Epidemiological investigation and economic analysis of Polycystic Ovary Syndrome (PCOS) for women in the UK

Tao Ding

Thesis submitted for the degree of PhD

Department of Statistical Science

University College London

Nov 2017

DECLARATION

I Tao Ding confirm that this thesis is a presentation of my original research work. While there is information derived from other sources, effort has been made to indicate clearly in the reference.

ACKNOWLEDGEMENTS

ABSTRACT

Background. Polycystic ovary syndrome (PCOS) is an endocrine disorder affecting millions of women worldwide. The pathophysiology of this condition is unclear but is believed to be caused by some genetic and environmental factors. PCOS is associated with a range of reproductive, metabolic and dermatological disorders and therefore, the economic burden of this condition can be potentially significant for the public health system in the UK.

Methods. The methodology used for this research includes several parts. Firstly, I conducted literature reviews to identify studies reporting the prevalence of PCOS and morbidities associated PCOS. The Bayesian hierarchical model was then applied to model data from the published studies. This forms the first part of this research for which the analysis was based on aggregate data. In the second part, I investigated the incidence and prevalence of PCOS under the specific UK context using data from The Health Improvement Network (THIN), a primary care database with over 500 general practices contributing data each year. I then used a multi-state Markov model to simulate the population dynamics of PCOS and evaluated the associated economic burden of care as well as the quality of life for the entire population with this condition in the UK.

Results. The prevalence estimates from community studies are generally much higher compared with that from database studies. The prevalence of PCOS varies for different diagnostic criteria and across distinct ethnic groups. Women with PCOS are at higher risk of type 2 diabetes, obesity, cardiovascular diseases and pregnancy complications and are more likely to experience psychological disorders. The prevalence of PCOS in the UK is estimated to be approximately 2% based on the primary care data, with an annual incidence rate of 2 per 1000 person-year. There is wide variation in the prescriptions initiated for the PCOS patients. The prevalence of

4

type 2 diabetes in the PCOS population is estimated to increase to approximately 26% in the next 25 years in the UK, which significantly reduces the quality of life for individual patients and incurs massive amount of healthcare-related costs for the National Health Service (NHS).

Conclusions. The large gap between the prevalence rates estimated from community and database studies suggests that PCOS is a condition without much public awareness and under-reporting is often observed. The differences in prevalence rates estimated according to different diagnostic criteria indicates the potential issue of under- and over-diagnosis of the condition at present. The ethnic variation in terms of the diagnostic criteria, disease monitoring and management may need to be considered carefully. The prescribing patterns of PCOS in the primary care suggest that currently, there is lack of most effective treatment for this condition and patients generally receive treatments tailored to their external symptoms. The prevalence of type 2 diabetes among PCOS patients is estimated to be high, resulting in massive amount of healthcare costs and reduced quality of life for PCOS population in the UK. Early screening is likely to help reduce the adverse outcomes associated with PCOS for this selected population and it may be cost effective to include them in the current Diabetes Prevention Programme. This may help improve the detection of symptoms indicative of diabetes in PCOS patients to allow early interventions and save significant amount of healthcare costs for the NHS from the country perspective.

Contents

THESIS AIM AND STRUCTURE

The aim of the research presented in the current thesis was to apply statistical approach for health research to explore various aspects of polycystic ovary syndrome (PCOS), a disease affecting millions of women worldwide.

Specific objectives were to:

- Conduct search to identify relevant literature reporting epidemiological outcomes of PCOS and use statistical approach to establish the prevalence of PCOS in population with different ethnicity backgrounds;

- Explore the incidence/prevalence and prescribing patterns of PCOS in the UK using primary care database;

- Investigate the population dynamics of PCOS and evaluate the healthcare-related economic burden and quality of life associated with this condition in the UK from 2014 to 2039.

The thesis consists of two main parts. In Part I, which consists of 3 chapters, I present existing evidence in the empirical studies and modelling results based on aggregate data from literature review for PCOS, which are presented in Chapter 1, 2 and 3. In Part II (including Chapter 4 and 5), I discuss analyses conducted for PCOS mainly under the specific UK context, including an exploratory analysis on the epidemiological outcomes and an economic evaluation using individual level patient data from primary care in the UK. Finally, Chapter 6 summarises and concludes the main findings from the current research and discusses the impact of the results from both a global and country perspective as well as the implication for practice and future research.

The thesis is structured as follows: Chapter 1 gives an overview of PCOS including the

discovery and major diagnostic criteria of PCOS, morbidities associated with PCOS and the common treatments for the condition. This chapter aims at providing the reader some background information of PCOS such that the reader can grasp some basic understanding about the disease.

Chapter 2 describes the proceedings and results of systematic literature reviews. This is of great importance because from the literature review, I obtained extensive knowledge about various aspects of PCOS. In addition, previous studies enable me to further investigate my main research questions into depths. This chapter introduces the methods and results of database searches, including details of search words, inclusion criteria for study selection, qualitative results from individual studies and quality assessment for each study considered. The majority of the data identified from this Chapter were further used to populate our statistical models, which is discussed in Chapter 3.

Chapter 3 introduces the methods and main findings for our prevalence estimation of PCOS and its associated diseases. I firstly provide some background information about evidence synthesis and meta-analysis, which are the main methodologies leading the whole chapter. I then discuss the specific modelling approach used to establish the prevalence of PCOS and its associated diseases by synthesising evidence from the published studies. Furthermore, the generalised meta-analytic models using Bayesian methods are described, which forms the basis of the package I created for R named bmeta. The bmeta package was motivated by the need to combine evidence from multiple sources for decision making. The implementation and graphical functions as well as the web application of bmeta are discussed through sample data for PCOS obtained from the literature review in Chapter 2.

Chapter 4 presents the methods and results from a database study where I used The Health Improvement Network (THIN) to evaluate the epidemiology and prescribing patterns of PCOS under the UK context. I firstly introduce the UK primary care and the motivation of this database study. The data analysis plan (i.e. patient identification process, statistical modelling approach for individual level patient data) and main findings for this database study are then described. Finally, I interpret the results and compare with evidence from empirical studies to make conclusions and discuss the potential implications.

In Chapter 5, I discuss the economic and utility analysis conducted for PCOS. I attempted to estimate the healthcare burden of PCOS in the UK in the next 25 years by using a Markov multi-

state model. Different health states through the disease progression were considered and included in the model. The model was populated by data from multiple sources and for parameters that lack enough data to inform, reasonable assumptions were used and uncertainties were assessed through sensitivity analysis. The results from the model included the epidemiology of type 2 diabetes in the PCOS population, disease burden and quality of life for PCOS population in the UK from 2014 to 2039. The implications of the main findings are described and some recommendations for improving the current screening and diagnosis of PCOS are raised.

Finally, in Chapter 6, I summarise the main findings throughout the chapters of the thesis and discuss the impact and implications of this research.

ABBREVIATIONS

Abbreviation	Meaning
ADA	American Diabetes Association
AES	Androgen Excess Society
AHD	Additional Health Data
BMI	Body Mass Index
CI	Confidence interval
CINAHL	Cumulated index to Nursing and Allied Health Literature
COC	Combined Oral Contraceptives
CPRD	Clinical Practice Research Database
CrI	Credible Interval
CVD	Cardiovascular Disease
DIC	Deviance Information Criteria
DM2	Type 2 diabetes
GP	General Practitioner
GPRD	General Practice Research Data
ICD	International Classification of Disease
ICU	Intensive Care Unit
IDF	International Diabetes Federation
IGT	Impaired glucose tolerance
INPS	In Practice Systems Ltd
IPD	Individual Patient Data

Abbreviation	Meaning
IUD	Intrauterine Devices
IVF	In Vitro Fertilisation
LOD	Laparoscopic Ovarian Drilling
MetS	Metabolic Syndrome
MM	Markov Model
NGT	Normal Glucose Tolerance
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIH	National Institute for Health
OCP	Oral Contraceptive Pills
ONS	Office for National Statistics
OR	Odds Ratio
PCO	Polycystic ovaries
PCOS	Polycystic ovary syndrome
POC	Progestin Oral Contraceptives
PVI	Postcode Variable Indicator
PY	Person Year
PYAR	Person Year At Risk
QALY	Quality Adjusted Life Years
QoL	Quality of life
SHBG	Sex Hormone Binding Globulin
THIN	The Health Improvement Network
VAMP	Value Added Information Medical Products
VRD	Vamp Research Databank
WHO	World Health Organisation

LIST OF TABLES

LIST OF FIGURES

Chapter 1

BACKGROUND

1.1 Chapter summary

Polycystic ovary syndrome (PCOS) has been recognised as the most common endocrine abnormality, affecting millions of women in reproductive age worldwide. In this chapter, I firstly introduce the discovery of PCOS and the diagnostic criteria associated with it. I then discuss PCOS-related diseases and common treatments prescribed to women with PCOS. This introductory chapter about PCOS aims at giving the reader a general understanding of the disease.

1.2 Discovery and definition of PCOS

Polycystic ovary syndrome (PCOS) remained clinically unexplored until 1935 when two American gynaecologists, Irving Stein and Michael Leventhal, reported seven women who presented with amenorrhea and excess growth of terminal hair had enlarged ovaries containing small collections of follicles (Adi and Tank, 2010). PCOS was called Stein-Leventhal Syndrome for a number of decades after their first description to honour the pioneering achievements by these two gynaecologists.

Since its discovery, many medical specialists have manifested great interests in this research area and, over the past years, the scientific community has witnessed a rapid increase in the amount of published literature for this complex syndrome. PCOS is nowadays recognised as a heterogeneous endocrine abnormality whose principle clinical characteristics include oligo-ovulation (light or infrequent menstrual periods) or anovulation (failure of the ovary to release

ova for more than 3 months), hyperandrogenism (elevated androgen level which may lead to acnes or excess growth of terminal hair on body and face), and polycystic ovaries (many small cysts in the ovary) (Sirmans and Pate, 2014). Its pathophysiology remains largely unknown but many believe that PCOS appears to be familial, with its various aspects differentially inherited from one generation to the next (Livadas and Diamanti-Kandarakis, 2013).

The phenotypes of PCOS are observed to vary in each case as well as in different ethnic groups and, for an individual case, symptoms and signs of PCOS also change gradually over time, making its diagnosis and evaluation challenging for clinical practice (Livadas and Diamanti-Kandarakis, 2013; Balen et al., 2005).

The three major diagnostic criteria for PCOS that are widely followed are the 1990 National Institutes of Health (which is abbreviate as "**NIH**", Zawadzki and Dunaif, 1992), the 2003 Rotterdam Consensus ("**Rotterdam**", Rotterdam and ASRM-Sponsored, 2004a,b) and the 2006 Androgen Excess Society ("**AES**", Azziz et al., 2006). Details of these 3 definitions are provided in Table 1.1.

Table 1.1: Three major diagnostic criteria of PCOS.

1990 NIH	**2003 Rotterdam**	**2006 AES**
Includes all of the followings: 1. Clinical and/or biochemical signs of hyperandrogenism 2. Menstrual dysfunction AND exclusion of other disorders*	Includes any 2 of the followings: 1. Clinical and/or biochemical signs of hyperandrogenism 2. Ovarian dysfunction (Oligo-ovulation or anovulation) 3. Polycystic ovaries AND exclusion of related disorders*	Includes all of the followings: 1. Hyperandrogenism (hirsutism and/or hyperandrogenemia) 2. Ovarian dysfunction (Oligo-anovulation and/or polycystic ovaries) AND exclusion of related disorders*

*Possibly exclude congenital adrenal hyperplasia, 21-hydroxylase-deficient non-classic adrenal hyperplasia, androgen-secreting neoplasms, androgenic/anabolic drug use or abuse, Cushing's syndrome, thyroid dysfunction, the syndromes of severe insulin resistance and hyperprolactinemia

Arguably, the NIH is the strictest criterion while the Rotterdam covers a broadest spectrum. The AES removes a combination of terms from the Rotterdam. Therefore, there is an inclusive and exclusive relationship between the three major criteria of PCOS: the Rotterdam contains the AES, which in turn contains the NIH.

1.3 PCOS-related disorders

PCOS is associated with a variety of reproductive, cardiometabolic and dermatologic abnormalities. Patients are often observed to have psychological problems as well as reduced quality of life (Azziz et al., 2005).

Menstrual irregularity, one of the major diagnostic criteria of this complex syndrome, is considered to be prevalent among the PCOS population in reproductive age (Fauser et al., 2012). The cyclic vaginal bleeding (generally known as "menses" or "periods"), normally occurs at intervals within the range of 22 to 35 days, with variation in length of no more than 2 or 3 days every month (Balen et al., 2005; Azziz et al., 2005). However, the frequency of ovulation for PCOS patients is largely unpredictable (Fauser et al., 2012). The presence and severity of ovulation dysfunction is related to many factors, for instance body mass index (obese or not), or whether the patient has other hormonal disorders (Balen et al., 2005). As menstrual disturbance is a common clinical presentation for PCOS patients, their overall fecundability is reduced and a great number of PCOS patients is reported to suffer from infertility (Azziz et al., 2005). Apart from infertility, there is evidence that menstrual dysfunction may result in an increased risk of endometrial hyperplasia (Cheung, 2001), which is caused by the continuous secretion of unopposed oestrogen without regular ovulation (Hardiman et al., 2003). This may further lead to various gynaecological neoplasia although currently we still lack enough supporting evidence due to the scarcity of large-scale studies (Balen et al., 2005).

A growing body of literature has suggested that PCOS patients are at increased risks for cardiometabolic disturbances (Moran et al., 2010b) because insulin resistance is observed to appear in approximately 50-70% of PCOS population (Ovalle and Azziz, 2002). The most common metabolic disorder associated with PCOS is type 2 diabetes and it is suggested that the healthcare-related economic burden of PCOS is largely attributed to diabetic care (Azziz et al., 2005). The elevated level of insulin also stimulates the ovary to generate excess androgen and inhibits the production of sex hormone-binding globulin (SHBG), with the latter process allowing more free androgen as well as oestrogen to be biologically active (Tsilchorozidou et al., 2004). This "vicious cycle" of continued excess hormonal secretion aggravates anovulation and possibly endometrial hyperplasia that mentioned previously (Teede et al., 2010a).

Not surprisingly, it is also common for PCOS patients to have androgen excess diseases such as hirsutism—the presence of excess terminal hairs in a male pattern in women (Bode et al.,

23

2012), acne and alopecia. Studies have indicated that acne and hirsutism appear more frequently in younger women while older women are more likely to experience alopecia (Balen et al., 2005). However, due to the fact that levels of androgen in circulation are negatively associated with age, it is expected that the degree of various symptoms of hyperandrogenism may decrease gradually over time (Winters et al., 2000; Piltonen et al., 2004).

As women with PCOS are plagued with a great number of complications across the lifespan, many of them experience different degree of psychological disorders and have reduced quality of life. For example, infertility may lead to tensions between family members; acne and hirsutism can result in altered self-perception and low self-esteem; metabolic disorders such as type 2 diabetes mellitus can cause severe problems at work (Balen et al., 2005; Li et al., 2011a).

Overall, it is universally accepted that PCOS is a complex and heterogeneous disorder adversely affecting women throughout the lifespan. This pinpoints the importance of early detection, long term monitoring and management of PCOS for clinical practice as well as the public awareness of this syndrome.

1.4 Treatments of PCOS

The exact ætiology of PCOS remains unclear and is believed to be due to some inherited genetic characteristic. Consequently, PCOS itself is largely incurable and treatment are tailored to various health consequences of PCOS (Barthelmess and Naz, 2014; Teede et al., 2010b).

As PCOS is largely caused by hormonal imbalances, many treatments are designed to break the "vicious cycle" of excess hormonal production. Oral contraceptive pills (OCP) are the most common treatment for PCOS women presenting menstrual disturbances but not desiring pregnancy (Barthelmess and Naz, 2014; Meyer et al., 2007). They help regulate the cycle of menses and protect from endometrial hyperplasia (Teede et al., 2010b), with an additional advantage of reducing androgen production, which relieves the clinical symptoms of hyperandrogenism (Barthelmess and Naz, 2014; Meyer et al., 2007). However, there are concerns that high doses of OCP may be a risk factor for both metabolic and cardiovascular diseases by worsening glucose tolerance and arterial stiffness (Meyer et al., 2007).

This draws the attention to another treatment, metformin, which can ameliorate clinical presentations of PCOS such as menstrual irregularity and hirsutism without negative cardiometabolic impacts (Teede et al., 2007). It is also suggested by the International Diabetes

Federation (IDF) that metformin can, to some extent, prevent diabetes when lifestyle modification is not an adequate approach (Meyer et al., 2007; Alberti et al., 2007).

While both oral contraceptive pills and metformin are recommended treatments for PCOS by numerous endocrine societies worldwide, it should be noted that neither of them are approved specifically for PCOS by a majority of regulatory bodies (Teede et al., 2010b).

Lifestyle modification, including dietary therapies and regular exercises, is another treatment option and has been proved to be especially suitable for obese PCOS women (Moran et al., 2009) as obesity worsens hormonal imbalances and exacerbates many adverse health consequences associated with PCOS (Legro, 2012). There is evidence that for the obese PCOS population, weight loss is significantly beneficial to ameliorate anovulation, menstrual irregularity, infertility, psychological disorders as well as cardiometabolic risk factors associated with PCOS (Clark et al., 1998; Huber-Buchholz et al., 1999). Even for PCOS women who have normal weight, weight control should still be emphasised (Barthelmess and Naz, 2014).

Several treatments have been proved to be effective for infertility. Ovulation induction using clomiphene citrate or clomiphene is probably the simplest one (Barthelmess and Naz, 2014; Azziz et al., 2005). Another method of ovulation induction called laparoscopic ovarian drilling (LOD) can be used in case that clomiphene citrate therapy fails to induce ovulation (Barthelmess and Naz, 2014). This has achieved a remarkably high success rate of 84% in infertile PCOS women (Gjønnaess, 1994). A major advantage of this method is that it improves hormonal balances. Specifically speaking, both insulin resistance and excess androgen production are improved by means of LOD (Seow et al., 2007), which also lowers risks of multiple pregnancies and ovarian hyperstimulation syndrome frequently associated with clomiphene citrate therapy (Flyckt and Goldberg, 2011). Furthermore, there is a higher chance for PCOS women to have a second child by using LOD (Nahuis et al., 2011).

In vitro fertilisation (IVF) is probably a last resort for infertile PCOS women if the above methods fail to create a pregnancy (Barthelmess and Naz, 2014), with similar rates achieved as women received a mechanical tubal factor (Homburg et al., 1993). A new method called in vitro maturation has been utilised although its implantation has not achieved rates as high as that of traditional IVF procedure. However, it reduces the multiple pregnancy rate while maintains a similar delivery rate compared to outcomes of IVF (Shalom-Paz et al., 2012).

