			South Asian			Mixed ethnicity			Other ethnic minorities		
	Nr	Dr	Incidence	Nr	Dr	Incidence	Nr	Dr	Incidence	Nr	Dr	Incidence	Nr	Dr	Incidence
1995	5	72511	6.9	0	9320	0	0	1902	0	0	299	0	0	7075	0
1996	6	78270	7.67	1	10099	9.9	0	2106	0	0	346	0	0	7490	0
1997	6	92057	6.52	0	11639	0	1	2463	40.6	0	413	0	0	8328	0
1998	3	111335	2.69	1	14223	7.03	0	3144	0	0	545	0	0	9968	0
1999	15	133836	11.21	2	16756	11.94	1	4079	24.52	1	694	144.12	0	11346	0
2000	13	170973	7.6	3	22126	13.56	2	5288	37.82	0	925	0	0	13702	0
2001	23	224273	10.26	1	29467	3.39	1	7008	14.27	0	1325	0	2	16484	12.13
2002	38	274772	13.83	5	35419	14.12	1	8568	11.67	0	1653	0	1	18245	5.48
2003	59	337686	17.47	6	43789	13.7	3	10547	28.44	0	2030	0	1	20590	4.86
2004	320	401894	79.62	67	52454	127.73	22	12731	172.81	3	2513	119.37	9	22545	39.92
2005	380	458787	82.83	59	60111	98.15	25	14959	167.13	2	3011	66.42	15	24134	62.15
2006	420	489035	85.88	70	63527	110.19	31	16852	183.96	5	3551	140.81	18	24479	73.53
2007	413	524849	78.69	65	66029	98.44	27	19885	135.78	5	4318	115.8	14	25826	54.21
2008	460	586857	78.38	71	69623	101.98	33	24289	135.86	4	5437	73.57	22	28014	78.53
2009	468	626990	74.64	59	71190	82.88	55	27392	200.79	5	6228	80.28	15	29244	51.29
2010	521	648537	80.33	71	72447	98	59	30447	193.78	9	7032	127.99	22	30173	72.91
2011	544	650437	83.64	78	73649	105.91	78	33112	235.56	8	7520	106.38	18	30837	58.37
2012	504	653160	77.16	63	74507	84.56	57	35613	160.06	9	8061	111.65	24	31616	75.91
2013	496	629701	78.77	65	71420	91.01	78	34742	224.51	10	8118	123.18	17	30829	55.14
2014	446	575138	77.55	51	65210	78.21	51	31030	164.36	8	7389	108.26	17	26624	63.85
2015	386	507498	76.06	52	55814	93.17	49	25568	191.65	10	6416	155.86	16	21374	74.86
2016	342	431694	79.22	31	46445	66.75	20	21491	93.06	10	5420	184.5	13	17424	74.61
2017	295	390188	75.6	38	41060	92.55	24	20146	119.13	6	5000	119.99	17	15597	109
2018	240	364366	65.87	25	36193	69.07	36	18320	196.51	4	4631	86.38	6	14251	42.1
2019	251	337624	74.34	41	31447	130.38	19	16250	116.93	2	4309	46.41	9	13028	69.08

Supplementary 19: Prevalence trend of PCOS based on diagnostic codes stratified by ethnicity

Year	Prevalence of PCOS per 1,000 women based on diagnostic code stratified by ethnicity														
	Caucasian			Black Afro-Caribbean			South Asian			Mixed ethnicity			Other ethnic minorities		
	Nr	Dr	Prevalence	Nr	Dr	Prevalence	Nr	Dr	Prevalence	Nr	Dr	Prevalence	Nr	Dr	Prevalence
1995	39	67179	0.58	9	8633	1.04	1	1812	0.55	0	277	0	4	6662	0.6
1996	45	74993	0.6	9	9642	0.93	1	2026	0.49	0	329	0	4	7137	0.56
1997	56	80210	0.7	10	10317	0.97	1	2223	0.45	0	373	0	5	7536	0.66
1998	72	100925	0.71	12	12681	0.95	2	2689	0.74	0	485	0	6	8796	0.68
1999	87	120792	0.72	17	15117	1.12	3	3585	0.84	0	634	0	6	10473	0.57
2000	139	155428	0.89	25	19810	1.26	4	4876	0.82	1	833	1.2	8	12536	0.64
2001	185	195680	0.95	29	25123	1.15	9	5993	1.5	2	1146	1.75	12	15073	0.8
2002	265	248791	1.07	38	32256	1.18	13	7776	1.67	2	1539	1.3	15	17093	0.88
2003	430	309941	1.39	63	39897	1.58	18	9646	1.87	3	1886	1.59	17	19220	0.88
2004	587	360884	1.63	79	46481	1.7	29	11441	2.53	3	2264	1.33	22	20923	1.05
2005	1138	434979	2.62	177	56332	3.14	60	14197	4.23	8	2817	2.84	37	23201	1.59
2006	1706	478214	3.57	278	62357	4.46	99	16124	6.14	11	3378	3.26	68	24244	2.8
2007	2321	504245	4.6	376	64319	5.85	155	18493	8.38	19	3935	4.83	95	24681	3.85
2008	3153	557729	5.65	492	67542	7.28	226	22196	10.18	35	4944	7.08	125	26715	4.68
2009	4098	615248	6.66	601	70346	8.54	307	26479	11.59	53	6004	8.83	155	28414	5.46
2010	4845	642691	7.54	683	71674	9.53	397	29234	13.58	62	6784	9.14	184	29464	6.24
2011	5506	656029	8.39	790	72417	10.91	479	32645	14.67	80	7452	10.74	218	30296	7.2
2012	6100	652704	9.35	891	74387	11.98	584	35176	16.6	93	7956	11.69	237	31203	7.6
2013	6636	651013	10.19	955	73507	12.99	657	36184	18.16	100	8320	12.02	264	31488	8.38
2014	6756	609738	11.08	982	68368	14.36	694	33562	20.68	102	7972	12.79	292	29239	9.99
2015	6485	546165	11.87	939	61439	15.28	637	29168	21.84	95	7097	13.39	265	23904	11.09
2016	5572	456654	12.2	832	49275	16.88	492	22555	21.81	78	5715	13.65	221	18295	12.08
2017	5218	411296	12.69	735	43070	17.07	479	21262	22.53	75	5247	14.29	214	16251	13.17
2018	4945	376085	13.15	678	38641	17.55	451	19905	22.66	78	4863	16.04	219	14712	14.89
2019	4860	352852	13.77	618	33312	18.55	427	17469	24.44	85	4516	18.82	225	13673	16.46

Supplementary 20: Interaction between ethnicity and BMI on the risk of PCOS diagnosis

Variables	Adjusted Hazard ratio (95% CI)
BMI # Ethnicity	
Underweight # Caucasian	0.74 (0.62-0.88)
Underweight # Black Afro-Caribbean	1.12 (0.74-1.68)
Underweight # South Asian	1.64 (1.12-2.40)
Underweight # Mixed ethnicity	0.66 (0.16-2.64)
Underweight # Other ethnic minorities	0.37 (0.14-1.00)
Underweight # missing ethnicity	0.63 (0.52-0.75)
Normal weight # Caucasian	Ref
Normal weight # Black Afro-Caribbean	1.54 (1.34-1.76)
Normal weight # South Asian	2.41 (2.08-2.81)
Normal weight # Mixed ethnicity	1.37 (0.95-1.98)
Normal weight # Other ethnic minorities	1.27 (1.01-1.59)
Normal weight # missing ethnicity	0.84 (0.78-0.90)
Overweight # Caucasian	1.83 (1.70-1.97)
Overweight # Black Afro-Caribbean	2.62 (2.25-3.05)
Overweight # South Asian	5.04 (4.28-5.92)
Overweight # Mixed ethnicity	1.95 (1.18-3.25)
Overweight # Other ethnic minorities	3.27 (2.55-4.19)
Overweight # missing ethnicity	1.79 (1.66-1.93)
Obese # Caucasian	5.14 (4.84-5.47)
Obese # Black Afro-Caribbean	6.03 (5.40-6.75)
Obese # South Asian	11.37 (9.82-13.17)
Obese # Mixed ethnicity	7.68 (5.63-10.47)
Obese # Other ethnic minorities	6.64 (5.31-8.30)
Obese # missing ethnicity	5.00 (4.70-5.31)
Missing BMI # Caucasian	0.75 (0.69-0.81)
Missing BMI # Black Afro-Caribbean	1.07 (0.88-1.30)
Missing BMI # South Asian	1.66 (1.37-2.01)
Missing BMI # Mixed ethnicity	0.78 (0.47-1.30)
Missing BMI # Other ethnic minorities	0.95 (0.66-1.36)
Missing BMI # missing ethnicity	0.58 (0.54-0.63)
Age category	
20-25	1.63 (0.58-4.59)
25-30	1.01 (0.16-6.35)
30-35	9.09 (0.89-92.81)
35-40	2.77 (0.13-57.16)
40-45	5.01 (0.11-231.28)
45-50	NE
Year	
1996	1.69 (0.63-4.58)
1997	1.49 (0.55-4.02)
1998	1.57 (0.60-4.07)
1999	2.72 (1.13-6.58)
2000	2.75 (1.16-6.54)

2001	4.59 (2.00-10.55)
2002	5.9 (2.59-13.40)
2003	7.33 (3.24-16.55)
2004	35.23 (15.79-78.63)
2005	38.81 (17.40-86.56)
2006	37.98 (17.03-84.71)
2007	31.47 (14.10-70.21)
2008	30.76 (13.79-68.63)
2009	30.21 (13.54-67.4)
2010	31.55 (14.14-70.38)
2011	32.77 (14.69-73.10)
2012	29.67 (13.29-66.19)
2013	30.97 (13.88-69.11)
2014	29.27 (13.11-65.34)
2015	30.04 (13.45-67.07)
2016	28.67 (12.83-64.06)
2017	29.18 (13.06-65.22)
2018	25.9 (11.58-57.92)
2019	29.77 (13.31-66.56)
2020	20.38 (9.08-45.72)

NE – Not estimated due to small sample size within this patient category

Supplementary 21: Impaired Glucose Regulation definitions based on NICE guidelines

Criteria	Definition
Impaired Glucose Regulation / Type 2 diabetes	Read code diagnosis OR HbA1c ≥ 6.0 % OR Fasting blood glucose > 6.0 0mmol/L
Type 2 diabetes	Read code diagnosis OR HbA1c ≥ 6.5 % OR Fasting blood glucose ≥ 7.0 0mmol/L
Gestational Diabetes	Read code diagnosis

Note: Due to under-recording of Oral Glucose Tolerance Test records, these were not included in this study to define impaired glucose regulation

Supplementary 22: Read code lists for diagnosis of impaired glucose regulation

READ_CODE	DESCRIPTION
R102.12	[D]Impaired glucose tolerance test
C313500	Glucose intolerance
R10E.00	[D]Impaired glucose tolerance
R10D000	[D]Impaired fasting glycaemia
C11y200	Impaired glucose tolerance
C11y300	Impaired fasting glycaemia
R102.11	[D]Prediabetes
44V2.00	Glucose tol. test impaired
R10D011	[D]Impaired fasting glucose
8HlS.00	Referral for management of impaired glucose tolerance
9NS0400	Referral for impaired glucose tolerance management offered
C11y400	Impaired glucose regulation
C11y500	Pre-diabetes
9mX..00	Impaired glucose regulation monitoring invitation
9mX0.00	Impaired glucose regulation monitoring invitation 1st letter
9mX4.00	Impaired glucose regulation monitoring verbal invitation
9mX1.00	Impaired glucose regulation monitoring invitation 2nd letter
9mX2.00	Impaired glucose regulation monitoring invitation 3rd letter
9mX3.00	Impaired glucose regulation monitoring telephone invitation
6AC..00	Review of impaired glucose tolerance
2126900	Impaired glucose tolerance resolved
2126B00	Impaired fasting glycaemia resolved
9m9..00	Impaired glucose tolerance monitoring administration
9m90000	Impaired glucose tolerance monitoring invitation 1st letter
9m90100	Impaired glucose tolerance monitoring invitation 2nd letter
9m90.00	Impaired glucose tolerance monitoring invitation
9m90200	Impaired glucose tolerance monitoring invitation 3rd letter
C317.00	Non-diabetic hyperglycaemia

Supplementary 23: Read code lists for type 2 diabetes

READ_CODE	DESCRIPTION
C100112	Non-insulin dependent diabetes mellitus
C10F.00	Type 2 diabetes mellitus
C10FJ00	Insulin treated Type 2 diabetes mellitus
66A4.00	Diabetic on oral treatment
C109.00	Non-insulin dependent diabetes mellitus
C109.11	NIDDM - Non-insulin dependent diabetes mellitus
66A3.00	Diabetic on diet only
C109700	Non-insulin dependent diabetes mellitus - poor control
C10FC00	Type 2 diabetes mellitus with nephropathy
C10F500	Type 2 diabetes mellitus with gangrene
C109600	Non-insulin-dependent diabetes mellitus with retinopathy

C109.12	Type 2 diabetes mellitus
C109G11	Type II diabetes mellitus with arthropathy
C109012	Type 2 diabetes mellitus with renal complications
C109.13	Type II diabetes mellitus
C109J12	Insulin treated Type II diabetes mellitus
C109J00	Insulin treated Type 2 diabetes mellitus
C10FM00	Type 2 diabetes mellitus with persistent microalbuminuria
C10FB00	Type 2 diabetes mellitus with polyneuropathy
C10F600	Type 2 diabetes mellitus with retinopathy
C10F000	Type 2 diabetes mellitus with renal complications
C10F.11	Type II diabetes mellitus
C109711	Type II diabetes mellitus - poor control
C109G00	Non-insulin dependent diabetes mellitus with arthropathy
C109C12	Type 2 diabetes mellitus with nephropathy
ZC2CA00	Dietary advice for type II diabetes
C10FQ00	Type 2 diabetes mellitus with exudative maculopathy
C10F700	Type 2 diabetes mellitus - poor control
C10FL00	Type 2 diabetes mellitus with persistent proteinuria
C109900	Non-insulin-dependent diabetes mellitus without complication
C10FN00	Type 2 diabetes mellitus with ketoacidosis
C10F200	Type 2 diabetes mellitus with neurological complications
C10FK00	Hyperosmolar non-ketotic state in type 2 diabetes mellitus
C109400	Non-insulin dependent diabetes mellitus with ulcer
C10FH00	Type 2 diabetes mellitus with neuropathic arthropathy
C109K00	Hyperosmolar non-ketotic state in type 2 diabetes mellitus
C109J11	Insulin treated non-insulin dependent diabetes mellitus
C10FF00	Type 2 diabetes mellitus with peripheral angiopathy
C109500	Non-insulin dependent diabetes mellitus with gangrene
C109H00	Non-insulin dependent d m with neuropathic arthropathy
C109612	Type 2 diabetes mellitus with retinopathy
C10F311	Type II diabetes mellitus with multiple complications
C109D00	Non-insulin dependent diabetes mellitus with hypoglyca coma
C109E12	Type 2 diabetes mellitus with diabetic cataract
C10FE00	Type 2 diabetes mellitus with diabetic cataract
C109B00	Non-insulin dependent diabetes mellitus with polyneuropathy
C109712	Type 2 diabetes mellitus - poor control
C109212	Type 2 diabetes mellitus with neurological complications
C109512	Type 2 diabetes mellitus with gangrene
C10FD00	Type 2 diabetes mellitus with hypoglycaemic coma
C10F711	Type II diabetes mellitus - poor control
C10F100	Type 2 diabetes mellitus with ophthalmic complications
C109B11	Type II diabetes mellitus with polyneuropathy
C109H11	Type II diabetes mellitus with neuropathic arthropathy
C10F900	Type 2 diabetes mellitus without complication
C109E11	Type II diabetes mellitus with diabetic cataract
C10F400	Type 2 diabetes mellitus with ulcer