The most common treatments of hirsutism are pharmacological and cosmetic therapies

(Moghetti and Toscano, 2006; Azziz et al., 2005). Pharmacological therapies aim at reducing the androgen level in hirsute PCOS women while cosmetic therapies help remove the excess terminal hair. In terms of pharmacological therapies, oral contraceptives, which suppress excess androgen production by the ovary, is widely used although limited efficacy is observed. Some studies have suggested that anti-androgen drugs such as cyproterone acetate, spironolactone and flutamide appear to be highly effective treatments for hirsutism (Moghetti and Toscano, 2006). Laser treatment, which can result in large reduction in the severity of excess facial hair, is generally considered to be an optimal alternative for hirsutism (Clayton et al., 2005). It is also worth mentioning that eflornithine cream has recently become a licensed treatment, with a main function of removing excess facial hair (Moghetti and Toscano, 2006).

Acne, another dermatological disorder caused by hyerandrogenism in PCOS patients, is often treated by a combination of anti-androgens (mentioned previously as an option for treating hirsutism) and traditional methods. Applications of retinoids, tretinoin, antibacterial agents, azelaic acid, benzoyl peroxide as well as antibiotics such as topical treatments have been proved to be effective for ameliorating acne but each of them has limitations (Moghetti and Toscano, 2006; Balen et al., 2005). For example, tretinoin may lead to skin irritation (Moghetti and Toscano, 2006).

In conclusion, treatments of health consequences associated with PCOS are of a wide variety and many of them are interrelated. Therefore, an integrated care plan accounting for various aspects of PCOS is probably in need to achieve best treatment outcomes for PCOS population.

In the next chapter, I will introduce the evidence on major aspects of PCOS obtained from systematic literature review, which provide further insight into the epidemiology of PCOS and its associated morbidities.

Chapter 2

PREVALENCE AND COMORBIDITIES OF PCOS: EVIDENCE FROM LITERATURE

2.1 Chapter summary

This chapter presents existing evidence on major aspects of PCOS identified through systematic literature reviews conducted in common online databases including PubMed, EMBASE, The Cochrane Library and CINAHL. This mainly includes data from published studies in terms of the epidemiology of PCOS in the general population and the prevalence of morbidities associated with PCOS (e.g. type 2 diabetes, metabolic syndrome) among PCOS population. The methods used for the literature search and results from qualitative analysis are presented below.

2.2 Aim

The aim of the reviews was firstly to expand my knowledge of PCOS and then collect relevant data to explore its epidemiology. Given there had been no review for the prevalence of PCOS in the general population when the search was conducted, it was hoped that the findings from the review can provide guidance for this aspect and if sufficient data were identified, statistical modelling can be used to investigate the epidemiology of PCOS under a specific country context or for different ethnic populations. Another objective was to collect relevant data for common

metabolic disorders associated with PCOS, which can give me a better understanding of the relationship between PCOS and other disease areas and potentially be relevant when proceeding to the latter part of this thesis, e.g. economic evaluation of PCOS in the UK (discussed in detail in Chapter 5).

2.3 Methods

2.3.1 Search strategies

A literature review for **prevalence studies** of PCOS was conducted up to January 2015 in the following electronic databases: PubMed, The Cochrane Library, EMBASE, CINAHL. These are common databases where medical literature was indexed. The following combination of essential search words were used to identify studies evaluating epidemiology of PCOS:

```
((Stein-Leventhal syndrome) OR (polycystic ovary syndrome) OR
(PCOS)) AND ((prevalence) OR (incidence) OR (epidemiology))
```

The search was restricted to the English language.

Another search was conducted to identify studies that reported relative risk of common metabolic disorders associated with PCOS, including impaired glucose tolerance (IGT), type 2 diabetes (DM2), metabolic syndrome(MetS). The methods for the key words search followed exactly the same as the previous review (Moran et al., 2010b). The updated search extended from March 2009 to January 2015 was conducted in the following databases: PubMed, EMBASE, CINAHL (Cumulated index to Nursing and Allied Health Literature). Only articles published in English language were considered.

2.3.2 Inclusion and exclusion criteria

The inclusion/exclusion criteria of **prevalence studies** were developed considering a list of factors, e.g. population, intervention, comparator, outcomes and study type (abbreviated as PICOS), as displayed in Table 2.1.

In terms of **PCOS-related metabolic disorders**, I followed the selection criteria as the previous review (Moran et al., 2010b).

Table 2.1: Eligibility criteria for study inclusion/exclusion.

Criterion	Description
Population	Restricted to the general population, excluding patients seeking care for specific diseases, e.g. type 2 diabetes, infertility
Intervention	Not applicable
Comparator	Not applicable
Outcomes	Epidemiological outcomes of PCOS, i.e. prevalence, incidence
Study type	Studies were excluded if they were: 1. Reviews of PCOS 2. Prevalence of PCOS-related disorders

2.3.3 Data extraction

The following information was extracted from each of the **prevalence studies** considered:

- General characteristics of the study including author, year of publication, study period and location

- Characteristics of the study sample, e.g. sampling scheme, sample size, number of PCOS cases, age range, body mass index (BMI) and ethnicity

- Definition of PCOS (3 major criteria as suggested in Table 1.1, ICD-9 codes, medical diagnosed PCOS, clinical PCOS, self-reported PCOS)

The data from studies of PCOS-related metabolic disorders were extracted in the same way as the previous review, including setting and recruitment methods, diagnostic criteria of PCOS and metabolic disorders (i.e. IGT, DM2, MetS), characteristics of sample population (e.g. age, BMI), the total number of individuals who developed metabolic disorders in cases and controls as well as the total number of individuals in case and control arm.

2.3.4 Quality assessment

To ensure the quality of studies included in the qualitative analysis and quantitative analysis in Chapter 3, a methodological evaluation was performed for all the **prevalence studies**. Given that few instruments have been designed to evaluate prevalence studies, I referred to the Newcastle-Ottawa Scale (Stang, 2010) and the Joanna Briggs Institute critical appraisal tool (Munn et al., 2014) and modified a few items within some of their categories. Seven items with respect to various aspects were included (shown in Table 2.2):

Table 2.2: Definition of items and score awarding criteria in quality assessment for prevalence studies.

Item assessed	Score 1	Score 0
Sampling	Appropriate sampling with target population clearly defined and probability sampling applied	Inappropriate sampling, i.e. convenient sampling **OR** cases based on medical records where schematic sampling was not applied **OR** sampling scheme not stated/unknown
PCOS measured reliably and objectively	Systematic screening performed for sample population and PCOS strictly defined	Medical records based on studies where no systematic screening was performed **OR** studies which used self-reported PCOS (e.g. based on questionnaire)
Response rate	Low non-participant rates (<30%) of the initial target sample population for further study (systematic screening) **OR** low non-response rate (<30%) of a deliberated designed questionnaire	High non-response rate or refusal rate to further study (≥30%) of the initial target sample population **OR** high non-response rate (≥30%) of a deliberated-designed questionnaire **OR** studies based on electronic medical record where incomplete patient information (i.e. missing data) is a routine problem, leading to incomplete ascertainment of cases **OR** not stated/unknown
Sample size	Sample size clearly stated	Sample size not stated/unknown
Crude number of cases	Crude number of cases clearly stated	Crude number of cases not clearly stated
Age range	Age range of the sample population is approximately same as the reproductive age, e.g. 15-45 years, 18-45 years, 17-45 years	Otherwise (narrower age range **OR** upper/lower bound of age range lying outside the limit, e.g. 18-24 years, 12-44 years **OR** not stated/unknown)
Ethnicity	Ethnicity of PCOS cases and sample population clearly stated	Ethnicity not clearly stated/unknown

The Newcastle-Ottawa scale was applied as the critical appraisal tool to assess the quality of studies of PCOS-related metabolic disorders, as suggested by the previous review by Moran et al. (2010b). The Newcastle-Ottawa scale for the current analysis mainly assesses the following three aspects:

- Selection of subjects in PCOS and control group

  (a) Adequate case definition

- yes, with independent validation[*] [1]

- yes, i.e. record linkage or based on self-reports

- no description

(b) Representativeness of cases

- consecutive or obviously representative series of cases[*]

- potential for selection bias or not stated

(c) Selection of controls

- community controls[*]

- hospital controls

- no description

(d) Definition of controls:

- no history of disease (end-point)[*]

- no description of source

- Comparability of cases and controls on the basis of the study design and analysis

  (a) control for important factor such as age and BMI (most important factors)[*]

  (b) control for any additional factors[*]

- Exposure

  (a) ascertainment of exposure

  - secure record (i.e. surgical records)[*]

  - structured interview where blind to case/control status[*]

  - interview not blinded to case/control status

  - written self-report or medical record only

  - no description

---

[1] A star (⋆) can be awarded to this item if a study provides adequate case definition. Similarly, for the remaining items, a star indicates providing adequate and quality information for the item assessed.

(b) same method of ascertainment for case and control group

- yes*

- no

(c) non-response rate

- same rate for both groups*

- non-respondents described

- rate different and no designation

Note that studies with 3 or 4 stars in the selection domain AND 1 or 2 stars in comparability domain AND 2 or 3 stars in exposure domain were considered to be of good quality; those with 2 stars in selection domain AND 1 or 2 stars in comparability domain AND 2 or 3 stars in exposure domain were considered to be of fair quality; those with 0 or 1 star in selection domain OR 0 stars in comparability domain OR 0 or 1 stars in exposure domain were considered to be of poor quality.

2.4 Results

2.4.1 Search results

For **prevalence studies**, the search initially produced 2420 citations in total. Papers with irrelevant titles or abstracts, e.g. reviews of epidemiology of PCOS or prevalence studies of PCOS-related diseases, were excluded leaving 41 records for further consideration. The reference list of included studies and relevant reviews were searched in order to locate other potential eligible articles and 3 additional studies were identified. Two presentation posters were excluded as these studies did not appear to have been published in peer reviewed journals. One study was also excluded after full-text screening because it evaluated prevalence of PCOS in patients with DM2. Figure 2.1 presents a consort diagram summarising the search.
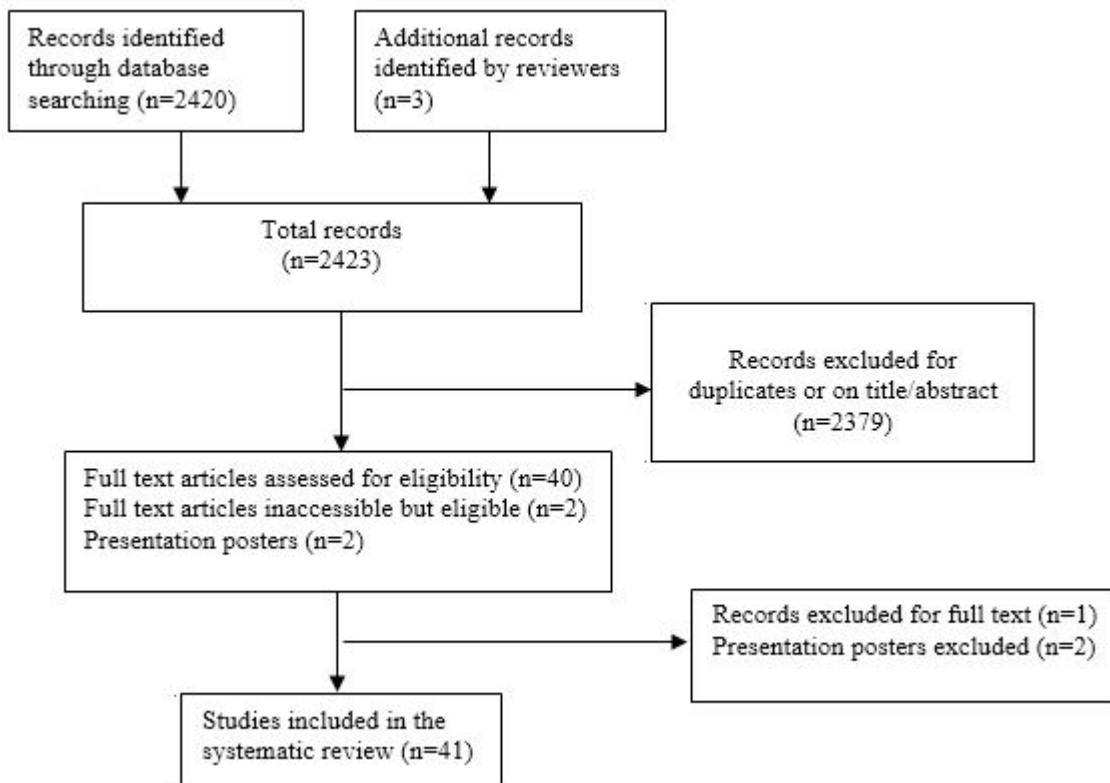
Figure 2.1: Flow chart of search process of prevalence studies.

In terms of major **PCOS-related diseases**, seven reviews and meta-analyses were identified and presented in Table 2.3. Notice that I further collected more studies for IGT, DM2 and MetS with the updated number of studies. The search end date and the new PCOS criterion included

are shown in the brackets. The consort diagram for the updated search for IGT, DM2 and metabolic syndrome related to PCOS is shown in Figure 2.2. The extended search initially yielded 2029 articles and after excluding irrelevant titles or abstracts, 128 studies were included for full-text screening. Of these, 76 were removed due to various reasons, e.g. duplicates (studies with different titles but included same study population thus reporting repeated results), lack of data or unable to determine prevalence or incidence of IGT, DM2 or MetS, studies without controls. In the previous review by Moran et al. (2010b), 35 studies were included and after this updated search, I eventually found 87 eligible studies for inclusion.

Table 2.3: Systematic reviews and meta-analyses of major PCOS-related diseases.

Article	Disease	No. of studies	Search end date	Sources	PCOS criteria
Moran et al. (2010b)	IGT DM2 MetS	35 (87)	Mar. 2009 (Jan.2015)	MEDLINE CINAHL EMBASE	NIH Rotterdam (AES)
Anderson et al. (2014)	Cardiovascular diseases	9	Up to 2014	MEDLINE EMBASE	WHO NIH Rotterdam AES Self-reported
Qin et al. (2013)	Pregnancy complications	27	Jul. 2012	PubMed MEDLINE EMBASE Cochrane	NIH Rotterdam AES
Barry et al. (2014)	Endometrial, ovarian and breast cancer	11	Oct. 2013	MEDLINE EMBASE	Most Self-reported 1 Rotterdam
Lim et al. (2012)	Obesity	106	Nov. 2010	MEDLINE EMBASE CINAHL Cochrane PSYCINFO	NIH Rotterdam
Barry et al. (2011)	Anxiety Depression	12	Dec. 2010	MEDLINE EMBASE Cochrane PubMed	Mostly Rotterdam Self-reported
Li et al. (2011b)	Quality of life	5	Dec. 2009	MEDLINE EMBASE CINAHL EMBR CNKI Wanfang Vip	2 NIH 3 Rotterdam
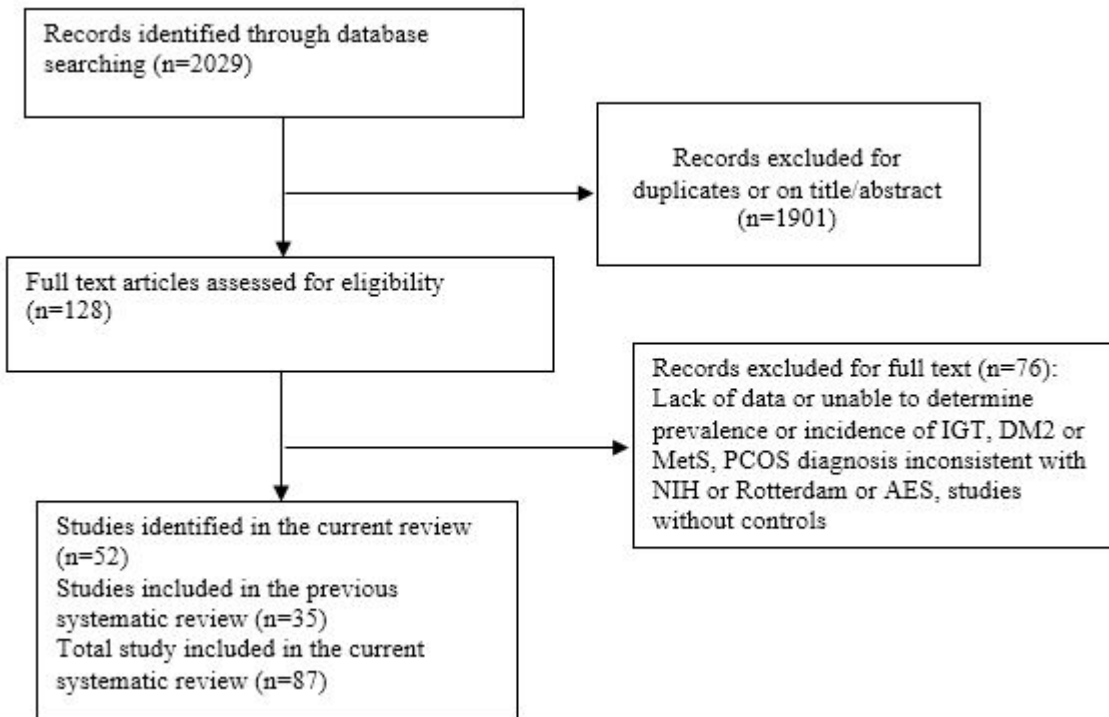
Figure 2.2: Flow chart of search process of metabolic disorders associated with PCOS.

2.4.2 Qualitative analysis

Results of prevalence

Forty-one studies were identified evaluating the epidemiology of PCOS in the general population but I only retrieved the full text for 39 of them, with the remaining 2 articles inaccessible: there was no link to follow on the websites where these two articles were indexed and attempts to contact the original authors also failed to receive responses. However, in the abstracts of these articles, Jiao et al. (2013) clearly presented source of recruitment, sample size, age range, ethnicity, definition of PCOS, prevalence data and the corresponding crude number of PCOS cases; Sung et al. (2010) provided source of recruitment, age range, definition of PCOS, prevalence data, with the crude number of PCOS cases and sample size irretrievable.