C10F611	Type II diabetes mellitus with retinopathy
C109G12	Type 2 diabetes mellitus with arthropathy
C109011	Type II diabetes mellitus with renal complications
C109100	Non-insulin-dependent diabetes mellitus with ophthalm comps
C10FB11	Type II diabetes mellitus with polyneuropathy
L180600	Pre-existing diabetes mellitus, non-insulin-dependent
C109A11	Type II diabetes mellitus with mononeuropathy
C10FP00	Type 2 diabetes mellitus with ketoacidotic coma
C109000	Non-insulin-dependent diabetes mellitus with renal comps
C10F911	Type II diabetes mellitus without complication
C109F00	Non-insulin-dependent d m with peripheral angiopath
C109F11	Type II diabetes mellitus with peripheral angiopathy
C109411	Type II diabetes mellitus with ulcer
C109200	Non-insulin-dependent diabetes mellitus with neuro comps
C109D11	Type II diabetes mellitus with hypoglycaemic coma
C107400	NIDDM with peripheral circulatory disorder
C10F011	Type II diabetes mellitus with renal complications
C109611	Type II diabetes mellitus with retinopathy
C10FG00	Type 2 diabetes mellitus with arthropathy
C109C00	Non-insulin dependent diabetes mellitus with nephropathy
C109111	Type II diabetes mellitus with ophthalmic complications
C109F12	Type 2 diabetes mellitus with peripheral angiopathy
C10FL11	Type II diabetes mellitus with persistent proteinuria
C109D12	Type 2 diabetes mellitus with hypoglycaemic coma
C109511	Type II diabetes mellitus with gangrene
C109300	Non-insulin-dependent diabetes mellitus with multiple comps
C10FA00	Type 2 diabetes mellitus with mononeuropathy
C10FR00	Type 2 diabetes mellitus with gastroparesis
C109C11	Type II diabetes mellitus with nephropathy
C10FJ11	Insulin treated Type II diabetes mellitus
C10F300	Type 2 diabetes mellitus with multiple complications
C109412	Type 2 diabetes mellitus with ulcer
C109H12	Type 2 diabetes mellitus with neuropathic arthropathy
C109211	Type II diabetes mellitus with neurological complications
C109E00	Non-insulin depend diabetes mellitus with diabetic cataract
C109112	Type 2 diabetes mellitus with ophthalmic complications
C109A00	Non-insulin dependent diabetes mellitus with mononeuropathy
C10FM11	Type II diabetes mellitus with persistent microalbuminuria
C10F411	Type II diabetes mellitus with ulcer
C10FE11	Type II diabetes mellitus with diabetic cataract
C10FA11	Type II diabetes mellitus with mononeuropathy
C10F211	Type II diabetes mellitus with neurological complications
C10FD11	Type II diabetes mellitus with hypoglycaemic coma
C10F111	Type II diabetes mellitus with ophthalmic complications
C10FC11	Type II diabetes mellitus with nephropathy
C10FG11	Type II diabetes mellitus with arthropathy

C10F511	Type II diabetes mellitus with gangrene
C10FF11	Type II diabetes mellitus with peripheral angiopathy
C109912	Type 2 diabetes mellitus without complication
C10FP11	Type II diabetes mellitus with ketoacidotic coma
C10FN11	Type II diabetes mellitus with ketoacidosis
C10FK11	Hyperosmolar non-ketotic state in type II diabetes mellitus
C10P100	Type II diabetes mellitus in remission
C109312	Type 2 diabetes mellitus with multiple complications
C109911	Type II diabetes mellitus without complication
C10FH11	Type II diabetes mellitus with neuropathic arthropathy
66AH300	Conversion to non-insulin injectable medication
C109B12	Type 2 diabetes mellitus with polyneuropathy
66o2.00	Diabetic on non-insulin injectable medication
66o5.00	Diabetic on oral treatment and glucagon-like peptide 1
C10P111	Type 2 diabetes mellitus in remission
C10FQ11	Type II diabetes mellitus with exudative maculopathy
C109311	Type II diabetes mellitus with multiple complications
C10FR11	Type II diabetes mellitus with gastroparesis

Supplementary 24: Read code lists for gestational diabetes mellitus (GDM)

READ_CODE	DESCRIPTION
L180900	Gestational diabetes mellitus
L180811	Gestational diabetes mellitus
L180800	Diabetes mellitus arising in pregnancy
L180.00	Diabetes mellitus during pregnancy/childbirth/puerperium
L180100	Diabetes mellitus during pregnancy - baby delivered
L180300	Diabetes mellitus during pregnancy - baby not yet delivered
L180z00	Diabetes mellitus in pregnancy/childbirth/puerperium NOS
L180000	Diabetes mellitus - unspec whether in pregnancy/puerperium
ZV13F00	[V]Personal history of gestational diabetes mellitus
ZC2CB00	Dietary advice for gestational diabetes

Supplementary 25: Read code lists for type 1 diabetes

READ_CODE	DESCRIPTION
C100011	Insulin dependent diabetes mellitus
C10E.00	Type 1 diabetes mellitus
C108.00	Insulin dependent diabetes mellitus
C108700	Insulin dependent diabetes mellitus with retinopathy
C108800	Insulin dependent diabetes mellitus - poor control
C10ED00	Type 1 diabetes mellitus with nephropathy
C10EM00	Type 1 diabetes mellitus with ketoacidosis
C10E.11	Type I diabetes mellitus
C108F11	Type I diabetes mellitus with diabetic cataract

C108.12	Type 1 diabetes mellitus
C108J12	Type 1 diabetes mellitus with neuropathic arthropathy
C10E700	Type 1 diabetes mellitus with retinopathy
C108.11	IDDM-Insulin dependent diabetes mellitus
C10EH00	Type 1 diabetes mellitus with arthropathy
C10E500	Type 1 diabetes mellitus with ulcer
C108012	Type 1 diabetes mellitus with renal complications
C10EP00	Type 1 diabetes mellitus with exudative maculopathy
C108.13	Type I diabetes mellitus
C108B00	Insulin dependent diabetes mellitus with mononeuropathy
C108400	Unstable insulin dependent diabetes mellitus
C10EL00	Type 1 diabetes mellitus with persistent microalbuminuria
C10EK00	Type 1 diabetes mellitus with persistent proteinuria
C108900	Insulin dependent diabetes maturity onset
C10E800	Type 1 diabetes mellitus - poor control
C108711	Type I diabetes mellitus with retinopathy
C10EE00	Type 1 diabetes mellitus with hypoglycaemic coma
C108J00	Insulin dependent diab mell with neuropathic arthropathy
C10E900	Type 1 diabetes mellitus maturity onset
C10EN00	Type 1 diabetes mellitus with ketoacidotic coma
C108712	Type 1 diabetes mellitus with retinopathy
C108C00	Insulin dependent diabetes mellitus with polyneuropathy
C108E11	Type I diabetes mellitus with hypoglycaemic coma
C10E200	Type 1 diabetes mellitus with neurological complications
C10E400	Unstable type 1 diabetes mellitus
C108F00	Insulin dependent diabetes mellitus with diabetic cataract
C108E00	Insulin dependent diabetes mellitus with hypoglycaemic coma
C108500	Insulin dependent diabetes mellitus with ulcer
C10E312	Insulin dependent diabetes mellitus with multiple complicat
C108812	Type 1 diabetes mellitus - poor control
C10EC00	Type 1 diabetes mellitus with polyneuropathy
C108811	Type I diabetes mellitus - poor control
C10E000	Type 1 diabetes mellitus with renal complications
C10E100	Type 1 diabetes mellitus with ophthalmic complications
C10E300	Type 1 diabetes mellitus with multiple complications
C108211	Type I diabetes mellitus with neurological complications
C10EF00	Type 1 diabetes mellitus with diabetic cataract
C10E411	Unstable type I diabetes mellitus
C10E.12	Insulin dependent diabetes mellitus
C108511	Type I diabetes mellitus with ulcer
C108300	Insulin dependent diabetes mellitus with multiple complicatn
C10EJ00	Type 1 diabetes mellitus with neuropathic arthropathy
C10E412	Unstable insulin dependent diabetes mellitus
C10EQ00	Type 1 diabetes mellitus with gastroparesis
C108D00	Insulin dependent diabetes mellitus with nephropathy
C108411	Unstable type I diabetes mellitus

C108J11	Type I diabetes mellitus with neuropathic arthropathy
C108600	Insulin dependent diabetes mellitus with gangrene
C108011	Type I diabetes mellitus with renal complications
C108212	Type 1 diabetes mellitus with neurological complications
C10EM11	Type I diabetes mellitus with ketoacidosis
C108H11	Type I diabetes mellitus with arthropathy
C10EA11	Type I diabetes mellitus without complication
C108911	Type I diabetes mellitus maturity onset
C108G00	Insulin dependent diab mell with peripheral angiopathy
C108H00	Insulin dependent diabetes mellitus with arthropathy
C10EN11	Type I diabetes mellitus with ketoacidotic coma
C108D11	Type I diabetes mellitus with nephropathy
C10EB00	Type 1 diabetes mellitus with mononeuropathy
C108512	Type 1 diabetes mellitus with ulcer
ZC2C900	Dietary advice for type I diabetes
C10EA00	Type 1 diabetes mellitus without complication
C10E600	Type 1 diabetes mellitus with gangrene
C108E12	Type 1 diabetes mellitus with hypoglycaemic coma
C10E812	Insulin dependent diabetes mellitus - poor control
C10E311	Type I diabetes mellitus with multiple complications
C10EC11	Type I diabetes mellitus with polyneuropathy
C10EG00	Type 1 diabetes mellitus with peripheral angiopathy
C10E712	Insulin dependent diabetes mellitus with retinopathy
C10E511	Type I diabetes mellitus with ulcer
C10E711	Type I diabetes mellitus with retinopathy
C108A11	Type I diabetes mellitus without complication
C10E911	Type I diabetes mellitus maturity onset
C108912	Type 1 diabetes mellitus maturity onset
C108412	Unstable type 1 diabetes mellitus
C10E912	Insulin dependent diabetes maturity onset
C10EP11	Type I diabetes mellitus with exudative maculopathy
C10E512	Insulin dependent diabetes mellitus with ulcer
C108B11	Type I diabetes mellitus with mononeuropathy
C10E111	Type I diabetes mellitus with ophthalmic complications
C10EE12	Insulin dependent diabetes mellitus with hypoglycaemic coma
C10EF12	Insulin dependent diabetes mellitus with diabetic cataract
C10EC12	Insulin dependent diabetes mellitus with polyneuropathy
C10E611	Type I diabetes mellitus with gangrene
C10ED12	Insulin dependent diabetes mellitus with nephropathy
C10EL11	Type I diabetes mellitus with persistent microalbuminuria
C108112	Type 1 diabetes mellitus with ophthalmic complications
66At011	Type 1 diabetic dietary review
C10E811	Type I diabetes mellitus - poor control
C108311	Type I diabetes mellitus with multiple complications
C10P000	Type I diabetes mellitus in remission
C10EQ11	Type I diabetes mellitus with gastroparesis

C10E612	Insulin dependent diabetes mellitus with gangrene
C10P011	Type 1 diabetes mellitus in remission
C10E011	Type I diabetes mellitus with renal complications
C108F12	Type 1 diabetes mellitus with diabetic cataract
C108A12	Type 1 diabetes mellitus without complication
C108D12	Type 1 diabetes mellitus with nephropathy
C10ED11	Type I diabetes mellitus with nephropathy
C108C11	Type I diabetes mellitus with polyneuropathy
C108612	Type 1 diabetes mellitus with gangrene

Supplementary 26: Incidence rate trend of impaired glucose regulation (prediabetes/type 2 diabetes) among women with PCOS defined using diagnostic codes and a combination of symptom codes fulfilling each of the diagnostic criteria

Year	Incidence per 100,000 person-years											
	PCOS definition - Diagnostic code			PCOS definition – combination of symptom codes fulfilling National Institute of Child Health and Human Development criteria for PCOS diagnosis			PCOS definition – combination of symptom codes fulfilling Rotterdam criteria for PCOS diagnosis			PCOS definition – combination of symptom codes fulfilling Androgen Excess Society criteria for PCOS diagnosis		
	Nr	Dr	Incidence rate	Nr	Dr	Incidence rate	Nr	Dr	Incidence rate	Nr	Dr	Incidence rate
1995	0	121.1225	0	2	2706.063	73.91	3	3299.498	90.92	3	2938.608	102.09
1996	1	139.7591	715.52	4	3436.99	116.38	4	4225.217	94.67	4	3740.398	106.94
1997	0	172.5448	0	5	4462.845	112.04	6	5522.038	108.66	5	4876.828	102.53
1998	2	206.4285	968.86	6	5866.124	102.28	9	7268.731	123.82	7	6462.229	108.32
1999	3	269.4319	1113.45	14	7345.592	190.59	18	9161.297	196.48	15	8132.524	184.44
2000	2	388.7996	514.4	29	9252.068	313.44	40	11643.61	343.54	36	10332.68	348.41
2001	5	568.8187	879.01	32	11409.69	280.46	46	14700.6	312.91	37	12994.2	284.74
2002	9	814.1396	1105.46	44	13165.62	334.2	71	17341.34	409.43	58	15254.59	380.21
2003	15	1192.201	1258.18	60	15474.08	387.75	96	20759.88	462.43	80	18205.34	439.43
2004	32	2057.238	1555.48	83	17911.97	463.38	121	24299.61	497.95	102	21321.13	478.4
2005	34	3716.775	914.77	94	20465.19	459.32	158	27825.49	567.82	129	24473.93	527.09
2006	59	5235.175	1126.99	90	22584.09	398.51	146	30582.92	477.39	119	27026.47	440.31
2007	74	6617.246	1118.29	112	24390.84	459.19	185	32974.63	561.04	152	29227.9	520.05
2008	60	7905.297	758.98	109	26458.72	411.96	174	35742.38	486.82	147	31779.34	462.56
2009	100	9015.503	1109.2	155	28247.32	548.72	233	38171.93	610.4	198	34097.89	580.68
2010	115	9980.923	1152.2	183	29463.79	621.1	305	39973.79	763	250	35775.61	698.8
2011	134	10720.81	1249.91	229	30008.74	763.11	367	41028.2	894.51	303	36700.76	825.6
2012	147	11597.96	1267.46	257	30674.87	837.82	392	42209.13	928.71	332	37756.47	879.32
2013	240	11812.71	2031.71	441	29802.26	1479.75	653	41397.58	1577.39	575	37001.37	1554
2014	200	11585.78	1726.25	424	27559.45	1538.49	619	38516.72	1607.09	541	34479.38	1569.05