Geographically speaking, there were 10 studies in America, 6 in Europe, 11 in Asia, 10 in the Middle East and 4 in Oceania. All studies were cross-sectional, with 4 (Lo et al., 2006a; Okoroh et al., 2012; Christensen et al., 2013; Sirmans et al., 2014a) in the US reported using databases

and 3 (Jiao et al., 2013; Al Khaduri et al., 2014; March et al., 2010) using a retrospective study design. The sample size of the studies ranged from 50 to 12,171,830. Most of the studies included sample population within a reproductive age range (15-45 years). The weight status of the sample population varied significantly for sample population in different studies. The characteristics of each study (displayed by geographical locations) are listed in Appendix A.

The prevalence of PCOS ranged from as low as 0.56% (Christensen et al., 2013) using the NIH criterion in the US population to as high as 22.5% (Joshi et al., 2014) using the Rotterdam criterion in Indian population, as marked by the data labels in Figure 2.3.

Studies of PCOS-related morbidities

The results from seven meta-analyses for PCOS-related morbidities suggested that the prevalence of a range of disorders is higher in PCOS patients and women with PCOS are more likely to suffer from psychological problems, thus having reduced quality of life compared with the healthy controls (shown in Table 2.4).

The updated search for **metabolic disorders associated with PCOS** included 52 studies in additional to previous reviews. The number of studies for IGT/DM2 and MetS distributed across different geographical locations is displayed in Figure 2.4.

The detailed characteristics of studies assessing IGT/DM2 and MetS in women with and without PCOS are presented in Appendix B and C, respectively. Of all the studies eligible for inclusion, we found 11 assessing PCOS by all the 3 major diagnostic criteria. Seven studies were identified to use more than one definition of metabolic syndrome. The majority of the studies assessed women of reproductive age (15-45 years) with the exception of 15 studies where sample population consisted of either adolescents only or post-menopausal women only. Most studies included overweight and/or obese PCOS patients with a mean body mass index (BMI) $\geq$ $25kg/m^2$. Only 15 studies reported including PCOS patients with normal weight as a subgroup or the mean or median BMI of sample population smaller than $25kg/m^2$ or $< 23kg/m^2$ for East Asian countries.

There are 50 studies assessing prevalence or incidence of either IGT or DM2 or both. The majority of the studies used the NIH criterion of PCOS and 5 studies were found to use more than one definition of PCOS. In terms of the definition of IGT and DM2, 14 studies used World Health Organisation (WHO) criteria of IGT (Alberti and Zimmet, 1998) and 16 used ADA

definition (American Diabetes Association, 2010). Six studies used OGTT but did not state the specific cut-off value for diagnosis of IGT and DM2. Fourteen used other definitions (National Diabetes Data Group, 1979) or did not specifically state.

For the assessment of MetS, a total number of 50 studies were identified. Most of the studies applied the Rotterdam criterion for PCOS. Ten studies used more than one diagnostic standard of PCOS and 7 studies reported using more than one definition of MetS. Twenty-eight studies used Adult Treatment Panel III or modifications (Williams et al., 2002) whereas 10 used International Diabetes Federation (IDF) or modifications (Alberti et al., 2005). Seven studies used American Heart Association/National Heart Lung Blood Institute or modifications (Grundy et al., 2005) and 4 used WHO criteria (Alberti and Zimmet, 1998). Fourteen studies were found to apply other criteria (Cook et al., 2003; de Ferranti et al., 2004; Coviello et al., 2006; Caliskan et al., 2007; Rotterdam and ASRM-Sponsored, 2004a,b) or not specifically state the definition of MetS.

The majority of the studies reported higher prevalence of IGT, DM2 or MetS among women with PCOS compared with controls. While most of the studies assessed the point prevalence of metabolic disorders of interest, a few long-term follow-up studies were identified. Five studies evaluated the incidence of IGT and DM2 comparing women with and without PCOS. Legro et al. (1999) performed a 2- to 3-year follow-up, reporting an odds ratio of 2.7 (95% CI: 0.7, 8.0) for conversion from normal glucose tolerance to IGT/DM2 in 71 women with PCOS and 25 controls. Boudreaux et al. (2006) conducted an 8-year follow-up study, presenting an odds ratio of 2.07 (95% CI: 0.68, 6.30) for incidence of DM2 in 97 PCOS patients and 95 controls. Hudecova et al. (2011a) followed age-matched populations over a period of 13.9 years and found that 9.5% of the women with PCOS developed IGT and 8.3% had DM2. The corresponding proportions for controls were 2.3% for IGT and 1.1% for DM2, respectively. Schmidt et al. (2011) conducted a prospective 21-year follow-up study using 35 PCOS patients and 120 age-matched controls. By the end of study, 7 out of 32 (22%) PCOS patients and 13 of 95 (14%) controls had diabetes although the difference was not statistically significant. Another study by Celik et al. (2014) reported that among women with PCOS who had normal glucose tolerance (NGT) at baseline, 11.5% were observed to develop IGT. For PCOS patients who had IGT at baseline, 33.3% had converted into DM2. In contrast, 2.3% of the controls with NGT at baseline had IGT at follow-up and no controls with IGT at baseline were observed to convert into DM2 later

on. One study by Hudecova et al. (2011b) was found to assess the incidence of MetS comparing 84 women with PCOS and 87 controls. This study reported that 23.8% of women with PCOS and 8% of controls had developed MetS by the end of study period.
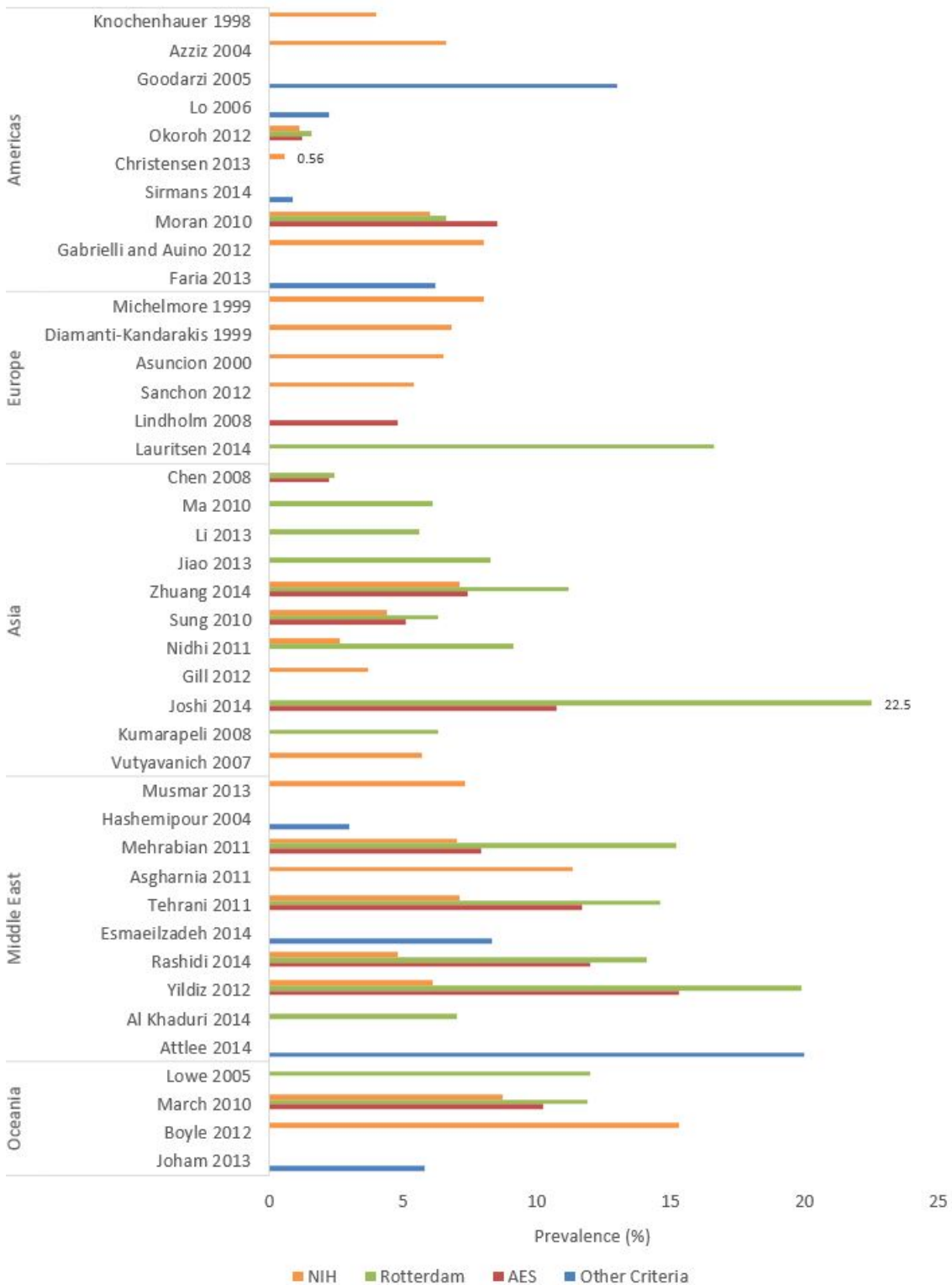
Figure 2.3: Summary of prevalence results reported by studies identified through literature review (categorised by geographical location). The results are presented based on the diagnostic criteria of PCOS. The two studies with a data label are the ones that reported the lowest and highest prevalence of PCOS.
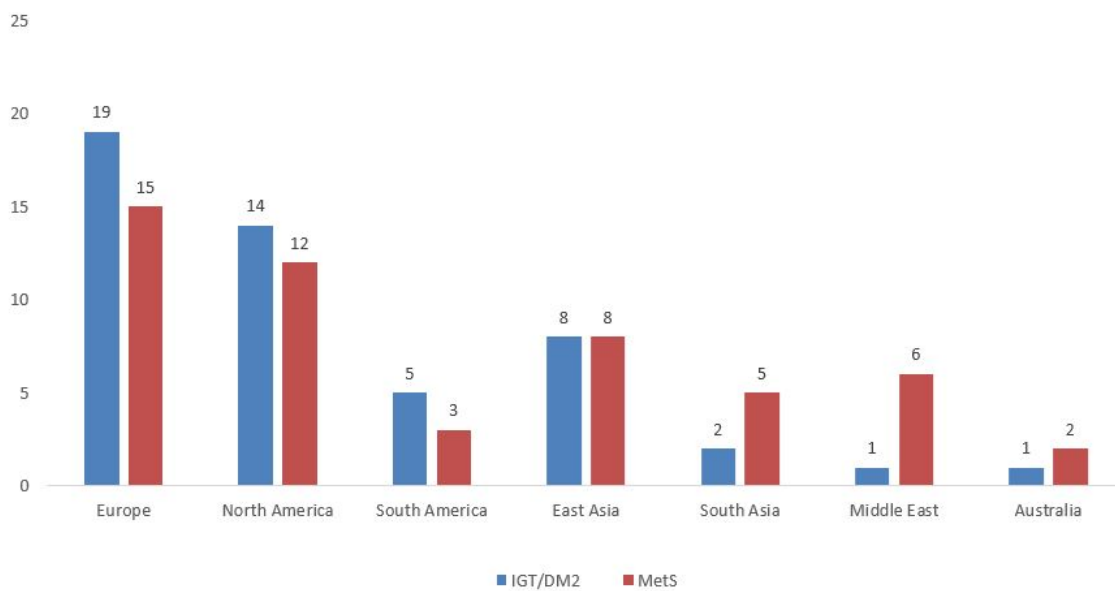
Figure 2.4: The distribution of studies across geographical locations. The bars in blue and red represent the number of studies identified in each continent reporting data on impaired glucose tolerance (IGT)/type 2 diabetes (DM2) and metabolic syndrome (MetS), respectively.

Table 2.4: Summary of results from reviews for PCOS-related morbidities.

Study	Sample size	Main results
Moran et al. (2010b)	IGT: 1403 DM2: 59064 MetS: 6386	Higher prevalence of IGT, DM2 and MetS in PCOS compared to controls: Odds Ratios (95% CI): IGT: 2.48 (95% CI: 1.63, 3.77) DM2: 4.43 (95% CI: 4.06, 4.82) MetS: 2.88 (95% CI: 2.40, 3.45)
Anderson et al. (2014)	Stroke: 3194 CHD: 3974	Higher prevalence of non-fatal stroke and coronary heart disease in women with PCOS but is not significant Odds Ratios (95% CI): Stroke: 1.61 (95% CI 0.82, 3.15) Coronary heart disease (CHD): 1.63 (95% CI 0.96, 2.78)

Study	Sample size	Main results
Qin et al. (2013)	124674	Higher prevalence of pregnancy complications in women with PCOS: Odds Ratios (95% CI): Gestational diabetes: 3.43 (95% CI: 2.49, 4.74) Pregnancy-induced hypertension: 3.43 (95% CI: 2.49, 4.74) Preeclampsia: 2.17 (95% CI: 1.91, 2.46) Preterm birth: 1.93 (95% CI: 1.45, 2.57) Caesarean section: 1.74 (95% CI: 1.38, 2.11) Babies delivered by PCOS patients have lower birth weight (weighted mean difference 0.11g 95% CI: -0.19, -0.03), and higher risk of admission to neonatal ICU (OR 2.32 95% CI: 1.40, 3.85)
Barry et al. (2014)	72973	Higher prevalence of endometrial cancer in women with PCOS (OR 2.79 95% CI 1.31, 5.95) Higher prevalence of ovarian and breast cancer in women with PCOS but is not significant (both ORs include 1)
Lim et al. (2012)	Overweight: 7633 Obesity: 9045	The prevalence of overweight and obesity is significantly higher in women with PCOS compared with controls (RR 1.95 95% CI: 1.52, 2.50; RR 2.77 95% CI: 1.88, 4.10 respectively)
Barry et al. (2011)	Depression: 2257 Anxiety: 377	Higher depression (mean difference: 0.82 95% CI: 0.73, 0.92) and anxiety (mean difference: 0.54 95% CI: 0.33, 0.75) measured in terms of score are observed in women with PCOS compared with controls

Study	Sample size	Main results
Li et al. (2011b)	708	Women with PCOS demonstrate a significant lower score in all SF-36 dimensions compared with controls, with a reduction in score as high as -23.86 (95% CI: 27.51, 20.21) for emotional role function and as low as -4.55 (95% CI: 7.99, 1.11) for body pain.

2.4.3   Results from quality assessment

The results from the quality assessment for **prevalence studies** are presented in Table 2.5 (ordered by geographical location). It should be noted that most of the prevalence studies scored 5 and above (n=27), with 8 studies rated very low score (i.e. ≤ 3).

Given that there is a large number of **studies reporting data on metabolic disorders associated with PCOS**, the results of quality assessment for these studies from my updated search are presented in Appendix D. In a nutshell, most of the studies identified are cross-sectional using non-randomised sampling. Consequently, the majority of the studies included were considered to be at least at some degree of selection bias and reporting bias. For diagnosis of PCOS, the exclusion of related reproductive disorders that may mimic symptoms of PCOS was absent in 23 studies and the medication use that might affect the outcomes was not assessed in 32 studies. The majority of the studies included used adequate diagnostic criteria for IGT, DM2 and MetS and all studies described the methods used to measure the outcomes. Eight studies did not report whether the same methods were applied to case and control to measure the outcomes of interest. Most studies did not state whether subjects had family history of IGT, DM2 and MetS. Four studies excluded controls with family history of DM, potentially introduce bias towards a more significant odds ratio. Comparability of subjects was evaluated based on age and BMI. Only 14 studies were found to strictly use age- and/or BMI-matched populations with an additional 24 reported non-significant difference in mean age and/or BMI between PCOS and control group.

Table 2.5: Summary of quality assessment of prevalence studies.

Study	Appropriate sampling	PCOS measured reliably and objectively	Response rate	Sample size	Crude number of cases	Age range	Ethnicity	Score
America								
Knochenhauer et al. (1998)	0	1	1	1	1	1	1	6
Azziz et al. (2004)	0	1	0	1	1	1	1	5
Goodarzi et al. (2005)	0	0	1	1	1	1	1	5
Lo et al. (2006a)	0	0	0	0	0	1	0	1
Okoroh et al. (2012)	0	0	0	1	0	1	0	2
Christensen et al. (2013)	0	0	0	1	1	0	0	2
Sirmans et al. (2014a)	0	0	0	1	1	1	1	4
Moran et al. (2010a)	0	1	1	1	1	1	1	6
Gabrielli and Aquino (2012)	1	1	1	1	1	1	1	7
Faria et al. (2013)	0	0	0	1	1	0	0	2
Europe								
Michelmore et al. (1999)	0	1	0	1	0	0	1	3
Diamanti-Kandarakis et al. (1999)	0	1	1	1	1	1	1	6
Asunción et al. (2000)	0	1	1	1	1	1	1	6
Sanchón et al. (2012)	0	1	1	1	1	1	0	5
Lindholm et al. (2008)	1	1	1	1	1	0	1	6
Lauritsen et al. (2014)	0	1	0	1	1	1	1	5
Asia								
Chen et al. (2008)	0	1	1	1	1	1	1	6

Study	Appropriate sampling	PCOS measured reliably and objectively	Response rate	Sample size	Crude number of cases	Age range	Ethnicity	Score
Ma et al. (2010)	1	1	1	1	1	1	1	7
Li et al. (2013a)	1	1	1	1	1	1	1	7
Jiao et al. (2013)	0	1	0	1	1	1	1	5
Zhuang et al. (2014)	1	1	1	1	1	0	1	6
Sung et al. (2010)	0	1	0	0	0	1	0	2
Nidhi et al. (2011)	0	1	1	1	1	0	1	5
Gill et al. (2012)	0	1	0	1	1	0	1	4
Joshi et al. (2014)	1	1	1	1	1	0	1	6
Kumarapeli et al. (2008)	1	1	1	1	1	1	1	7
Vutyavanich et al. (2007)	0	1	1	1	1	1	1	6
Middle East								
Musmar et al. (2013)	0	1	1	1	1	0	1	5
Hashemipour et al. (2004)	1	1	1	1	1	0	1	6
Mehrabian et al. (2011)	0	1	0	1	1	0	1	4
Asgharnia et al. (2011)	1	1	1	1	1	0	1	6
Tehrani et al. (2011)	1	1	1	1	1	1	1	7
Esmaeilzadeh et al. (2014)	1	1	1	1	1	0	1	6
Rashidi et al. (2014)	1	1	1	1	1	1	1	7
Yildiz et al. (2012)	0	1	1	1	1	1	1	6
Al Khaduri et al. (2014)	0	0	0	1	1	1	0	3
Attlee et al. (2014)	0	0	1	1	1	0	1	4

Study	Appropriate sampling	PCOS measured reliably and objectively	Response rate	Sample size	Crude number of cases	Age range	Ethnicity	Score
Australia								
Lowe et al. (2005)	0	0	1	1	1	0	0	3
March et al. (2010)	0	1	0	1	1	0	1	4
Boyle et al. (2012)	0	1	1	1	1	1	1	6
Joham et al. (2013)	1	0	0	1	1	0	1	4

2.5 Discussion

2.5.1 Prevalence studies identified

For prevalence studies identified from the literature, it is suggested that for the same study, the prevalence under the NIH criterion is always the lowest, followed by that estimated using the AES criterion and then the Rotterdam criterion. This verifies the inclusive and exclusive relationship between the 3 major criteria in Table 1.1.

In term of sample size, there are 4 studies (Lo et al., 2006a; Okoroh et al., 2012; Christensen et al., 2013; Sirmans et al., 2014a) in the US using fairly large sample sizes based on electronic medical records. However, all these studies presented a much lower prevalence of PCOS compared to community-based studies, underlying the fact that PCOS is a syndrome without much awareness by both the public and the healthcare workers. As a consequence, women with PCOS often do not seek for care and the diagnosis of PCOS is likely to be under-reported by clinicians.

However, there are some concerns about the quality of the studies included which may limit the interpretation of results from this qualitative analysis. As there are few instruments specifically designed (i.e. the Newcastle-Ottawa scale for case-control and cohort studies) for prevalence studies, the formal evaluation of included studies was challenged. It is suggested by Munn et al. (2014) that the followings are all essential factors to be considered for prevalence studies: sampling scheme, sample representativeness, recruitment strategy, sample size calculation, description of study subjects and settings, response rate, standard criteria used for measurement of a specific condition, reliable measurement instrument, appropriate statistical analysis. However, few studies met all of the above criteria. For example, there were differences in the outcomes due to hormonal assays and measurement techniques (Knochenhauer et al., 1998; Asunción et al., 2000). The understanding and application of the diagnostic criteria of PCOS in distinct studies also varied as some studies reported using less stringent definitions (Asunción et al., 2000). Selection bias was also observed because not all studies used an entirely randomised sampling scheme (Diamanti-Kandarakis et al., 1999). Other concerns with respect to study design included low response rate (Azziz et al., 2004), self-rated questionnaire (Gabrielli and Aquino, 2012), biases against the diagnosis of PCOS (e.g. a negative evaluation is easier to be made compared with a positive one since a confirmed diagnosis of PCOS requires a full blood test whereas there were potential cases with suggestive symptoms who did not complete

the full evaluation) (Azziz et al., 2004). The above factors all have influence on the accuracy of the prevalence estimates in individual studies.

2.5.2 Studies for PCOS-related morbidities

The results from the reviews suggested that the prevalence of common metabolic disorders and cardiovascular diseases is generally higher in women with PCOS compared with the general population. PCOS patients are also at elevated risk of pregnancy complications and their babies tend to have higher chance of suffering from neonatal complications. In addition, women with PCOS are more likely to develop endometrial cancer due to irregular menstruation. The consequences of PCOS exert significantly negative psychological impact on patients, which further lead to an overall reduced quality of life. Methodological issues raised by these reviews include that confoundings were not able to be controlled for when examining the association between PCOS-related disorders and PCOS. For example, these confoundings may include age, concomitant treatments, diet history, socio-economc status (Anderson et al., 2014; Lim et al., 2012).

The results from the updated search for metabolic disorders associated with PCOS also largely supported the main conclusions raised in the previous review on this aspect. However, methodological quality is a concern, which limit the interpretation of results from studies included in the systematic review. For example, the majority of the studies used non-randomised sampling scheme and therefore, were found to be at high risk of bias. In view of study design, most studies used cross-sectional design, with PCOS and metabolic risk factor status being evaluated concurrently at a single time point, which raised the issue of 'prevalence bias'. For instance, PCOS patients who suffer from the disease for a longer time are more likely to develop metabolic disturbances whereas the sample is taken at a single time point. Consequently, the prevalence odds ratios tend to overestimate the true risk factor because for most studies, PCOS patients with different durations of disease history are equally weighted, with the former being over-represented. There is also concern about the optimal comparison group as the control population in many studies comprised of population controls or pre-determined clinically examined individuals without PCOS or even subject with partial PCOS features. Evidence has been found to support an unselected age-matched control population rather than predefined non-PCOS population to be the optimal comparison group such that this prevalence bias can

be minimised (Bloom et al., 2006). The previous review by Moran et al. (2010b) suggested a prospective (or retrospective) cohort study design as it allows the timing and directionality of event (incidence) to be clearly established.

Despite the challenges mainly related to study design, I was able to explore the associations of PCOS with a range of diseases and found that PCOS is likely to worsen many health aspects of women, urging the need for early screening and monitoring as well as an integrated health plan.

In the next chapter, I will introduce the modelling approach and results obtained from the modelling for the prevalence of PCOS in different ethnic groups and the prevalence of morbidities associated with PCOS. These are largely based on modelling the data identified in the systematic literature review in this chapter.

Chapter 3

ESTIMATING ETHNICITY-SPECIFIC PREVALENCE OF PCOS AND THE RISK OF MORBIDITIES ASSOCIATED WITH PCOS

3.1 Chapter summary

In this chapter, I firstly introduce background information of evidence synthesis and meta-analysis given that these are the key concepts leading through the whole chapter. I then describe the methods of prevalence estimation for both PCOS and its associated morbidities. After that, I discuss the generalised models for meta-analysis using a Bayesian approach, which was motivated by the need to integrate data from multiple sources for comparing the relative risk of common metabolic disorders associated with PCOS. I then present results from evidence synthesis and a package created for R named `bmeta`. The package is largely based on the generalised meta-analytic models and various functions of this package are illustrated through examples, with sample data from the published studies for PCOS. The web application of the package (`bmetaweb`) is also presented, which was created to facilitate users without knowledge

of `R`. It should be noted that a paper derived from prevalence estimation of PCOS among different ethnic groups in this chapter has been published in `Oncotarget` (Ding et al., 2017).

3.2 Background

It is often the case that important medical questions are studied more than once by different research teams at different locations and the outcomes from small studies can be diverse and conflicting, which may result in difficulties in making clinical decisions (Haidich, 2010). For example, as described in Figure 3.1 (a fictional example is used here), suppose that four studies conducted in different countries have been found to provide odds ratio (OR) comparing the relative risk of developing type 2 diabetes in PCOS and general population. In such case, combining available information from multiple sources to generate an integrated result may provide more indications for decision making.
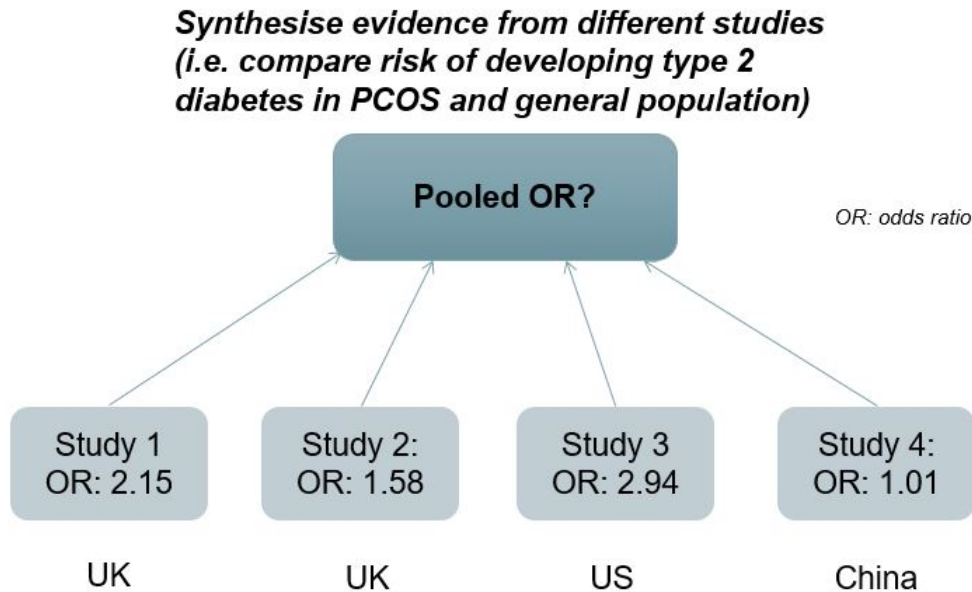
Figure 3.1: A fictional example: synthesise evidence from individual studies conducted in different countries to compare the relative risk of diabetes in PCOS and general population. Here four studies conducted in three countries all provided odds ratio and it is of our interest to estimate the 'true' effect of PCOS on the risk of diabetes.

Meta-analysis is a commonly used statistical approach to achieve this goal by integrating results from independent studies and is considered to play an essential role in evidence-based

medicine (Sackett et al., 1996). While meta-analysis based on Frequentist methods has been widely applied in medical research, the application of a Bayesian approach can be beneficial in the context of meta-analysis for several reasons (Sutton and Abrams, 2001a; Spiegelhalter et al., 2004; Gelman et al., 2014). The main advantage of a Bayesian approach is that the observed data can be complemented by a formal representation of prior belief to produce posterior probability distributions for parameters of interest, especially when there is little information provided by the observed data. Incidentally, this is exactly the case for studies that estimated the prevalence of PCOS in different ethnic groups as only a limited number of high quality studies has been identified through the literature search. For example, the prevalence of many chronic diseases is not expected to go beyond 50% in the general population and prior distributions based on this information can be included in the models. This helps stabilise the estimates, e.g. if we know the prevalence of a disease is expected to be around 5% and is unlikely to go beyond 10%, a small variance associated with the parameter representing the prevalence of the disease can be defined. This reduces the variances of the estimates we finally obtain, thus increasing the precision of the estimates. Moreover, compared to the traditional Frequentist methods where the parameters are often assumed to be fixed values (i.e. maximum likelihood estimation), the Bayesian approach allows to take fuller account of the uncertainties related to models and parameter values by producing a distribution of parameters (also called the posterior distribution) to sample from. This is a crucial aspect in decision making as in such way, the uncertainties can be propagated, e.g. in models that are used for health economic evaluation. These advantages motivate me to use the Bayesian approach in prevalence estimation for both PCOS and its associated morbidities as well as to develop a package including a set of Bayesian meta-analytic models.

3.3 Aim

In Chapter 2, 41 prevalence studies have been identified, with a few of high quality and potentially eligible for modelling. Therefore, the first objective was to establish the prevalence of PCOS among different ethnic populations. The second objective was to establish the prevalence of PCOS-related disorders in PCOS population. Along with this, I also aimed at creating some basic templates of meta-analytic models using Bayesian approach. This is because that although the idea of performing meta-analysis within the Bayesian framework has been widely accepted and recommended (Sutton and Abrams, 2001b; Higgins et al., 2009) due to obvious advantages

as discussed previously, there is no package identified that designed for Bayesian meta-analysis (e.g. a set of models with vague prior which can be modified later on according to users' need) and its relevant post-hoc analysis. In other words, by then, most applied implementation of meta-analysis are conducted under the Frequentist paradigm. For example, standard packages such as 'RevMan' developed by the Cochrane Centre (Collaboration et al., 2011), 'meta' or 'metareg' in STATA (Sterne et al., 2008) and the 3 existing packages 'metafor','rmeta' and 'meta' in R (Viechtbauer et al., 2010; Lumley, 2009; Schwarzer, 2007) all compute maximum likelihood estimates. Therefore, it would provide much more convenience if such a package that can perform Bayesian meta-analysis is available.

3.4 Methods

3.4.1 Estimating prevalence of PCOS

Modelling approach

This section presents the Bayesian hierarchical model used for the prevalence estimation. The underlying reason for using a Bayesian approach is discussed previously in the background of this chapter. The software used to perform the analysis is Just Another Gibbs Sampler (JAGS, Plummer et al., 2003). It is a software that can perform simulations from Bayesian hierarchical models using Markov Chain Monte Carlo (MCMC) and the concepts related to MCMC will be discussed later. The modelling approach is described through an explicit example of the White population and the same method was applied to establish the prevalence of PCOS for all the ethnic groups.

The model includes several modules. Suppose that there are I studies in total for a certain ethnic group (e.g. Caucasian) and for each study $i = 1, 2, \ldots, I$, the number of PCOS cases out of the total female population was observed. A Binomial distribution was used to model data in these studies:

$$y_i \quad \sim \quad \text{Binomial}(\pi_i, m_i),$$

where y_i is the total number of women observed to develop PCOS in the ith study; π_i represents the probability of developing PCOS for the population in study i and m_i is the total number of

women in the study.

A logistic link (defined as $\log(\frac{p}{1-p})$, where p is the probability of a binary outcome, i.e. success or failure) was used to model the probability of developing PCOS in this case. The most obvious reason for applying a logit link is that it rescales a probability with values varying between $-\infty$ and ∞.

The model is specified as

$$\text{logit}(\pi_i) = \beta_i \sim \text{Normal}(\mu, \sigma^2),$$

where β_i is the study-level prevalence of PCOS on the logit scale and was further assumed to follow a Normal distribution with a mean μ and variance σ^2. The parameter μ represents the pooled mean prevalence of PCOS (on the logit scale) for all the studies using population of the same ethnicity. This can be transformed back to the natural scale and the overall prevalence p was calculated as follows:

$$p = \frac{\exp(\mu)}{1 + \exp(\mu)}.$$

This module was completed by including some prior distributions for μ and σ and several versions of priors were attempted. For example, the simplest model specification assumes:

$$
\begin{aligned}
\mu &\sim \text{Normal}(\hat{\mu}, \sigma^2) \\
\sigma &\sim \text{Uniform}(0, 2).
\end{aligned}
$$

It is often the case that "off-the-shelf" vague priors are used for the parameters in the model. This would mean that the probability of developing PCOS is 50% and one does not favour any value between 0 and 1 (which can be represented by $\hat{\mu} = \text{logit}(p) = \log(\frac{0.5}{1-0.5}) = 0$ and a large value such as 10, on the logit scale, for σ). However, this was considered to be inappropriate because we do not expect such high prevalence of PCOS in the general population (e.g. >20%) associated with high uncertainties. Therefore, some common-sense information was used on what the likely range of probability of the condition may be and the prior for μ was informed accordingly. For example, values ranging from 2.5% to 15% for $\hat{\mu}$ were considered to be reasonable.

It should be noted that for most ethnic groups, fewer than 5 studies met the inclusion criteria for modelling (e.g. pooling is inappropriate for studies that used electronic databases and those that sampled study population from the community). Given that the amount of information present in the data is limited, the estimation for the parameter σ is likely to be associated with high variability. In particular, if a very vague prior specification, i.e. Uniform(0,k), for some relatively large value k was selected, it is possible that the data would not be able to update this in a consistent way, resulting in a highly unstable posterior distribution for statistical inference. To mitigate this problem, it is possible to include more information in the prior, for example, here, a half-Cauchy distribution was specified as prior for σ. This distribution can be expressed as: $\sigma \sim$ half-Cauchy(B), where B is the population median residual standard deviation (also called the scale parameter). The half-Cauchy distribution belongs to the half-t family of distributions and is recommended when a more informative prior is needed (Gelman et al., 2006). Its probability density function is defined as:

$$f(x) \quad = \quad \frac{2}{\pi B} \frac{1}{1 + (x/B)^2}, x > 0$$

Note that the scale parameter B, can be modelled assuming a vague prior, e.g. a uniform distribution with lower and upper bound as 0 and 100. By assigning proper values for the scale parameter B, a weakly but reasonably informative prior can be included. For example, although having heavy tails, the half-Cauchy distribution is less likely to produce extreme values and distort inferences in the area of high likelihood. Therefore, including a half-Cauchy distribution for the variance is advantageous in terms of allowing for outliers and accommodating small variances close to zero.

One technical complication is that the half-Cauchy distribution is not one of the standard distribution in software such as JAGS. However, it can be proved that if

$$
\begin{aligned}
\varepsilon_\sigma &\sim \quad \text{Gamma}(0.5, 0.5) \\
Z_\sigma &\sim \quad \text{Normal}(0, \sigma_{Z_\sigma}^2) \\
\sigma_{Z_\sigma}^2 &= \quad \frac{1}{B_\sigma^2},
\end{aligned}
$$

where $B_\sigma \sim$ Uniform$(0, k)$ for some upper limit k. Then $\sigma = \frac{|Z_\sigma|}{\sqrt{\varepsilon_\sigma}}$ follows a half-Cauchy distri-
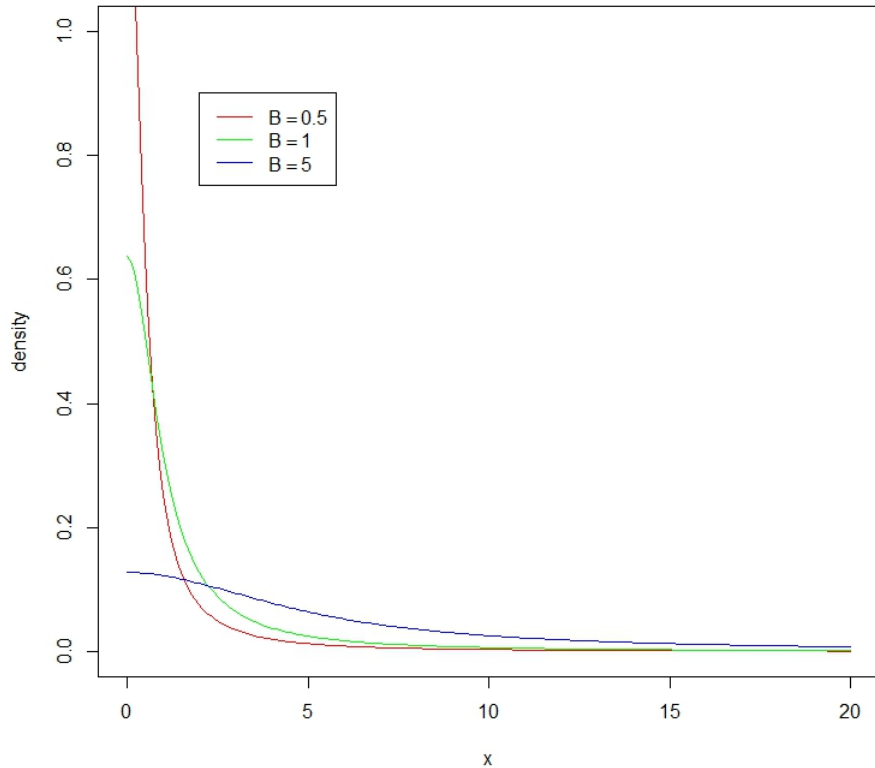
Figure 3.2: Graphical representation of the half-Cauchy distribution with different scale parameters.

bution with a median of B_σ (Lunn et al., 2012). Using this coding trick allows us to implement the half-Cauchy model in a relatively simple way.