2015	254	10662.46	2382.19	397	24371.46	1628.95	585	34113.68	1714.85	508	30569.13	1661.81
2016	233	9162.938	2542.85	363	19780.38	1835.15	543	27924.44	1944.53	488	25008.4	1951.34
2017	169	8402.679	2011.26	306	17273.39	1771.51	447	24548.11	1820.91	403	21963.67	1834.85
2018	185	8075.505	2290.88	381	15830.07	2406.81	530	22565.02	2348.77	479	20214.78	2369.55
2019	228	7707.078	2958.32	383	14508.33	2639.86	520	20695.87	2512.58	472	18569.22	2541.84

Supplementary 27: Incidence rate trend of type 2 diabetes among women with PCOS defined using diagnostic codes and a combination of symptom codes fulfilling each of the diagnostic criteria

Year	Incidence per 100,000 person-years											
	PCOS definition - Diagnostic code			PCOS definition – combination of symptom codes fulfilling National Institute of Child Health and Human Development criteria for PCOS diagnosis			PCOS definition – combination of symptom codes fulfilling Rotterdam criteria for PCOS diagnosis			PCOS definition – combination of symptom codes fulfilling Androgen Excess Society criteria for PCOS diagnosis		
	Nr	Dr	Incidence rate	Nr	Dr	Incidence rate	Nr	Dr	Incidence rate	Nr	Dr	Incidence rate
1995	0	121	0	2	2706	73.91	3	3299	90.92	3	2939	102.09
1996	1	140	715.52	3	3437	87.28	3	4225	71	3	3741	80.2
1997	0	173	0	4	4465	89.59	5	5524	90.52	4	4879	81.99
1998	2	206	968.86	3	5872	51.09	6	7275	82.47	4	6468	61.84
1999	3	269	1113.45	10	7356	135.94	14	9174	152.61	11	8144	135.07
2000	1	389	256.87	22	9273	237.25	31	11667	265.71	27	10355	260.75
2001	5	571	875.27	26	11440	227.27	37	14737	251.07	30	13028	230.27
2002	7	819	854.87	36	13202	272.69	60	17390	345.02	48	15297	313.79
2003	11	1199	917.79	42	15526	270.51	70	20835	335.97	59	18269	322.95
2004	26	2073	1254.16	59	17990	327.96	95	24411	389.16	79	21418	368.85
2005	32	3751	853.03	63	20572	306.23	107	27984	382.36	85	24607	345.43
2006	39	5293	736.78	65	22716	286.14	108	30793	350.73	87	27205	319.8
2007	47	6706	700.89	81	24554	329.88	127	33236	382.11	109	29445	370.19
2008	52	8021	648.33	82	26649	307.7	127	36059	352.2	106	32039	330.85
2009	71	9150	775.95	105	28468	368.84	169	38534	438.57	140	34399	406.99
2010	98	10146	965.94	134	29736	450.64	240	40397	594.1	193	36137	534.08
2011	107	10908	980.92	178	30332	586.85	284	41519	684.03	235	37121	633.07
2012	101	11825	854.16	198	31033	638.04	308	42756	720.36	264	38218	690.77
2013	194	12087	1604.99	355	30190	1175.89	521	42003	1240.4	461	37510	1229
2014	174	11877	1464.98	354	27960	1266.09	532	39143	1359.12	453	35016	1293.7

2015	213	10952	1944.91	316	24766	1275.94	467	34724	1344.89	408	31095	1312.1
2016	190	9411	2018.86	292	20134	1450.29	425	28475	1492.55	382	25484	1498.97
2017	149	8638	1724.98	260	17622	1475.43	388	25082	1546.95	350	22430	1560.38
2018	144	8314	1731.93	315	16171	1947.88	431	23091	1866.51	397	20674	1920.29
2019	186	7966	2334.95	326	14870	2192.4	446	21232	2100.6	408	19039	2143.02

Supplementary 28: Prevalence trend of impaired glucose regulation (prediabetes / type 2 diabetes) among women with PCOS defined using diagnostic codes and a combination of symptom codes fulfilling each of the diagnostic criteria

Year	Prevalence per 1,000 women											
	PCOS definition - Diagnostic code			PCOS definition – combination of symptom codes fulfilling National Institute of Child Health and Human Development criteria for PCOS diagnosis			PCOS definition – combination of symptom codes fulfilling Rotterdam criteria for PCOS diagnosis			PCOS definition – combination of symptom codes fulfilling Androgen Excess Society criteria for PCOS diagnosis		
	Nr	Dr	Prevalence	Nr	Dr	Prevalence	Nr	Dr	Prevalence	Nr	Dr	Prevalence
1995	0	117	0	3	2392	1.25	5	2920	1.71	5	2596	1.93
1996	0	134	0	4	3071	1.3	8	3766	2.12	6	3340	1.8
1997	2	157	12.74	11	3860	2.85	17	4781	3.56	14	4215	3.32
1998	4	197	20.3	29	5194	5.58	35	6441	5.43	33	5708	5.78
1999	7	243	28.81	44	6691	6.58	58	8321	6.97	51	7393	6.9
2000	11	358	30.73	76	8431	9.01	104	10593	9.82	90	9381	9.59
2001	15	500	30	110	10656	10.32	160	13584	11.78	135	12011	11.24
2002	23	704	32.67	154	12469	12.35	229	16278	14.07	190	14337	13.25
2003	37	1062	34.84	223	14596	15.28	346	19465	17.78	281	17054	16.48
2004	60	1453	41.29	305	17003	17.94	481	23065	20.85	396	20180	19.62
2005	125	3009	41.54	402	19625	20.48	632	26809	23.57	518	23483	22.06
2006	203	4753	42.71	515	22416	22.97	818	30541	26.78	673	26874	25.04
2007	300	6353	47.22	624	24254	25.73	994	33031	30.09	818	29162	28.05
2008	401	7699	52.08	721	25943	27.79	1162	35336	32.88	962	31277	30.76
2009	483	9058	53.32	827	28376	29.14	1343	38531	34.86	1114	34255	32.52
2010	597	10176	58.67	955	29944	31.89	1537	40880	37.6	1287	36453	35.31
2011	738	11181	66	1136	30907	36.76	1825	42393	43.05	1539	37828	40.68
2012	859	12123	70.86	1313	31551	41.62	2110	43648	48.34	1768	38944	45.4
2013	990	12874	76.9	1493	31978	46.69	2390	44570	53.62	2012	39691	50.69
2014	1219	13101	93.05	1742	30492	57.13	2780	42896	64.81	2354	38218	61.59
2015	1256	12500	100.48	1884	28016	67.25	2933	39543	74.17	2503	35277	70.95

2016	1234	10885	113.37	1838	23189	79.26	2798	32801	85.3	2394	29259	81.82
2017	1293	10003	129.26	1886	20176	93.48	2881	28820	99.97	2496	25714	97.07
2018	1334	9556	139.6	1919	18195	105.47	2923	26179	111.65	2557	23356	109.48
2019	1455	9414	154.56	2169	17335	125.12	3265	24936	130.94	2870	22286	128.78

Supplementary 29: Prevalence trend of type 2 diabetes among women with PCOS defined using diagnostic codes and a combination of symptom codes fulfilling each of the diagnostic criteria