In our case, we set $B \sim \text{Uniform}(0, 0.5)$ and Figure 3.2 shows different half-Cauchy distributions for varying values of the scale B. As is possible to see, using an upper limit of 0.5 still ensured that the prior for σ is concentrated closed to 0, while still allowing for large values (up to around 20 — recall that σ is actually defined in the log scale).

The computation approach for this analysis is Markov Chain Monte Carlo (MCMC), which is a widely used sampling method (Gamerman and Lopes, 2006; Gilks et al., 1996). A fundamental concept of MCMC methods is Markov chain. Consider a sequence of random variables $X_0, X_1, X_2, \ldots$ where we assume that the observation X_{t+1} only depends on the current one X_t

but does not depend on any observations before the previous one. This is equivalent to:

$$p(x_{t+1} \mid x_0, x_1, \ldots, x_t) = p(x_{t+1} \mid x_t).$$

Note that MCMC methods largely depend on the construction of a Markov chain that can converge to a target distribution p, from which one can simulate the distribution of interest, e.g. a posterior for a set of parameters. The distribution p is formally called the stationary distribution of the Markov chain. A stationary distribution can be described as that the probability distribution remains unchanged as time t progresses and is independent of the initial value of X_0, i.e. the distribution p no longer changes. In other words, the chain forgets where it starts in the end and enters into an "equilibrium" state where there is no more change in the distribution even we run the chain for a further number of iterations. In practice, under fairly general regularity conditions (MCMC works if the chains are irreducible and aperiodic such that convergence to the stationary distribution can be reached when the number of iterations becomes larger) (Brooks et al., 2011), the chain will converge after a sufficient large number of iterations (also called burn-in). Once convergence is reached, it is possible to compute any statistics such as mean and standard deviation from this distribution using random sampling (i.e. Monte Carlo).

One of the most commonly used algorithms for simulating Markov chains is the Gibbs sampling (Geman and Geman, 1984). The idea of this sampling method is to generate posterior samples by blocking one parameter each time to sample from its conditional distribution, with the remaining parameters fixed to their present values. More specifically, suppose the parameters of interest are defined as $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$ and the observed data are represented by y, then the sampling methods can be described as follows:

- Select arbitrary starting values for all the parameters of interest, e.g. $\theta_1^{(0)}, \theta_2^{(0)}, \ldots, \theta_k^{(0)}$;

- Sample a new value, $\theta_1^{(1)}$, for θ_1 from the conditional distribution $p(\theta_1 \mid \theta_2^{(0)}, \theta_3^{(0)}, \ldots, \theta_k^{(0)}, y)$;

  Sample a new value, $\theta_2^{(1)}$, for θ_2 from the conditional distribution $p(\theta_2 \mid \theta_1^{(1)}, \theta_3^{(0)}, \ldots, \theta_k^{(0)}, y)$;

  Note that since a new value for θ_1 has been sampled before, this most "recent" value is included as an element to be conditioned upon.

  This procedure continues until a new value of all the parameters has been sampled from the conditional distribution, e.g. sample $\theta_k^{(1)}$ from $p(\theta_k \mid \theta_1^{(1)}, \theta_2^{(1)}, \ldots, \theta_{k-1}^{(1)}, y)$

- Repeat the previous step a sufficient large number of times until the chain converges and then generate a sample from this distribution p.

MCMC is particularly important in Bayesian inference because it is often difficult to perform analytic examination on the posterior distributions given its complexity. MCMC allows one to take random samples from a distribution even without any knowledge of its mathematical properties.

Given the paucity of data, the estimates obtained from the posterior distribution are expected to be largely driven by the priors. I attempted to address this by looking at the Deviance Information Criterion (DIC, Spiegelhalter et al., 2002) of each model. The DIC is a common measure for comparing Bayesian models and can be used to perform model selection. The definition of DIC is described in the following steps. The model deviance, a likelihood-based measure of the model fit, is defined as:

$$D(\theta) = -2 \log p(y \mid \theta),$$

where θ represents the parameters of interest.

It is often the case that models with a large number of parameters are more likely to be subjected to 'over-fitting'. In other words, these models may fit the observed data (also called training dataset) quite well, while perform much worse for other datasets of similar structure. However, since these model fit the training dataset quite well, they are associated with smaller values of $D(\theta)$ and therefore, it may be of interest to penalise the complexity of a set of different models. This can be quantified by the following function:

$$
\begin{aligned}
p_D &= E_{\theta|y}[D(\theta)] - D(E_{\theta|y}[\theta]) \\
&= \overline{D} - D(\overline{\theta}),
\end{aligned}
$$

where $\overline{D}$ is the posterior mean deviance and $D(\overline{\theta})$ represents deviance evaluated at the posterior mean of the parameters. p_D measures the complexity of the model structure and the strengths of the prior at the same time. The DIC is then defined as follows:

$$
\begin{aligned}
\text{DIC} &= D(\overline{\theta}) + 2p_D \\
&= \overline{D} + p_D.
\end{aligned}
$$

Models with lower values of DIC are considered to perform better. However, only differences in the DIC are of interest whereas its absolute value is not relevant. This is because that the computation of model deviance depends on the observed data. A set of alternative models can then be compared by DIC however it may be possible that none of the models in comparison are correct.

It should be noted that the use of DIC may be a concern when the models are only populated by limited amount of information. This is due to the fact that the posterior distribution of the parameters of interest can be approximated to a Normal distribution only when the sample size of the observed data is large enough.

Despite this limitation, the DIC is simple to be obtained in the MCMC procedures and it can also be used to produce results by model averaging. Model averaging is a formal approach to account for model uncertainties in the predictions from a series of competing models. From a Bayesian perspective, the main interest of model averaging would be laid on the posterior predictive distributions, which can be computed as a weighted average of the model specific posterior predictive distribution (Hoeting et al., 1999). Suppose there is a set of models $M = (M_1, M_2, \ldots, M_H)$, the weights attached to individual models can be defined using the following formula (Baio, 2012):

$$ w_h \quad = \quad \frac{\exp(-0.5\Delta\text{DIC}_h)}{\Sigma_{h=1}^{H} \exp(-0.5\Delta\text{DIC}_h)}, \tag{3.1} $$

where $\Delta\text{DIC}_h =\mid \min_h(\text{DIC}_h) - \text{DIC}_h \mid$ and $h = 1, ..., H$, indicating the set of models.

For the current analysis, models with relatively small DIC were identified and the model averaging methods introduced above was applied to combine results from these models. This means that models with lower DIC were still favoured, however, instead of choosing a discrete cut-off (the lowest DIC) and discarding information from all the other models, our assessment is more continuous and effective by using the results from all the models in a principled way.

Data source and selection criteria for modelling

The source of data for prevalence estimation were from the literature review conducted in Chapter 2. Studies rated less than 5 in the quality assessment were deemed as low quality studies and therefore, were not included for further analysis (refer to Table 2.2). It should be noted

that I also excluded studies that did not apply the same diagnostic criterion as the majority of the studies reporting data for the same ethnicity (see Table 3.1). For example, Lauritsen et al. (2014) applied the Rotterdam criterion of PCOS whereas the other studies conducted in the Europe all followed the NIH criterion of PCOS. Another 5 studies were removed (Musmar et al., 2013; Hashemipour et al., 2004; Mehrabian et al., 2011; Asgharnia et al., 2011; Zhuang et al., 2014) due to that the age range of study population recruited was either narrower or wider than the reproductive age range (e.g. 17-18 years, 14-18 years, 12-44 years). Similarly, given that the age range of the study population are all different, statistical modelling was considered to be inappropriate for certain ethnic groups, i.e. South Asians (from India and Sri Lanka) and Australians. Likewise, modelling was not applied to ethnic populations such as Mexicans and Thai since only one study was identified to report relevant data. This is because if the result from a single study is unreliable, one would have no choice but to include it as the 'best' source of data, which largely bias the estimation.

For studies that reported data for more than one ethnic group (i.e. White and Black population), data for subgroups were used to populate the model. In other words, data from all eligible studies for a certain ethnic groups were included for modelling regardless of the country origin of the studies. For example, studies conducted in the US and Europe providing data for the Caucasians were synthesised together to generate estimates for the White population.

3.4.2   Modelling approach for comorbidities of PCOS

Meta-regression is an extension of standard meta-analysis in order to control for the moderator effect such as age or ethnicity (Borenstein et al., 2009). However, it should be noted that meta-regression accounts for the effect of study level covariates (e.g. mean age of the population in one study) rather than those at individual level.

The random-effects meta-regression for binary outcome was used to establish the prevalence of morbidities associated with PCOS, i.e. impaired glucose tolerance (IGT), type 2 diabetes (DM2) and metabolic syndrome (MetS). Meta-regression rather than meta-analysis was used because variation in the prevalence of IGT, DM2 and MetS in different geographical locations was expected and therefore needed to be controlled for. The meta-regression model was fitted using Bayesian methods such that the observed data were complemented by information from empirical studies reporting prevalence estimates of metabolic disorders in the general population

residing in different geographical locations. A random-effects model was selected as it is more sensible to assume that the 'true' study effect varies across studies in a same region.

The prevalence of DM2 in PCOS patients (using the NIH criterion) in comparison with controls was used as an illustrated example. The results for other combinations of metabolic morbidities and diagnostic criteria of PCOS were obtained using the same method. The model includes several modules.

Suppose that there are I eligible studies and each study $i = 1, 2, \ldots, I$ has 2 arms. The total number of events out of the total number of PCOS patients and controls was observed. Let y_{i0}, y_{i1} denote the total number of events observed in case and control arm, respectively, for study i. Similarly, n_{i0} and n_{i1} were defined as the total number of cases and controls. A Binomial distribution was used to model these studies:

$$y_{i0} \sim \text{Binomial}(\pi_{i0}, n_{i0})$$
$$y_{i1} \sim \text{Binomial}(\pi_{i1}, n_{i1}),$$

where π_{i0} and π_{i1} represent the probability of developing DM2 in the control arm and among women with PCOS, respectively.

A logistic regression was used to obtain the pooled mean probability of developing DM2 in women with PCOS and controls, respectively, in study i:

$$\text{logit}(\pi_{i0}) = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3}$$
$$\text{logit}(\pi_{i1}) = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \delta_i,$$

where α is the baseline odds for controls in studies conducted in the Europe and the prevalence of DM2 is reported to be 6.8% in this region; δ_i is the incremental risk of developing DM2 for women with PCOS, in study i. X_{i1}, X_{i2}, X_{i3} are indicators for region and were assumed to generate some impacts on the study effect. Studies have suggested that the prevalence of DM2 is the highest in North America (9.6%), then in a descending order by South East Asia (8.7%), South America (8.2%), with the lowest prevalence reported in the Europe (IDF, 2013). Since Europe was assumed to be the reference category, $\beta_1, \beta_2, \beta_3$ were used to represent the incremental prevalence of DM2 in North America, South East Asia and South America, respectively, compared with that in the Europe, e.g. β_1 was included to reflect the mean difference in prevalence of DM2

comparing North America and Europe.

The model was completed by including prior distributions for all the parameters. For example, the following distribution was included for α, which indicates that the prevalence of DM2 in the general population residing in the Europe is expected to be 6.8% with some small variability:

$$\alpha \quad \sim \quad \text{Normal}(-2.62, 0.5).$$

Note that $\text{logit}(0.068) = -2.62$ and in this way, evidence from pilot studies was converted to work as the prior distribution to inform our model. The variance of this prior distribution for the reference group is small because the prevalence of DM2 is expected to be within some reasonable range, e.g. 0-30% (Figure 3.3), with probabilities ranging from 0-15% more likely to be favoured when the simulation procedure proceeds.

The incremental prevalence of DM2 for the remaining regions in comparison with Europe is represented by the prior distribution of $\beta_1, \beta_2, \beta_3$:

$$\beta_1 \quad \sim \quad \text{Normal}(0.375, 0.01)$$
$$\beta_2 \quad \sim \quad \text{Normal}(0.266, 0.01)$$
$$\beta_3 \quad \sim \quad \text{Normal}(0.202, 0.01).$$

Here the expected prevalence (mean) estimates correspond exactly to the regional estimates mentioned previously, i.e. 9.6% in the North America, 8.7% in the South East Asia and 8.2% in the South America. The variance of these prior distributions is small because these are all population level estimates and therefore, high uncertainties are not expected. Note that when two independent Normal distributions are summed up, the resulting distribution is also Normal and the variance is equal to the sum of the variances of the two Normal distributions. Taking studies evaluating the prevalence of DM2 in North America as an example, the regional indicator β_1 is equal to 1 so the probability of the disease is in fact based on the sum of the distribution of α and β_1. Since again the prevalence of DM2 in the general population residing in North America is believed to lie within some reasonable range, the variance of the prior distribution included for β_1 is expected to be small. The resulting distribution after summing up the distribution of α and β_1 is shown in Figure 3.4. We can see that most of the bars lie within the range of 0 and 0.15, suggesting that the prevalence of DM2 is most likely to be in a range of 0-15% but
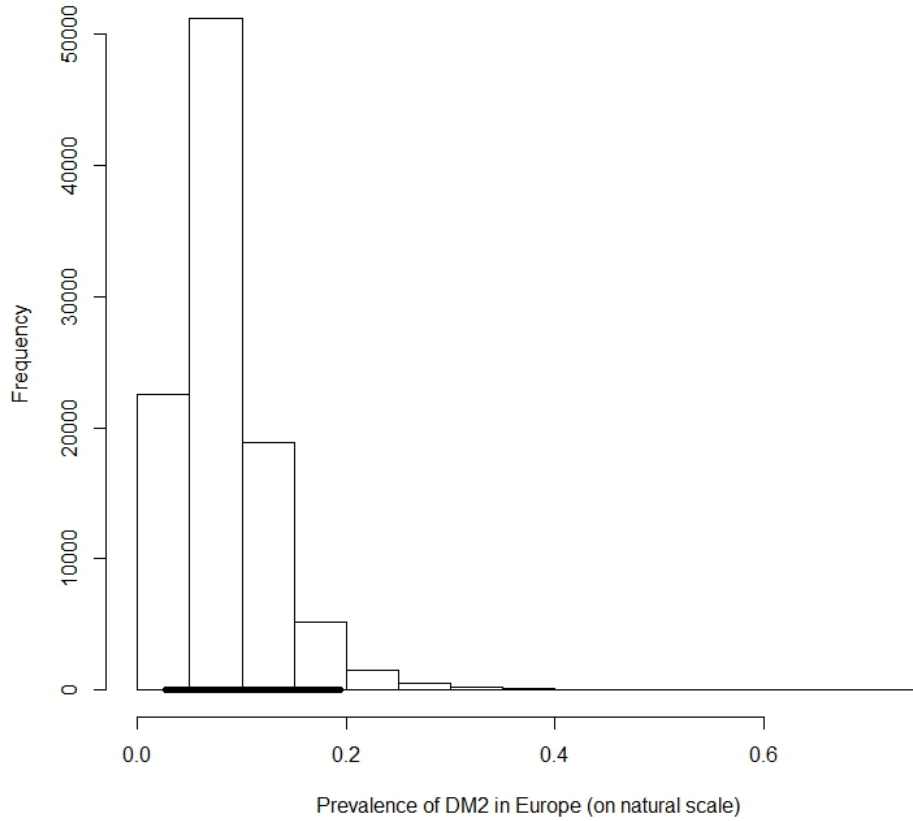
Figure 3.3: Graphical representation of the prior distribution included for the prevalence on a natural scale, which reflects the assumption made for the reference region (Europe) that the prevalence of DM2 is centred at 6.8%. The horizonal bar in bold represents the 95% prior interval.

unlikely to go beyond 25% for this given region. This figure reflects the assumptions made for the prevalence of DM2 in North America.

The study-specific odds ratio (on log scale), δ_i, was assumed to follow a Normal distribution with a mean of μ and a variance of σ^2 and μ is the pooled odds ratio for all studies.

$$\delta_i \quad \sim \quad \text{Normal}(\mu, \sigma^2)$$

I selected a prior for μ following a Student t-distribution. This is because the data included for this analysis are quite heterogeneous (e.g. some of the studies reported a high risk of diabetes in the PCOS patients compared with the controls in contrast with other studies) and the Student
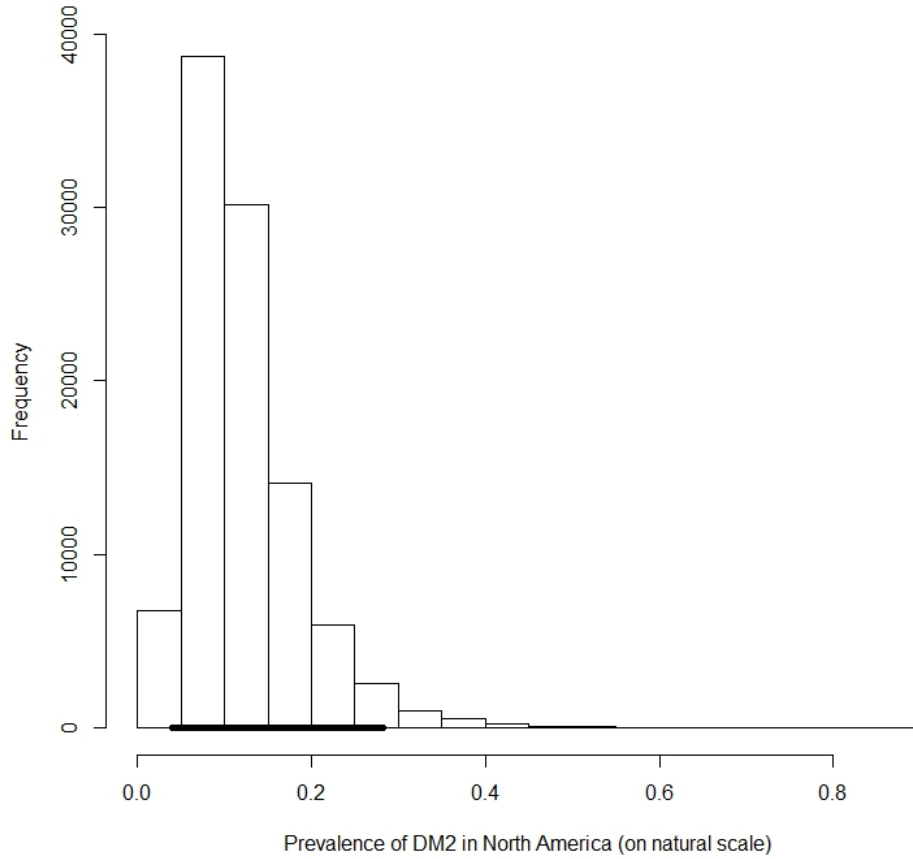
Figure 3.4: Graphical representation of the prevalence of DM2 in North America, which is the resulting distribution after summing up the distribution of α and β_1. This is consistent with the assumption made for North America that the prevalence of DM2 is centred at 9.6%. The horizonal bar in bold represents the 95% prior interval.

t-distribution has heavy tails which can accommodate outliers. The degree of freedom of the t-distribution is assumed to follow a uniform distribution within a range of 0 to 8. This is because a degree of freedom of 4 is suggested to limit shrinkage too much (i.e. can generate too extreme results). However using a Uniform(0,8) ensures on average a prior value for the degree of freedom of 4, while allowing for some variability around it (Berger, 1990).

$$\mu \quad \sim \quad t_v$$
$$v \quad \sim \quad \text{Uniform}(0, 8)$$

Note that JAGS uses a "generalised" t-distribution where centrality and precision (i.e. inverse

of the variance) need to be specified. In our case, a zero mean and a precision of 0.5 are included and the variance can then be computed using the formula $\frac{v}{\tau(v-2)}$ for $v \geq 0$ where τ is the precision (Lunn et al., 2012). This is more or less equivalent to a Student t-distribution with zero mean and a variance of 4, which is reasonably vague given that this is on the log scale. The standard deviation σ is assumed to follow a Uniform distribution with a lower and upper bound of zero and one, respectively (also considered reasonably large on the log scale).

$$\sigma \quad \sim \quad \text{Uniform}(0,1).$$

The graphical representation of the prior distribution for this parameter obtained from forward sampling (running the simulation without including the observed data) is illustrated in Figure 3.5. The majority of the bars lie within a range of 2 to 6, suggesting that the risk of diabetes comparing PCOS patients and the general population is most likely to be in this range. This assumption was validated with some clinical experts: the graphical illustration of the prior distributions was presented to clinicians, who considered these distributions sensible to work as prior information to complement the observed data.

The data from literature review were used to populate the meta-regression model and the results from quality assessment for all the studies considered are introduced in Chapter 2.
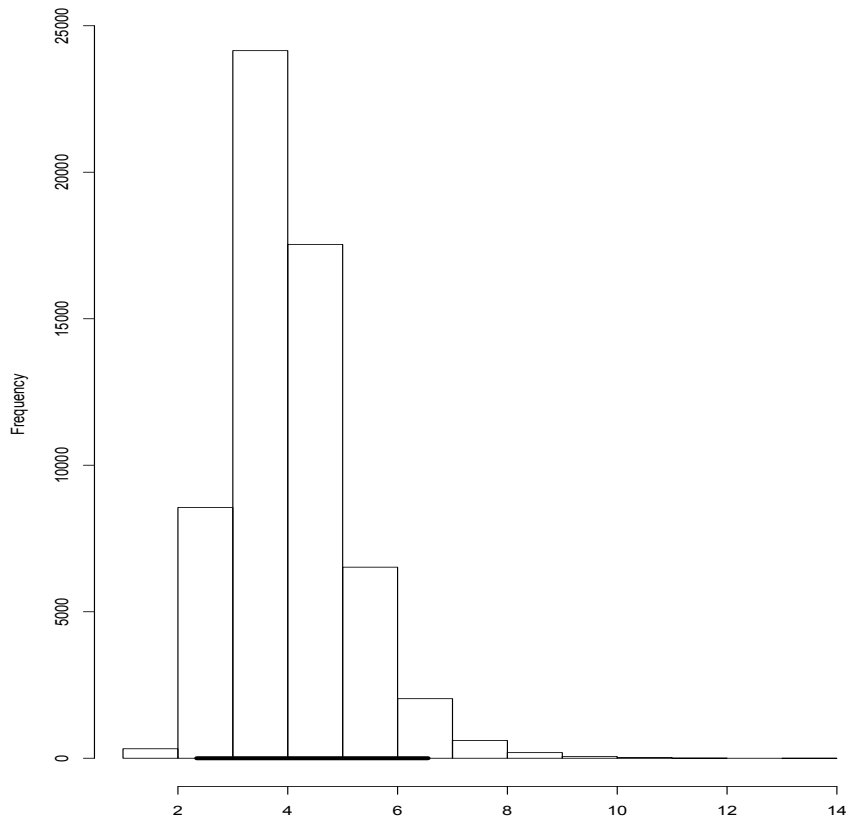
Figure 3.5: Graphical representation of the prior distribution included for the pooled odds ratio μ, which reflects the assumption that women with PCOS are generally 2-6 times more likely to develop DM2 compared with the general female population.

3.4.3 Generalised models for meta-analysis using Bayesian approach

In this section, I introduce generalised meta-analytic models using Bayesian approach, which forms the basis of all models included in the bmeta package for R (Ding and Baio, 2016). The modelling approach for the binary outcomes in the previous section was extended to fit other types of outcomes such as continuous and count outcomes.

The framework of models included can be summarised as follows. Suppose y_{ik} denotes the number of events or mean (i.e. continuous data) observed in arm k of study i. This is generally modelled as

$$y_{ik} \quad \sim \quad p(y_{ik} \mid \boldsymbol{\theta}_{ik}),$$

where the possible distributional assumptions are

$$p = \begin{cases} \text{Binomial,} & \text{binary outcome} \\ \text{Normal,} & \text{continuous outcome} \\ \text{Poisson,} & \text{count outcome} \end{cases}$$

and $\boldsymbol{\theta}_{ik} = (\phi_{ik}, \boldsymbol{\psi_{ik}})$ is a vector of parameters, specific to the chosen distribution (i.e. the mean and the variance, in the case of a Normal distribution). Here ϕ_{ik} represents the parameter of interest whereas $\boldsymbol{\psi_{ik}}$ is a vector representing a list of other parameters in the model.

Model		Parameter of interest (ϕ_{ik})		Nuisance parameters ($\boldsymbol{\psi_{ik}}$)
Binomial	π_{ik}	arm-specific probability of the event	n_{ik}	sample size of arm k
Normal	μ_{ik}	arm-level mean	τ_{ik}	arm-level precision
Poisson	λ_{ik}	arm-specific event rate	t_{ik}	person-time in the follow-up in arm k

A link function $g(\cdot)$ is then used to construct a linear predictor for the main parameter of interest, possibly as a function of observed covariates. For binary, continuous and count outcome, this is considered to be the logit, identity and log, respectively. In general terms, this is constructed as

$$g(\phi_{ik}) = \begin{cases} \alpha + \boldsymbol{X_i}\boldsymbol{\beta}, & \text{if } k\text{=}0 \text{ (control arm)} \\ \alpha + \boldsymbol{X_i}\boldsymbol{\beta} + \delta_i, & \text{if } k\text{=}1 \text{ (case arm)} \end{cases}$$

where α represents the 'baseline' information of parameter of interest for the control arm (e.g. probability of an event on logit scale for the control arm when the outcome is binary). δ_i is the study-specific impact of an exposure (for observational studies) or intervention effects (for randomised control trials). $\boldsymbol{X_i}$ is a matrix with covariates information (i.e. binary, continuous or categorical variables) for the $i-$th study and $\boldsymbol{\beta}$ is a vector of coefficients for regression parameters of interest. $\boldsymbol{X_i}$ and $\boldsymbol{\beta}$ only exist when a meta-regression is performed.

For fixed-effects models, it is assumed that all studies estimate the same intervention effect. While for the random-effects models, the assumption of exchangeability between studies is made and therefore, the 'true' intervention effect is modelled probabilistically:

$$\begin{aligned} \delta_i &= \mu & \text{Fixed-effects model} \\ \delta_i &\sim p(\delta_i \mid \mu) & \text{Random-effects models} \end{aligned}$$

where μ is the pooled estimate and p is a probability distribution of μ.

The model is completed by including minimally informative prior for the model parameters. For example, for binary outcome, a Normal prior is included, e.g. Normal(0,10000), for the pooled estimate μ. This indicates that the probability of developing a certain disease is 50% but is not likely to favour any particular value within the range of 0 and 1. This is a vague prior with large variability to represent the uncertainties so the results would be driven by the data.

These generalised models can be applied to synthesise evidence from multiple sources for the major data types in the medical research.

3.5 Results

3.5.1 Prevalence of PCOS in different ethnic populations

Using the modelling approach described in the Section 3.4.1, the following results were obtained (see Table 3.1).

For the NIH criterion, the prevalence of PCOS for White women was estimated at 5.5% and the 95% credible interval (CrI) is 4.8-6.3%. In other words, there is 95% chance that the mean prevalence of PCOS for the White women lie in the range between 4.8% and 6.3%. The corresponding figures for Black women and women residing in the Middle East are 7.4% (95% CrI: 6.3-8.7%) and 6.1% (95% CrI: 5.3-7.1%), respectively.

Using the Rotterdam criterion, prevalence estimation is only feasible for Chinese women (5.6%, 95% CrI: 4.4-7.3%) and women in the Middle East (16.0%, 95% CrI: 13.8-18.6%).

The prevalence for females in Middle East according to different criteria were estimated, which is 6.1% (95% CrI: 5.3-7.1%) under the NIH criterion, 16.0% (95% CrI: 13.8-18.6%) under the Rotterdam criterion, and 12.0% (95% CrI: 11.3-14.2%) under the AES criterion.

3.5.2 Prevalence of metabolic disorders in PCOS population

The odds ratios comparing the prevalence of impaired glucose tolerance (IGT), type 2 diabetes (DM2) and metabolic syndrome (MetS) in women with and without PCOS are presented in the Table 3.2. The prevalence of all the metabolic disorders examined is notably higher in women with PCOS given that all the 95% credible intervals exclude 1, regardless of the diagnostic criteria of PCOS applied. For example, under the NIH criterion, the risk of IGT is 3.86 times

Table 3.1: Estimated prevalence of PCOS using 3 major criteria for different ethnic groups. The results are presented as posterior mean with 95% credible interval (CrI) after weighted averaging results from a set of models based on Deviance Information Criteria (DIC).

Ethnicity	Estimated Prevalence (%) of PCOS in general population (95% CrI)		
	1990 NIH	2003 Rotterdam	2006 AES
White (Caucasian)	5.5 (4.8-6.3)	-	-
Black (African-American and Afro-Brazilian)	7.4 (6.3-8.7)	-	-
Chinese	-	5.6 (4.4-7.3)	-
Middle East (Iranian and Turkish)	6.1 (5.3-7.1)	16.0 (13.8-18.6)	12.6 (11.3-14.2)

higher for women with PCOS compared with the general population and there is 95% probability that the true odds ratio lie within the range of 2.28 and 6.47. It should be noted that for IGT and DM2, the odds ratio is larger for studies that used the Rotterdam criterion compared with the NIH criterion. While for MetS, the odds ratio is the largest for studies that applied the AES criterion, followed by the NIH criterion and then the Rotterdam criterion.

Table 3.2: The estimated odds ratio of IGT, DM2 and MetS comparing women with and without PCOS. The posterior means with 95% credible interval (CrI) are presented for each combination of diagnostic criteria and PCOS-related comorbidity.

	Prevalence of metabolic disorders in women with and without PCOS (OR with 95% CrI)		
	1990 NIH	2003 Rotterdam	2006 AES
IGT	3.86 (2.28, 6.47)	4.17 (2.42, 6.77)	-
DM2	3.95 (2.33, 6.40)	4.71 (1.85, 10.59)	-
MetS	3.14 (2.04, 4.69)	2.81 (1.85, 4.11)	3.89 (2.15, 6.54)

3.6   Discussion

3.6.1   Prevalence of PCOS in different ethnic groups

In terms of disease risk, Caucasian women in the US and Europe are less likely to develop PCOS compared with women residing in the Middle East. The Black women (the majority are African-Americans and Afro-Brazilians) are at the highest risks of PCOS. The 95% CrI of the

PCOS prevalence for White female and Black females do not overlap, indicating that White and Black females are at substantially different risk of PCOS. For Chinese women, the prevalence is merely 5.6% even when the Rotterdam criterion was applied. This figure is comparable with the prevalence for White females under the NIH criterion. Although estimates for Chinese women under the NIH criterion and the AES criterion was not modelled due to lack of data, it is expected that the prevalence using these two criteria would be even smaller, had data been available. The credible intervals of the estimates under the NIH criterion and the AES criterion for women in Middle East indicate that the prevalence of PCOS for women residing in this region differs according to these two criteria. It should be noted that the prevalence of PCOS is quite high for women in the Middle East when using the Rotterdam criterion, i.e. 16% (95% CrI: 13.8-18.6%). However, this finding is supported by a recent study where the prevalence was found to be 20% in this ethnic group based on the same criterion (Dargham et al., 2017). The prevalence of PCOS using different criteria for women in the Middle East supports the inclusive and exclusive relationship of the 3 major diagnostic criteria (i.e. the AES criterion is an extension of the NIH criterion and the Rotterdam criterion covers the broadest spectrum) as discussed in Chapter 1.

Generally speaking, under the same diagnostic criterion, Asian women are at a lowest risk of developing PCOS and then, in ascending order through Caucasians, Middle Eastern women, with the highest prevalence reported in Black women.

The results from the current analysis may be interpreted by the genetic ancestry data (Louwers et al., 2013). Evidence has suggested that the risk variants of PCOS in Korean women are not replicated in female Caucasians (Kim et al., 2014) and there is wide variation in the clinical presentations, hormonal and metabolic characteristics of PCOS across different ethnic populations (Wang and Alvero, 2013). As a consequence, the ethnic variation in terms of the diagnostic criteria, disease monitoring and management needs to be considered carefully.

Our results also indicated that even for the same ethnic group (e.g. women in the Middle East), there is large variation in the prevalence of PCOS when different diagnostic criteria were applied. This potentially pointed out the issue of under- or over-diagnosis of PCOS at present. Since the major concern for PCOS patients is the long-term metabolic risk, the clinical management of PCOS is often suggested to be at the earliest possible when a confirmed diagnosis is made because with proper interventions, the rapid conversion into complications such as type

2 diabetes may be controlled or even avoided (Conway et al., 2014). However, on the other hand, healthcare workers should be aware to avoid over-diagnosis of PCOS because it can negatively affect the mood of potential cases, which may consequently worsen some major symptoms of PCOS (i.e. menstrual dysfunction), thus increasing the chance for a potential case to be qualified as a true case. A recent study supported this argument by stating that the occurrence of severe dysmenorrhoea and irregular cycles can be induced by high stress (Kollipaka et al., 2013).

Admittedly, there are some limitations in this analysis, which is largely attributed to the paucity of data identified from the literature. Given that the published studies only provided limited prevalence data for different ethnic groups, the extrapolation of prevalence was largely driven by the prior distributions. The prior included may have some influence on the pooled estimates. However since experts' opinions were consulted, they were considered to be sensible. For example, the prevalence of PCOS was expected to be lie within a range of 2-20% and this information was represented by some suitable prior distributions (i.e. the prevalence is bounded within this range with certain degree of variability). Note that a prominent advantage of Bayesian methods is that the prior distributions can be updated by the observed data to generate posterior distributions of parameters of interest through the application of simulations. For example, in our case, the posterior distributions of prevalence in different ethnic groups were generated. Random sampling can then be proceeded by drawing a large sample size from the posterior distributions and this is expected to provide much more reliable results. Moreover, sensitivity analysis was conducted by attempting different versions of prior for our Bayesian model to identify the ones with relatively lower DIC and pooled estimates were obtained from model averaging based on Formula 3.1 in Section 3.4.1. Specifically, models with smaller DIC, indicating a better fit to the data, were weighted up while those with larger DIC were weighted down. In this way, information from models with a slightly higher DIC but provide reasonable estimates was not discarded. This tends to improve the accuracy of the estimation.

It is also suggested that the prevalence of PCOS varies in individuals with different weight categories (i.e. normal, overweight, obese) (Yildiz et al., 2008). However, the model used did not account for this factor due to lack of high quality body mass index (BMI) data from individual studies and the significant heterogeneity in the way of how BMI data were provided by different studies (e.g. some studies provided mean BMI of the entire study population whereas the other studies provided the proportion of patients who were overweight or obese).

73

There is also evidence that lifestyle modification ameliorates presentations of PCOS (Domecq et al., 2013). Consequently, the assumption that individuals of the same ethnicity have similar risk of developing PCOS may not be sensible in reality (as they may have quite different lifestyles) although we have accounted for individual level variability for prevalence estimates in our model.

3.6.2 Prevalence of morbidities associated with PCOS

In the current analysis, a larger number of studies were included compared with the previous review, allowing the studies to be categorised by diagnostic criteria and geographical location. The latter factor was controlled for in the meta-analytic model.

The results from our meta-analysis supported that significant greater prevalence of IGT, DM2 and MetS is observed in women with PCOS compared with controls regardless of which diagnostic standard of PCOS is applied: the odds ratios in Table 3.2 suggested that women with PCOS are generally 2-6 times more likely to develop various metabolic abnormalities than the general female population. Since the exact pathophysiology of PCOS remains largely unclear, the metabolic disorders associated with PCOS is often considered to be attributed to a combination of factors. For example, the elevated risk of IGT and DM2 in PCOS may be caused by peripheral insulin resistance, insulin hypersecretion and elevated β-cell function (Dunaif et al., 1989; O'Meara et al., 1993; Holte et al., 1995; Ciampelli et al., 1997; Vrbíková et al., 2002; VrbIkova et al., 2004). It should be noted that insulin resistance appears in 50-70% of women with PCOS (Legro et al., 2004), which augments androgen production and increases the level of free androgens through reducing sex hormone binding globulin (Teede et al., 2010a). The underlying hormonal imbalances form a vicious circle, which underpin PCOS.

Some studies have suggested that women with hyperandrogenic PCOS are more likely to present a worse cardiometabolic profile (Daan et al., 2014). The results from the current study tend to agree with this argument with respect to MetS as slightly higher ORs were observed under both the NIH and the AES criterion where hyperandrogenism is a prerequisite for diagnosis of PCOS. On the contrary, PCOS patients who met the Rotterdam criterion seem to be at a higher risk of IGT and DM2 compared with those who met the NIH criterion. However, note that there are uncertainties around estimates for IGT and DM2 (as suggested by the much wider 95% CrI compared with that of estimates for MetS) because relatively more studies for MetS were identified. Many studies investigating DM2 reported absence of DM2 in the control population,

which potentially generates strong impact on the pooled estimates. The absence of DM2 in the controls can be due to different reasons, e.g. selection bias, as discussed in the methodological issues of the studies included in Section 2.5.2. Further research may need to be warranted to investigate the metabolic features of PCOS population presenting different phenotypes.