Year	Prevalence per 1,000 women											
	PCOS definition - Diagnostic code			PCOS definition – combination of symptom codes fulfilling National Institute of Child Health and Human Development criteria for PCOS diagnosis			PCOS definition – combination of symptom codes fulfilling Rotterdam criteria for PCOS diagnosis			PCOS definition – combination of symptom codes fulfilling Androgen Excess Society criteria for PCOS diagnosis		
	Nr	Dr	Prevalence	Nr	Dr	Prevalence	Nr	Dr	Prevalence	Nr	Dr	Prevalence
1995	0	117	0	3	2392	1.25	5	2920	1.71	5	2596	1.93
1996	0	134	0	4	3071	1.3	8	3766	2.12	6	3340	1.8
1997	2	157	12.74	10	3860	2.59	16	4781	3.35	13	4215	3.08
1998	4	197	20.3	25	5194	4.81	31	6441	4.81	29	5708	5.08
1999	7	243	28.81	36	6691	5.38	49	8321	5.89	43	7393	5.82
2000	11	358	30.73	60	8431	7.12	85	10593	8.02	73	9381	7.78
2001	13	500	26	86	10656	8.07	131	13584	9.64	108	12011	8.99
2002	19	704	26.99	120	12469	9.62	187	16278	11.49	152	14337	10.6
2003	33	1062	31.07	182	14596	12.47	290	19465	14.9	233	17054	13.66
2004	51	1453	35.1	241	17003	14.17	389	23065	16.87	316	20180	15.66
2005	102	3009	33.9	312	19625	15.9	503	26809	18.76	408	23483	17.37
2006	164	4753	34.5	397	22416	17.71	635	30541	20.79	517	26874	19.24
2007	225	6353	35.42	478	24254	19.71	762	33031	23.07	624	29162	21.4
2008	295	7699	38.32	548	25943	21.12	877	35336	24.82	729	31277	23.31
2009	363	9058	40.08	631	28376	22.24	1015	38531	26.34	844	34255	24.64
2010	450	10176	44.22	717	29944	23.94	1157	40880	28.3	967	36453	26.53
2011	562	11181	50.26	842	30907	27.24	1376	42393	32.46	1153	37828	30.48
2012	661	12123	54.52	982	31551	31.12	1608	43648	36.84	1341	38944	34.43
2013	738	12874	57.32	1130	31978	35.34	1829	44570	41.04	1540	39691	38.8
2014	922	13101	70.38	1364	30492	44.73	2172	42896	50.63	1843	38218	48.22
2015	968	12500	77.44	1496	28016	53.4	2328	39543	58.87	1978	35277	56.07

2016	979	10885	89.94	1482	23189	63.91	2245	32801	68.44	1922	29259	65.69
2017	1057	10003	105.67	1539	20176	76.28	2343	28820	81.3	2026	25714	78.79
2018	1109	9556	116.05	1596	18195	87.72	2431	26179	92.86	2124	23356	90.94
2019	1210	9414	128.53	1819	17335	104.93	2740	24936	109.88	2409	22286	108.09

Supplementary 30: Incidence trend of gestational diabetes mellitus (GDM) among pregnant women with PCOS defined using diagnostic codes and a combination of symptom codes fulfilling each of the diagnostic criteria

Year	Incidence per 1,000 pregnant women											
	PCOS definition - Diagnostic code			PCOS definition – combination of symptom codes fulfilling National Institute of Child Health and Human Development criteria for PCOS diagnosis			PCOS definition – combination of symptom codes fulfilling Rotterdam criteria for PCOS diagnosis			PCOS definition – combination of symptom codes fulfilling Androgen Excess Society criteria for PCOS diagnosis		
	Nr	Dr	Incidence	Nr	Dr	Incidence	Nr	Dr	Incidence	Nr	Dr	Incidence
2000	0	24	0	3	1107	2.71	7	1384	5.06	3	1182	2.54
2001	0	56	0	7	1457	4.8	9	1824	4.93	7	1542	4.54
2002	1	75	13.33	6	1583	3.79	10	2090	4.78	8	1817	4.4
2003	4	121	33.06	7	1879	3.73	9	2540	3.54	8	2133	3.75
2004	4	195	20.51	14	2078	6.74	22	2868	7.67	17	2439	6.97
2005	4	421	9.5	14	2382	5.88	22	3163	6.96	15	2740	5.47
2006	10	546	18.32	5	2654	1.88	10	3559	2.81	8	3145	2.54
2007	14	934	14.99	24	2792	8.6	33	3759	8.78	26	3257	7.98
2008	12	929	12.92	25	3173	7.88	33	4115	8.02	32	3685	8.68
2009	12	1159	10.35	32	3473	9.21	52	4622	11.25	46	4053	11.35
2010	40	1181	33.87	32	3536	9.05	51	4652	10.96	43	4106	10.47
2011	28	1278	21.91	51	3442	14.82	86	4781	17.99	68	4191	16.23
2012	37	1330	27.82	43	3622	11.87	80	4791	16.7	58	4327	13.4
2013	32	1311	24.41	54	3072	17.58	87	4445	19.57	71	3941	18.02
2014	28	1314	21.31	38	2723	13.96	75	3888	19.29	63	3457	18.22
2015	42	1081	38.85	51	2442	20.88	88	3388	25.97	73	2954	24.71
2016	47	1021	46.03	43	1962	21.92	68	2668	25.49	53	2375	22.32
2017	54	965	55.96	23	1585	14.51	50	2261	22.11	43	1982	21.7
2018	43	793	54.22	53	1520	34.87	71	2129	33.35	61	1922	31.74
2019	18	700	25.71	32	1330	24.06	37	1875	19.73	34	1685	20.18

Supplementary 31: ICD-10 codes for outcome ascertainment – preterm birth

ICD-10 CODE	DESCRIPTION
O60	Preterm labour and delivery
O60.1	Preterm spontaneous labour with preterm delivery
O60.2	Preterm spontaneous labour with term delivery
O60.3	Preterm delivery without spontaneous labour
P07.2	Extreme immaturity
P07.3	Other preterm infants
P59.0	Neonatal jaundice associated with preterm delivery
P61.2	Anaemia of prematurity

Supplementary 32: OPCS codes for outcome ascertainment – mode of delivery

	OPCS CODE	DESCRIPTION
Emergency Caesarean section	R180	Other caesarean delivery
	R181	Upper uterine segment caesarean delivery NEC
	R182	Lower uterine segment caesarean delivery NEC
	R188	Other specified other caesarean delivery
	R189	Unspecified other caesarean delivery
Elective or other unspecified Caesarean section	R170	Elective caesarean delivery
	R171	Elective upper uterine segment caesarean delivery
	R172	Elective lower uterine segment caesarean delivery
	R178	Other specified elective caesarean delivery
	R179	Unspecified elective caesarean delivery
	R251	Caesarean hysterectomy
Instrumental vaginal delivery	R210	Forceps cephalic delivery
	R211	High forceps cephalic delivery with rotation
	R212	High forceps cephalic delivery NEC
	R213	Mid forceps cephalic delivery with rotation
	R214	Mid forceps cephalic delivery NEC
	R215	Low forceps cephalic delivery

	R218	Other specified forceps cephalic delivery
	R219	Unspecified forceps cephalic delivery
	R220	Vacuum delivery
	R221	High vacuum delivery
	R222	Low vacuum delivery
	R223	Vacuum delivery before full dilation of cervix
	R228	Other specified vacuum delivery
	R229	Unspecified vacuum delivery
Spontaneous or other unspecified vaginal delivery	R201	Spontaneous breech delivery
	R230	Cephalic vaginal delivery with abnormal presentation of head at delivery without instrument
	R231	Manipulative cephalic vaginal delivery with abnormal presentation of head at delivery without instrument
	R232	Non-manipulative cephalic vaginal delivery with abnormal presentation of head at delivery without instrument
	R238	Other specified cephalic vaginal delivery with abnormal presentation of head at delivery without instrument
	R239	Unspecified cephalic vaginal delivery with abnormal presentation of head at delivery without instrument
	R240	Normal delivery
	R249	All normal delivery
	R190	Breech extraction delivery
	R191	Breech extraction delivery with version
	R198	Other specified breech extraction delivery
	R199	Unspecified breech extraction delivery
	R200	Other breech delivery
	R208	Other specified other breech delivery
	R209	Unspecified other breech delivery
	R202	Assisted breech delivery
	R252	Destructive operation to facilitate delivery
	R258	Other specified other methods of delivery
	R259	Unspecified other methods of delivery