Although the current model allowed to adjust for prevalence of IGT, DM2 and MetS in populations living in different geographical locations, a range of other confounding factors were not included. For example, it is well documented that obesity is a risk factor for glucose abnormalities and cardiovascular disorders, including metabolic syndrome and dyslipidaemia (Legro et al., 1999, 2001; Ehrmann et al., 2006). There is also evidence that even the body weight or BMI is controlled, central obesity may remain a contributing factor of increased cardiometabolic risk in PCOS population (Escobar-Morreale and San Millán, 2007; Godoy-Matos et al., 2009; Moran and Teede, 2009). Another example of confounding factor may be age as it is proposed that glucose abnormalities often worsen gradually with aging (Aguiree et al., 2013). While the confounding factors may bias the estimates, it was not possible to account for all of these variables due to lack of high quality data reported by empirical studies. In addition, the variation across studies in terms of ethnicity, medication use, recruitment source of sample population, family history of DM2 and definition of controls were also considered to substantially raise the statistical heterogeneity in the current analysis. Further observational studies regarding these aspects (e.g. matching by age, BMI and ethnicity) are in need to explore the causal relationship between common metabolic disorders and PCOS.

3.7   The `bmeta` package for `R`

Based on the modelling approach introduced in Section 3.4.3, I also developed a specific `R` package, named `bmeta`. The purpose of creating this package is to keep the users at least for a selected range of model specifications by coding up a large set of predefined models. In this way, all required for performing the Bayesian meta-analysis is a simple call to an `R` function and a decision in terms of what type of model to fit the data. After model specification, the package automatically calls `JAGS` and launches it in the background to conduct simulations. The results obtained from simulations (e.g. estimates from posterior distributions) are then stored back to `R` for post-hoc analysis. The package provides a relatively large class of models for conducting Bayesian meta-analysis in `R`, which is an open source software and is easy to interface with

Bayesian/MCMC software, e.g. `WinBUGS` (Spiegelhalter et al., 2003) or `JAGS` (Plummer et al., 2003), for simulations. The package includes functions for the calculation of various effect size or outcome measures (e.g. odds ratios, mean difference and incidence rate ratio) for different types of data and users are allowed to fit fixed- and random-effects models with different priors for the parameters. When effects of additional covariates are observed, meta-regression (Dias et al., 2013) can be performed to adjust for moderators. The sample data collected from the literature review for PCOS were used to present functions of this package.

3.7.1 Models included

There are 22 fixed-/random-effects meta-analysis/meta-regression included in the `bmeta` package (see Table 3.3 below for more details), which can be used to model different types of data (i.e. binary, continuous and count). Users can select fixed- or random-effects standard meta-analysis or meta-regression in conjunction with different types of prior distributions for the pooled mean outcome. In this way, users can compare the effect of different models (e.g. prior effect) fitted to their data and use the results from the optimal ones based on the diagnostic statistics.

The t-distribution prior was included for modelling the pooled estimates for binary outcome due to the fact that it has heavier tails and therefore was considered to be more suitable (compared with a Normal distribution prior) for dataset with outliers. The half-Cauchy distribution was included as prior for modelling the variance of the pooled estimate for count outcome. The underlying reason is that it has heavy tails and therefore can allow for outliers and accommodate small variances close to zero. The heavy tail of the half-Cauchy distribution can lead to less shrinkage to the population mean (in our case, the pooled estimates from all studies). However, this is considered to be appropriate for certain types of dataset where there are outlying data points. In terms of continuous outcome, some studies may report mean and standard deviation for both the case and control arm separately. Conversely, other studies may only report a mean difference between the two arms and pooled standard deviation. Thus, models that can fit data in different formats were created.

Table 3.3: Models provided by `bmeta`.

Outcome	Model	Model explanation	Type
Binary	`std.norm`	meta-analysis with normal prior	fixed/random
	`std.dt`	meta-analysis with t-distribution prior	fixed/random
	`reg.norm`	meta-regression with normal prior	fixed/random
	`reg.dt`	meta-regression with t-distribution prior	fixed/random
Continuous	`std.ta`	meta-analysis for studies reporting two arms separately	fixed/random
	`std.mv`	meta-analysis for studies reporting mean difference and variance only	fixed/random
	`reg.ta`	meta-regression for studies reporting two arms separately	fixed/random
	`reg.mv`	meta-regression for studies reporting mean difference and variance only	fixed/random
Count	`std`	meta-analysis	fixed
	`std.unif`	meta-analysis with uniform prior	random
	`std.hc`	meta-analysis with half-Cauchy prior	random
	`reg`	meta-regression	fixed
	`reg.unif`	meta-regression with uniform prior	random
	`reg.hc`	meta-regression with half-Cauchy prior	random

3.7.2 Implementation

The models introduced in the above section can be selected by users and `bmeta` calls `JAGS` to perform simulations with output stored back in `R`. However, before making model selection, the user firstly needs to create or load the data. For example, data stored in the `Excel` spreadsheet can be loaded using the following code:

```
> data=read.csv("IGT-NIH.csv")
```

Once the data are loaded into `R`, the user needs to format the data as required by `bmeta` using the `list` command. Moreover, the user needs to name the input variables with specific names (e.g. `y1`, `n1`, `y0`, `n0`) for the package to process and perform further analysis. The recommended format of the input data is shown below.

```
                Study Year y1  n1 y0  n0 Region X0 X1 X2
1             dos Reis 1995  7  29  0  19      1  1  0  1
2               Dunaif 2001  3  14  0  12      1  1  0  1
3               Yarali 2001  1  30  0  30      2  0  1  1
4               Faloia 2004  3  50  1  20      2  0  1  1
5                  Phy 2004  4   7  2  18      1  1  0  1
6  Diamanti-Kandarikis 2005  1  29  0  22      2  0  1  1
7       Sawathiparnich 2005  3   6  0   6      3  0  0  1
```

```
8        Alvarez-Blasco 2006  4  32  8  72      2  0  1  1
9              Echiburu 2008 17 159  1  93      1  1  0  1
10        Bhattacharya 2009 32 264 11 116      3  0  0  1
11               Kawai 2009 72 185 40 120      1  1  0  1
12                 Wei 2009 27 356 28 974      3  0  0  1
13               Huang 2010 11  90  0  40      3  0  0  1
14      Luque-Ramirez 2010 12 112  2  86      2  0  1  1
15                Pall 2010  1  75  0  23      2  0  1  1
16             Wiltgen 2010 22 195  0  25      1  1  0  1
17               Huang 2010 13 128  0  40      3  0  0  2
18               Amato 2011  7 125  2 144      2  0  1  1
19             Wickham 2012  4  13  0  13      1  1  0  1
20             Guleria 2014  5  50  1  50      3  0  0  1
```

The first two columns show the study characteristics including author and year of publication. Column 4-7 display the number of events and the total number of individuals in the case and control arm. Column 8 contains the regional indicator and this categorical variable was further converted into binary variables X_0, X_1, X_2, which can only take either 0 (no) or 1 (yes), as required by bmeta. Note that the data in Column 4-7 are the necessary information required by bmeta to perform a standard meta-analysis and if a meta-regression model is selected, then data from Column 9 and onwards are considered necessary depending on the number of covariates included for the model. The variables (e.g. Study, year, y1) do not necessarily have to be named as what are shown in the above table because the user will have to rename the input variables to feed bmeta using the list command mentioned previously:

```
> data.list<-list(y0=data$y0,y1=data$y1,n0=data$n0,n1=data$n1,X=cbind(data$X0,data$X1,
  data$X2))
```

After this, the user can proceed with model selection. Here, the random-effects meta-regression model with normal prior for binary outcome was selected and the object x included all the results.

```
> x<-bmeta(data=data.list,outcome="bin",model="reg.norm",type="ran")
```

The following output table is automatically generated by bmeta after the model specification. The second column presents the posterior mean of each model parameter and the associated standard deviation is presented in Column 3. Columns 4-5 are the lower and upper bound of 95% credible interval of the posterior mean. Notice that alpha[] represents the study-specific baseline odds whereas delta[] and gamma[] are the study-specific odds ratio (comparing the relative risk of events in case and control arm) on log scale and natural scale, respectively. The

pooled odds ratio `rho` is high (6.02) with a wide 95% credible interval of (2.60, 13.39). It is worth mentioning that `n.eff` (effective sample size) for certain parameters of interest is low, indicating potential autocorrelation. Note that under the context of MCMC, autocorrelation is a measure of the independence of different samples drawn from the posterior distribution and lower autocorrelation indicates more independent results and conversely, if there is high autocorrelation, the sample drawn does not accurately represent the posterior distribution.

```
Inference for Bugs model at "model.txt", fit using jags,
 2 chains, each with 10000 iterations (first 5000 discarded), n.thin = 5
 n.sims = 2000 iterations saved
          mu.vect  sd.vect    2.5%    97.5%  Rhat n.eff
alpha[1]   -3.718    1.261  -7.073  -2.094 1.093    55
alpha[2]   -3.623    0.994  -5.713  -1.902 1.003   550
alpha[3]   -6.062    1.484  -8.869  -3.489 1.063    31
alpha[4]   -3.974    0.991  -6.237  -2.314 1.070    42
......
alpha[18]  -4.422    0.594  -5.553  -3.297 1.013   120
alpha[19]  -3.530    1.070  -5.860  -1.850 1.025    90
alpha[20]  -3.986    0.712  -5.421  -2.667 1.073    27
delta[1]    2.394    1.265   0.721   5.413 1.020    88
delta[2]    2.050    0.946   0.449   4.145 1.001  1500
delta[3]    2.077    1.118   0.048   4.483 1.004   750
delta[4]    1.259    0.894  -0.287   3.200 1.059    34
......
delta[18]   1.542    0.642   0.318   2.841 1.010  1300
delta[19]   2.376    1.063   0.774   4.713 1.012   230
delta[20]   1.691    0.729   0.385   3.306 1.027    68
gamma[1]  178.820 1921.234   2.057 224.366 1.095    58
gamma[2]   13.654   27.510   1.567  63.125 1.001  1500
gamma[3]   15.484   24.030   1.049  88.506 1.004   750
gamma[4]    5.784   10.635   0.750  24.530 1.059    34
......
gamma[18]   5.786    4.355   1.375  17.140 1.010  1300
gamma[19]  24.141   75.662   2.169 111.374 1.012   230
gamma[20]   7.253    7.059   1.470  27.267 1.027    68
mu          1.695    0.433   0.956   2.595 1.003   520
rho         6.016    3.055   2.600  13.394 1.003   510
sigma       1.060    0.424   0.383   2.054 1.003   720
tau         1.587    2.298   0.237   6.805 1.003   720
deviance  141.753    8.141 127.402 159.204 1.014   110

For each parameter, n.eff is a crude measure of effective sample size,
and Rhat is the potential scale reduction factor (at convergence, Rhat=1).
```

```
DIC info (using the rule, pD = var(deviance)/2)
pD = 32.9 and DIC = 174.6
DIC is an estimate of expected predictive error (lower deviance is better).
```

Apart from the output table, the JAGS file for the required model configuration (random vs fixed effects; choice of outcome and prior distributions; presence of covariates) is saved in the working directory. This is considered as a sort of "template" and can then be modified to extend the modelling by changing priors in a way that is not automatically done by bmeta or saved for future reference. For example, if the model convergence is poor, one may consider changing the prior to be more informative as the prevalence of most diseases is likely to be restricted within some reasonable range. The output table above suggests that more informative priors based on pilot studies may need to be included in the analysis for this particular dataset.

3.7.3 Graphical functions

The bmeta package includes a range of graphical functions to produce not only standard plots for output display but also diagnostics to assess heterogeneity between studies, publication bias as well as model fit and convergence (Table 3.4).

Table 3.4: Plots produced by bmeta.

Plots	Function
posterior.plot	Posterior distribution plot of the summary estimate and between-study standard deviation
forest.plot	Graphical display of study-specific estimates and the pooled estimate. Need to install package 'forestplot' from R library to implement command
funnel.plot	Scatter plot to present publication bias
diag.plot	Diagnostic plot to present Gelman-Rubin statistic (Rhat) or effective sample size from MCMC simulations
traceplot.bmeta	Trace plot to examine the model convergence for each node
acf.plot	Autocorrelation plot to examine model convergence for each node

The function posterior.plot provides the posterior distribution of the summary estimate and the between-study standard deviation (when a random-effects model is selected). As all the models in bmeta employ the Bayesian approach, it is natural to look at the posterior distributions of the parameters of interest. Figure 3.6 shows the posterior plot for the summary estimate from the previous output table (Figure 3.6). We can see that all the bars lie to the right

of the reference line $x = 0$, suggesting that the prevalence of DM2 is notably higher in PCOS patients compared with the general population. The horizontal line in bold on the bottom of the bars represents the 95% credible interval.
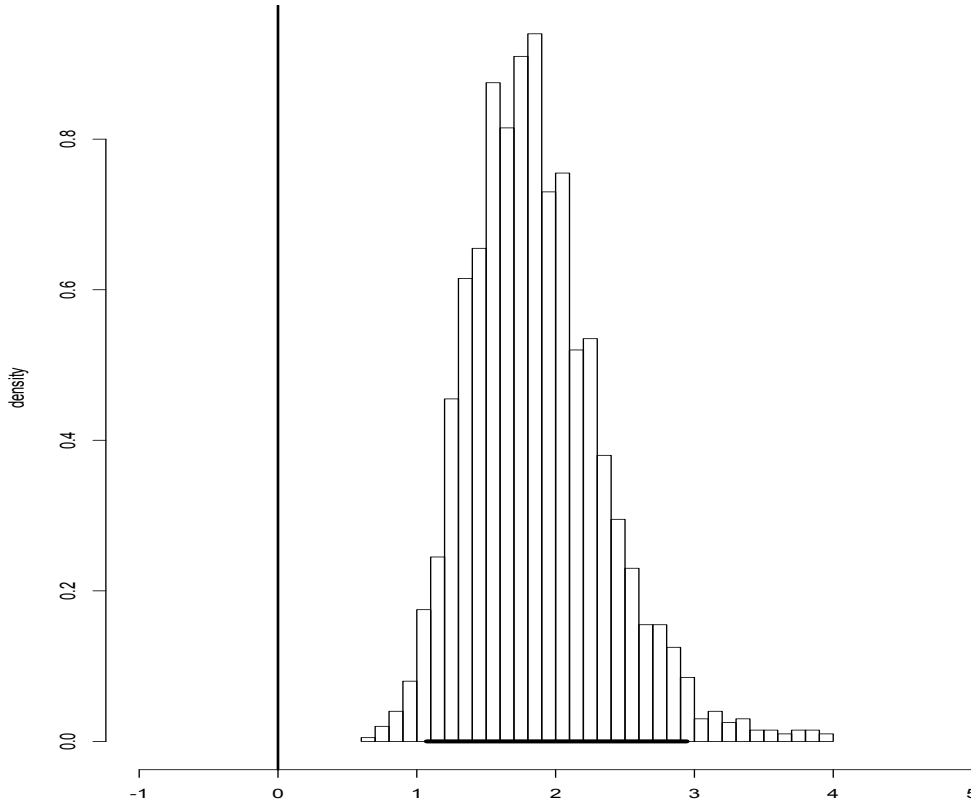
Figure 3.6: Posterior plot for the overall estimate obtained from pooling results from all studies included. The results are presented on a log scale and the reference line is therefore x=0. The horizonal line in bold represents the 95% credible interval of the estimate.

The forest plot is commonly used to graphically illustrate the outcomes obtained from meta-analysis. In a forest plot, a box or a circle is normally used to represent the study-specific estimates, which is displayed somewhere on a line that represents the 95% Credible Interval (CrI) of the estimate. The overall estimate obtained from pooling all studies is drawn like a diamond and a vertical reference line is used to show whether the outcome favours intervention or control. It should be noted that this function in bmeta is based on calling the R package 'forestplot' (note that bmeta automatically loads the forestplot package). A set of arguments is provided to edit the plot according to the user's preference. The user can change

the scale of the x-axis easily to make the estimates to be shown either on the log or natural scale. The color of the boxes and diamond can be changed and titles and study details including author and the year of publication can easily be added on the plot. For example, the following command defines a log scale and all the study characteristics are added, with the colour of the boxes and the diamond being specified as blue and orange, respectively.

```
> forest.plot(x,log=T,study.label=c(paste0(data$Study ,",", data$Year),"Summary estimate
    "), box="blue", summary="orange")
```

Figure 3.7 is a forest plot generated from the previous output table for the relative risk of IGT comparing the PCOS patients and controls. The boxes in blue represent study-specific odds ratio and the diamond in orange represents the summary estimate obtained by pooling all studies together. Note that the size of diamond is proportional to the uncertainties in the pooled estimate. There are 20 studies and for each of them, the study-specific log odds ratio are presented, with the pooled estimate at the bottom. We can see that all the boxes lie to the right of the reference line x=0, indicating significant increased prevalence of IGT in PCOS patients compared with controls.

Another important aspect when performing a meta-analysis is to examine the publication bias and Figure 3.8 presents a sample funnel plot produced by bmeta for the previous sample data of IGT. The function funnel.plot in bmeta displays study effect on log scale against the standard error of the study effect. It is suggested that when bias and between-study heterogeneity are completely absent, the scatter resembles a symmetrical funnel and the triangle area formed by connecting the centred summary estimate with its 2.5% and 97.5% quantiles on either side includes about 95% of the studies if the fixed-effects model assumption holds (i.e. all the studies estimate the same effect). Therefore, for the current plot, we can see that the scatter is not quite symmetrical with several dots lying outside the triangle, suggesting some extent of violation of the fixed-effects model assumption.

It is also worth commenting that bmeta allows to visualise diagnostics to assess model convergence, i.e. diagnostic plot, trace plot and autocorrelation plot, and these diagnostics are provided for all parameters in the output table.