Supplementary 33: ICD-10 codes for outcome ascertainment – mode of delivery

	ICD-10 CODE	DESCRIPTION
Emergency Caesarean section	O82.1	Delivery by emergency caesarean section
Elective or other unspecified Caesarean section	O82.0	Delivery by elective caesarean section
	O82.2	Delivery by caesarean hysterectomy
	O82.8	Other single delivery by caesarean section
	O82.9	Delivery by caesarean section unspecified
	O84.2	Multiple delivery all by caesarean section
	O83.3	Delivery of viable fetus in abdominal pregnancy
Instrumental vaginal delivery	O81.0	Low forceps delivery
	O81.1	Mid-cavity forceps delivery
	O81.2	Mid-cavity forceps with rotation
	O81.3	Other and unspecified forceps delivery
	O81.4	Vacuum extractor delivery
	O81.5	Delivery by combination of forceps and vacuum extractor
	O84.1	Multiple delivery all by forceps and vacuum extractor
Spontaneous or other unspecified vaginal delivery	O80.0	Spontaneous vertex delivery
	O80.1	Spontaneous breech delivery
	O80.8	Other single spontaneous delivery
	O80.9	Single spontaneous delivery unspecified
	O84.0	Multiple delivery all spontaneous
	O83.0	Breech extraction
	O83.1	Other assisted breech delivery
	O83.2	Other manipulation-assisted delivery
	O83.8	Other specified assisted single delivery

	O83.9	Assisted single delivery unspecified
	O84.8	Other multiple delivery
	O84.9	Multiple delivery unspecified

Supplementary 34: ICD-10 code for outcome ascertainment – stillbirth

ICD-10_CODE	**DESCRIPTION**
Z37.1	Single stillbirth

Supplementary 35: **ICD-10 code for outcome ascertainment – High birthweight**

ICD-10_CODE	**DESCRIPTION**
P08.0	Exceptionally large baby

Supplementary 36: Baseline characteristics of women with PCOS and age matched controls – Sensitivity Analysis

Variables	Deliveries of women with PCOS*	Age matched deliveries of women without PCOS
All	**(n=4,559)**	**(n=18,236)**
Age at delivery [Mean (SD)]	31.13 (5.10)	31.12 (5.05)
Age at delivery [Median (IQR)]	31.00 (27.00-34.00)	31.00 (27.00-34.00)
Age categories, n (%)		
14 - 20 years	34 (0.75)	137 (0.75)
20 - 30 years	1850 (40.58)	7302 (40.04)
30 - 40 years	2491 (54.64)	10107 (55.42)
40 - 50 years	184 (4.04)	690 (3.78)
Pre-gravid BMI [Mean (SD)]	28.34 (6.94)	25.12 (5.46)
Pre-gravid BMI [Median (IQR)]	27.00 (22.00-32.00)	23.00 (21.00-27.00)
BMI Categories, n (%)		
<25 kg/m2	1656 (36.32)	9863 (54.09)
25-30 kg/m2	1115 (24.46)	4258 (23.35)
30-35 kg/m2	786 (17.24)	1680 (9.21)
35-40 kg/m2	477 (10.46)	641 (3.52)
>40 kg/m2	275 (6.03)	358 (1.96)
Missing	250 (5.48)	1436 (7.87)
IMD, n (%)		
1 Most deprived)	552 (12.11)	2218 (12.16)
2	457 (10.02)	1783 (9.78)
3	460 (10.09)	1866 (10.23)
4	445 (9.76)	1774 (9.73)
5	464 (10.18)	1970 (10.80)
6	377 (8.27)	1750 (9.60)
7	449 (9.85)	1720 (9.43)
8	450 (9.87)	1724 (9.45)
9	458 (10.05)	1668 (9.15)
10 (Least deprived)	445 (9.76)	1753 (9.61)
Missing	2 (0.04)	10 (0.05)
Ethnicity, n (%)		
White Caucasian	2401 (52.67)	8445 (46.31)
South Asian	370 (8.12)	573 (3.14)
Black Afro-Caribbean	286 (6.27)	854 (4.68)
Mixed Race	46 (1.01)	97 (0.53)
Others	118 (2.59)	416 (2.28)
Missing	1338 (29.35)	7851 (43.05)
Record of symptoms and measurements at baseline, n (%)		
PCO	1580 (34.66)	0 (0)
Hair Loss	295 (6.47)	425 (2.33)
Hirsutism	491 (10.77)	120 (0.66)
Anovulation	1910 (41.90)	1864 (10.22)

High Testosterone (serum testosterone level ≥ 2.0 nmol/L)	835 (18.32)	92 (0.50)
Other comorbidities, n (%)		
Type 2 diabetes	142 (3.11)	207 (1.14)
Prediabetes	265 (5.81)	333 (1.83)
Hypertension	114 (2.50)	212 (1.16)
Thyroid disorders	251 (5.51)	415 (2.28)
Pregnancy related variables	265 (5.81)	333 (1.83)
Number of babies at the delivery, n (%)		
1	4492 (98.53)	17908 (98.20)
2	64 (1.40)	318 (1.74)
3	3 (0.07)	9 (0.05)
4+	0 (0)	1 (0.01)

*Patients with a diagnostic code for PCOS only

PCOS: Polycystic Ovary Syndrome; PCO: Polycystic ovaries; SD: Standard Deviation; IQR: Interquartile Range; BMI: Body Mass Index; IMD: Index of Multiple Deprivation

Supplementary 37: Risk of primary obstetric outcomes among women with PCOS compared to women without PCOS – Sensitivity Analysis

Outcomes	Deliveries of women with PCOS*	Age matched deliveries of women without PCOS
Preterm		
Number of patients	4559	18236
Outcome events, n (%)	391 (8.58%)	1223 (6.71%)
Unadjusted OR (95% CI)	1.31 (1.15-1.48)	
Adjusted OR (95% CI) (Model 1)	1.29 (1.11-1.44)	
Adjusted OR (95% CI) (Model 2)	1.23 (1.06-1.44)	
Adjusted OR (95% CI) (Model 3)	1.29 (1.13-1.47)	
Adjusted OR (95% CI) (Model 4)	1.31 (1.13-1.52)	
Model of delivery		
Number of patients	4559	18236
Outcome events, n (%)		
Emergency CS	610 (13.38%)	1986 (10.89%)
Elective/Other/Unspecified CS	798 (17.50%)	2535 (13.90%)
Instrumental Vaginal	536 (11.76%)	2051 (11.25%)
Spontaneous/Other/Unspecified Vaginal	2615 (57.36%)	11664 (63.96%)
Unadjusted OR (95% CI)		
Emergency CS	1.37 (1.23-1.53)	
Elective/Other/Unspecified CS	1.40 (1.27-1.55)	
Instrumental Vaginal	1.17 (1.05-1.30)	
Spontaneous/Other/Unspecified Vaginal	Ref	
Adjusted OR (95% CI) (Model 1)		
Emergency CS	1.37 (1.23-1.54)	
Elective/Other/Unspecified CS	1.38 (1.25-1.52)	
Instrumental Vaginal	1.17 (1.05-1.30)	
Spontaneous/Other/Unspecified Vaginal	Ref	
Adjusted OR (95% CI) (Model 2)		
Emergency CS	1.31 (1.17-1.47)	
Elective/Other/Unspecified CS	1.36 (1.23-1.50)	
Instrumental Vaginal	1.18 (1.06-1.32)	
Spontaneous/Other/Unspecified Vaginal	Ref	
Adjusted OR (95% CI) (Model 3)		
Emergency CS	1.32 (1.17-1.48)	
Elective/Other/Unspecified CS	1.37 (1.24-1.51)	
Instrumental Vaginal	1.18 (1.06-1.32)	
Spontaneous/Other/Unspecified Vaginal	Ref	
Adjusted OR (95% CI) (Model 4)		
Emergency CS	1.17 (1.04-1.31)	
Elective/Other/Unspecified CS	1.25 (1.13-1.38)	
Instrumental Vaginal	1.24 (1.11-1.38)	
Spontaneous/Other/Unspecified Vaginal	Ref	
Adjusted OR (95% CI) (Model 5)		
Emergency CS	1.15 (1.02-1.30)	

Elective/Other/Unspecified CS	1.03 (1.02-1.03)	
Instrumental Vaginal	1.00 (1.00-1.00)	
Spontaneous/Other/Unspecified Vaginal	Ref	
High birthweight >4 kg (for at least one of the baby)		
Number of patients	4559	18236
Outcome events, n (%)	500 (10.97%)	1846 (10.12%)
Unadjusted OR (95% CI)	1.09 (0.98-1.22)	
Adjusted OR (95% CI) (Model 1)	1.12 (0.95-1.33)	
Adjusted OR (95% CI) (Model 2)	1.11 (0.94-1.31)	
Adjusted OR (95% CI) (Model 3)	1.11 (0.94-1.31)	
Adjusted OR (95% CI) (Model 4)	0.97 (0.84-1.12)	
Adjusted OR (95% CI) (Model 5)	1.00 (0.88-1.13)	
Low birthweight <2.5 kg (for at least one of the baby)		
Number of patients	4559	18236
Outcome events, n (%)	277 (6.08%)	1001 (5.49%)
Unadjusted OR (95% CI)	1.11 (0.96-1.29)	
Adjusted OR (95% CI) (Model 1)	1.09 (0.95-1.24)	
Adjusted OR (95% CI) (Model 2)	1.06 (0.93-1.21)	
Adjusted OR (95% CI) (Model 3)	1.09 (0.94-1.26)	
Adjusted OR (95% CI) (Model 4)	1.18 (1.00-1.39)	
Adjusted OR (95% CI) (Model 5)	1.03 (0.77-1.37)	
Stillbirth		
Number of patients	4559	18236
Outcome events, n (%)	16 (0.35%)	82 (0.45%)
Unadjusted OR (95% CI)	0.78 (0.46-1.32)	
Adjusted OR (95% CI) (Model 1)	0.74 (0.44-1.26)	
Adjusted OR (95% CI) (Model 2)	0.70 (0.27-1.77)	
Adjusted OR (95% CI) (Model 3)	0.66 (0.25-1.72)	
Adjusted OR (95% CI) (Model 4)	0.52 (0.27-1.02)	

*Patients with a diagnostic code for PCOS only

PCOS: Polycystic Ovary Syndrome; CS: Caesarean Section; OR: Odds Ratio

Model 1: Adjusted for age, ethnicity, and deprivation

Model 2: Adjusted for age, ethnicity, deprivation, baseline dysglycaemia, hypertension and thyroid disorders

Model 3: Adjusted for age, ethnicity, deprivation, baseline dysglycaemia, hypertension, thyroid disorders, and numbers of babies born at the delivery

Model 4: Adjusted for age, ethnicity, deprivation, baseline dysglycaemia, hypertension, thyroid disorders, numbers of babies born at the delivery, and pre-gravid body mass index

Model 5: Adjusted for age, ethnicity, deprivation, baseline dysglycaemia, hypertension, thyroid disorders, numbers of babies born at the delivery, pre-gravid body mass index, and gestational age

Supplementary 38: Risk of secondary obstetric outcomes among women with PCOS compared to women without PCOS – Sensitivity Analysis

Outcomes	Deliveries of women with PCOS*	Age matched deliveries of women without PCOS
Very Preterm (<32 weeks of gestational age at delivery)		
Number of patients	4559	18236
Outcome events, n (%)	125 (2.74%)	346 (1.90%)
Unadjusted OR (95% CI)	1.46 (1.16-1.84)	
Adjusted OR (95% CI) (Model 1)	1.44 (0.98-2.11)	
Adjusted OR (95% CI) (Model 2)	1.44 (0.87-2.39)	
Adjusted OR (95% CI) (Model 3)	1.43 (0.98-2.08)	
Adjusted OR (95% CI) (Model 4)	1.42 (0.88-2.31)	
Extremely preterm (<28 weeks of gestational age at delivery)		
Number of patients	4559	18236
Outcome events, n (%)	59 (1.29%)	137 (0.75%)
Unadjusted OR (95% CI)	1.73 (1.25-2.41)	
Adjusted OR (95% CI) (Model 1)	1.86 (1.33-2.60)	
Adjusted OR (95% CI) (Model 2)	1.97 (1.43-2.71)	
Adjusted OR (95% CI) (Model 3)	1.93 (1.23-3.05)	
Adjusted OR (95% CI) (Model 4)	1.86 (1.31-2.65)	
Large for gestational age >90th percentile (for at least one of the babies)		
Number of patients	4559	18236
Outcome events, n (%)	925 (20.29%)	3161 (17.33%)
Unadjusted OR (95% CI)	1.22 (1.11-1.33)	
Adjusted OR (95% CI) (Model 1)	1.23 (1.13-1.34)	
Adjusted OR (95% CI) (Model 2)	1.21 (1.11-1.32)	
Adjusted OR (95% CI) (Model 3)	1.21 (1.11-1.31)	
Adjusted OR (95% CI) (Model 4)	1.08 (0.99-1.18)	
Small for gestational age <10th percentile (for at least one of the babies)		
Number of patients	4559	18236
Outcome events, n (%)	146 (3.20%)	753 (4.13%)
Unadjusted OR (95% CI)	0.77 (0.64-0.92)	
Adjusted OR (95% CI) (Model 1)	0.72 (0.57-0.91)	
Adjusted OR (95% CI) (Model 2)	0.72 (0.55-0.95)	
Adjusted OR (95% CI) (Model 3)	0.70 (0.53-0.93)	
Adjusted OR (95% CI) (Model 4)	0.74 (0.59-0.94)	

*Patients with a diagnostic code for PCOS only

PCOS: Polycystic Ovary Syndrome; CS: Caesarean Section; OR: Odds Ratio

Model 1: Adjusted for age, ethnicity, and deprivation

Model 2: Adjusted for age, ethnicity, deprivation, baseline dysglycaemia, hypertension and thyroid disorders

Model 3: Adjusted for age, ethnicity, deprivation, baseline dysglycaemia, hypertension, thyroid disorders, and numbers of babies born at the delivery

Model 4: Adjusted for age, ethnicity, deprivation, baseline dysglycaemia, hypertension, thyroid disorders, numbers of babies born at the delivery, and pre-gravid body mass index

Supplementary 39: Read codes used for the ascertainment of outcome (confirmed/suspected COVID)

Read Code	Description
G558500	Cardiomyopathy due SARS-CoV-2
A795200	COVID-19 confirmed by laboratory test
G520800	Myocarditis due to SARS-CoV-2
H051100	URTI due to SARS-CoV-2
A076400	Gastroenteritis due to SARS-CoV-2
F529.00	Otitis media due to SARS-CoV-2
F289.00	Encephalopathy due to SARS-CoV-2
A795300	COVID-19 confirmed using clinical diagnostic criteria
H204.00	Pneumonia due to SARS-CoV-2
43hF.00	Detection of SARS-CoV-2 by PCR
4J3R100	2019-nCoV (novel coronavirus) detected
9N31200	Telephone consultation for suspected 2019-nCoV (novel coronavirus)
A795100	Disease caused by 2019-nCoV (novel coronavirus)
1JX1.00	Suspected disease caused by 2019-nCoV (novel coronavirus)
1JX..00	Suspected coronavirus infection
43dt400	Has immunity to SARS-CoV-2
43dtA00	SARS-CoV-2 IgG detected
43dtG00	SARS-CoV-2 IgM detected