Figure 3.9 presents the diagnostic plot for all nodes and we can see that the number of effective sample size for many parameters is quite low (much smaller than the actual number of simulated iterations as represented by the dash line at the top), suggesting issue of high
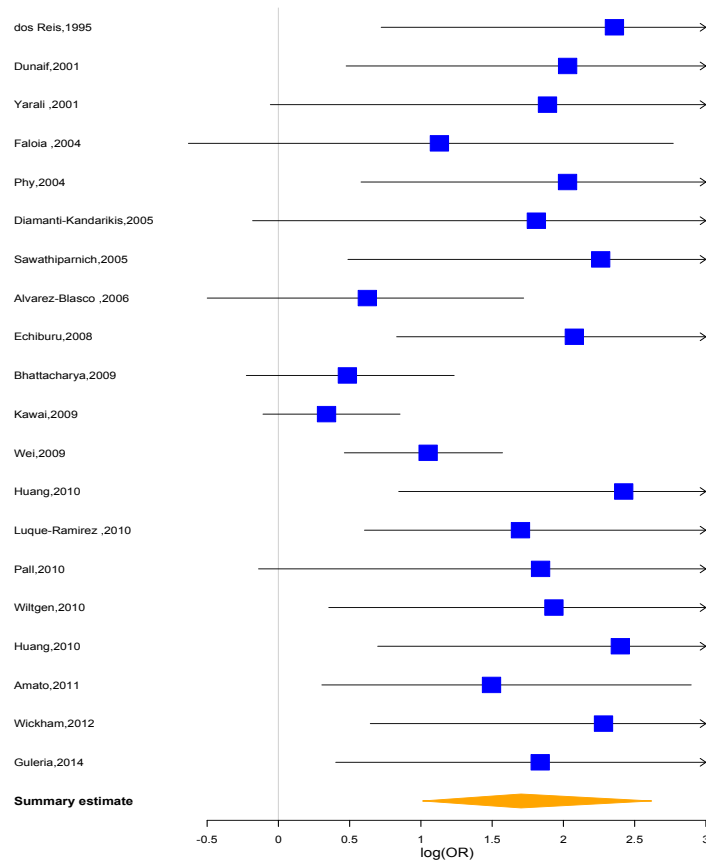
Figure 3.7: Forest plot for studies comparing relative risk of IGT in women with and without PCOS. The blue boxes represent the study-specific odds ratios and the lines are the corresponding 95% credible intervals. The diamond on the bottom represents the summary estimate obtained from pooling results from all the studies. The odds ratios are presented on a log scale so the reference line is x=0.

autocorrelation.

This is confirmed by Figure 3.10 where a certain node gamma[18] in the output table is presented. The autocorrelation plot often helps to check dependency among Markov chain samples. The reason behind is that the distribution of the current observation always depends on that of the previous one (a critical property of MCMC simulation), therefore the iterations of a Markov chain are observed to be correlated. However, it is expected that the kth lag autocorrelation (defined as the correlation between every draw and its kth lag) gets smaller as k increases (i.e. the 5th and 100th draws should be less correlated than 5th and 10th draws). Therefore, if autocorrelation continues to be high for larger values of k, we would suspect that there is a slow mixing of chains and high degree of correlation between draws. The plot presented
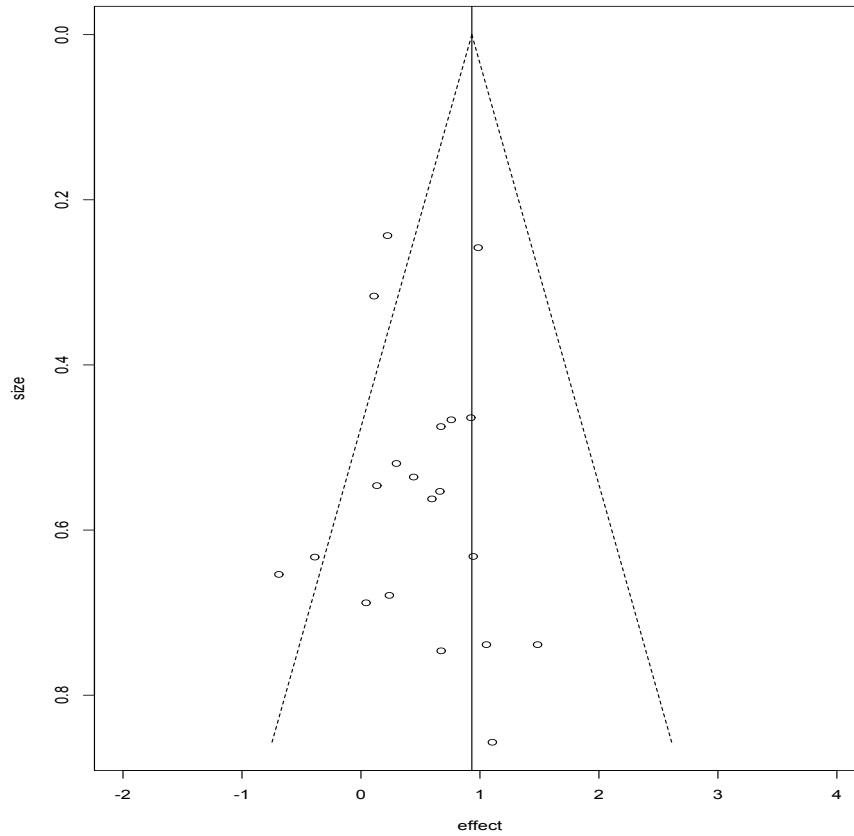
Figure 3.8: Funnel plot for studies comparing relative risk of IGT in women with and without PCOS.

here shows moderate autocorrelation as although the correlation (represented by each vertical line) between every draw of this node and its kth lag decreases when k increases, high correlation was still observed for larger value of k (e.g. $k=25$), with the vertical lines exceeding the border formed by the two dash lines.

Convergence diagnostics
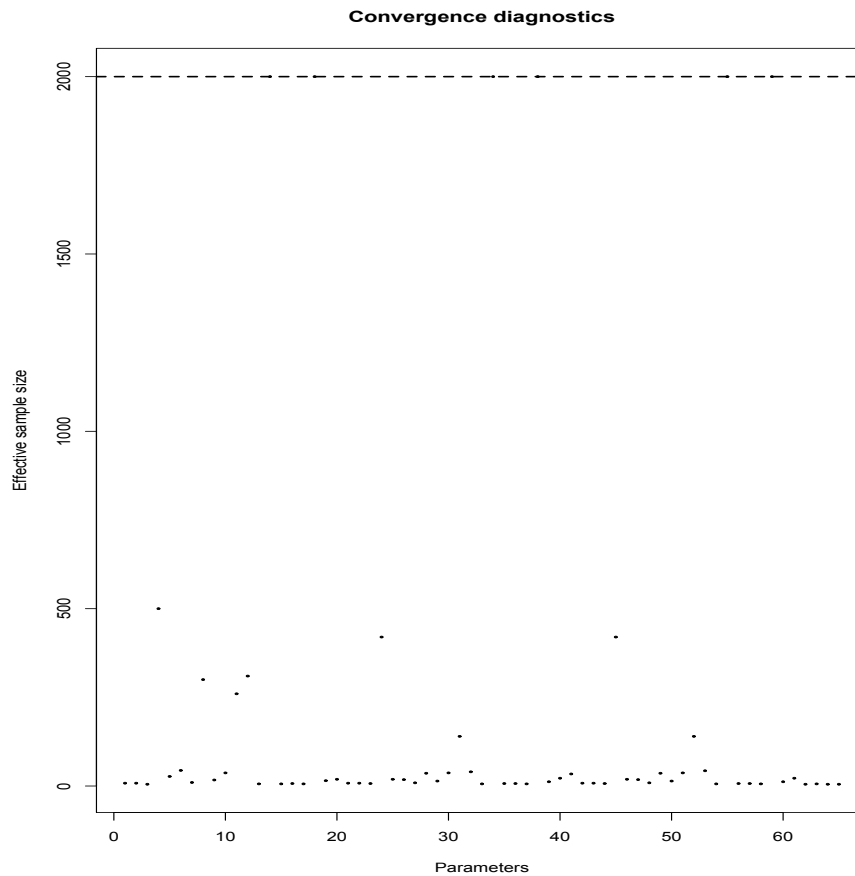
Figure 3.9: Diagnostic plot (display the number of effective sample size against the total number of saved simulations after burn-in) for main parameters in the model.

The traceplot of the node `gamma[18]` is provided (Figure 3.11). The two chains in the traceplot do not mix quite well and fluctuate around a certain horizontal line (represents a fixed value) even after a large number of burn-ins. This indicates that the model convergence has not been reached. It could be caused by the minimally informative prior in the model template provided by the package. Therefore, more informative prior distributions for this dataset may be considered or a larger number of iterations may need to be specified.
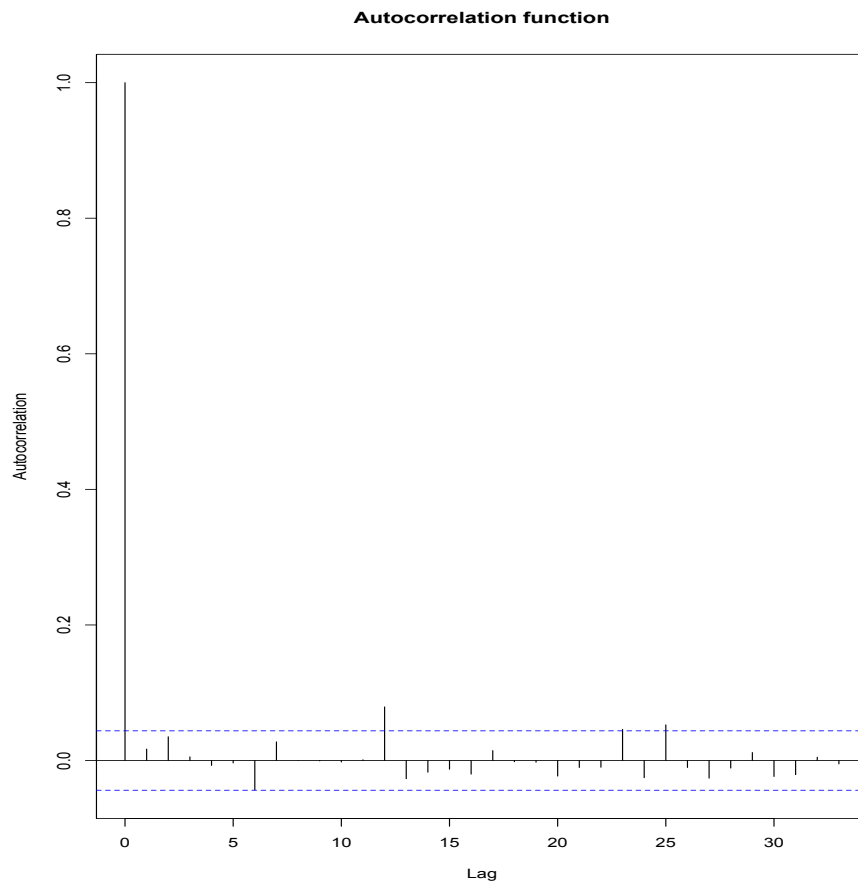
Figure 3.10: Autocorrelation plot produced by `bmeta`.

Note that the `bmeta` package has been released to CRAN, the offcial repository of R packages.
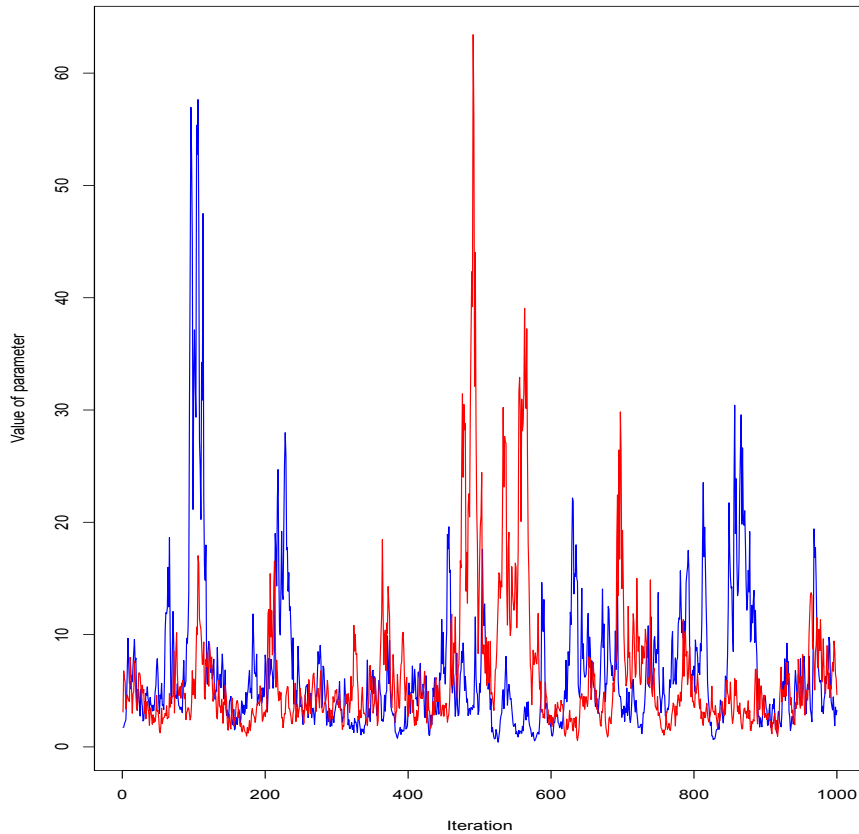
Figure 3.11: Traceplot produced by `bmeta`.

3.7.4 Web application of **bmeta**

A web application of `bmeta` has been created to facilitate users without direct knowledge of `R`. By using `bmetaweb`, users only need to upload data in the required format in `Excel` spreadsheet and specify the model to populate the data. `bmetaweb` runs the package in the background and the outcomes are presented in fairly straightforward way on the website page. A structured report in either `PDF` or `Word` can be downloaded with some pre-formatted texts aiming at guiding users through the interpretation of the results. The official website of `bmetaweb` is: `https://egon.stats.ucl.ac.uk/projects/bmetaweb/`.

Some screen shots of the `bmetaweb` are presented in Figure 3.7.4. The top graph in Figure 3.7.4 shows the welcome page of the website. This web page introduces the concepts of Bayesian meta-analysis and gives instructions of the data format required by the website. The bottom graph shows one of the function tabs where users can make model selection and upload data

87

for analysis. The right side of the window presents the output table obtained from the model specified by the user.
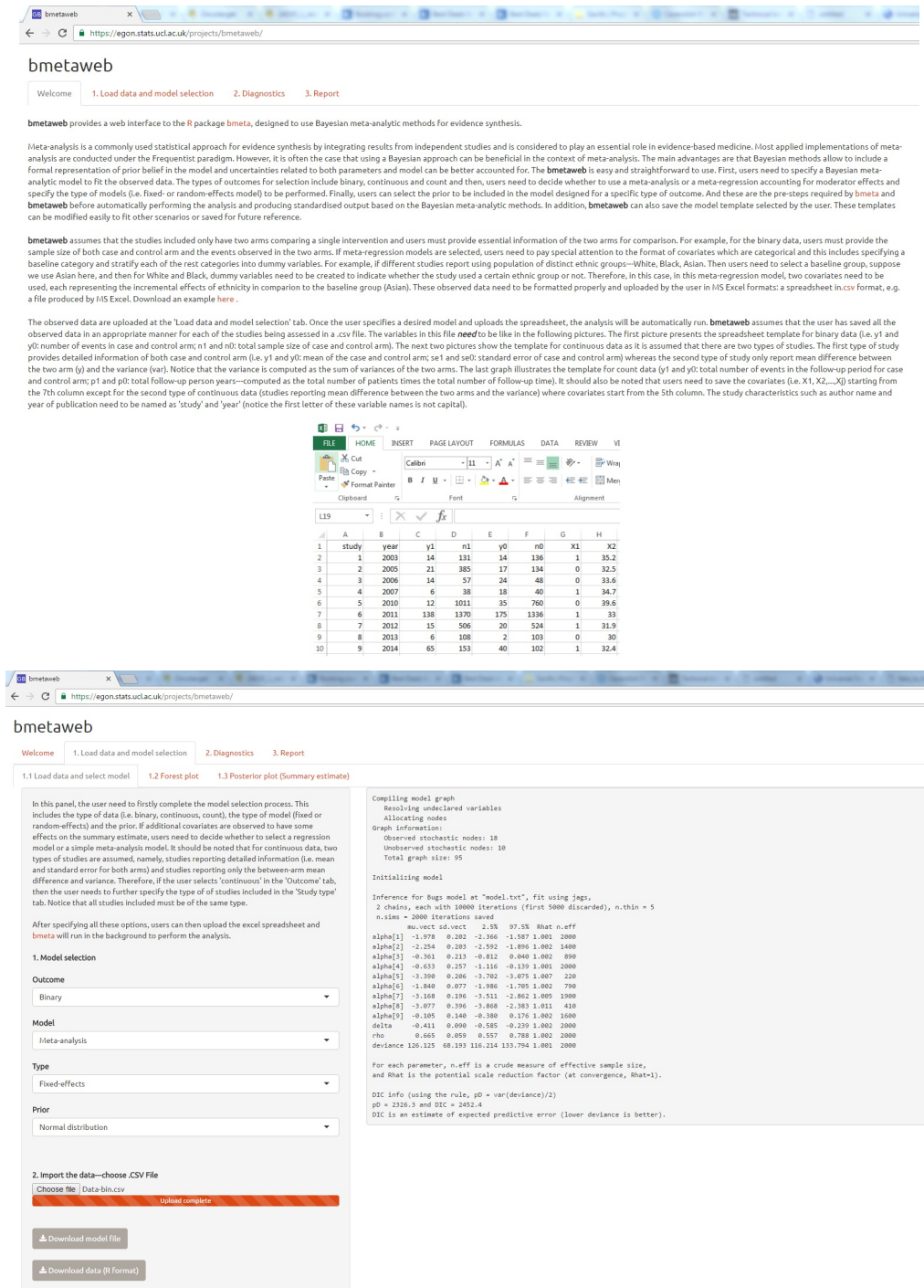
Figure 3.12: Screen shots of `bmetaweb`. The top plot is the welcome page which gives some general introduction about Bayesian meta-analysis and then instructs the user of the data format required by the website. The bottom plot shows one of the function tabs where users can make model selection and upload data for analysis. The output table is automatically generated on the right side of the window.

Chapter 4

EVALUATING THE PREVALENCE, INCIDENCE AND PRESCRIBING PATTERNS OF PCOS UNDER THE UK CONTEXT

4.1 Chapter summary

This chapter presents the results of incidence, prevalence and prescribing patterns of PCOS from a database study conducted in the UK context. The motivation of this database study is firstly discussed and then I give an overview of the UK primary care and the database used specifically for the current research. The methods for patient identification from the database and the analysis plan are introduced afterwards. The epidemiological and pharmaceutical results are presented and compared with studies conducted under similar settings. Finally, I describe the main conclusions from this database study and give some suggestions on the clinical management of PCOS from a primary care perspective. The main findings in this chapter have been converted into a paper published in the `BMJopen` (Ding et al., 2016).

4.2 Background

In Chapter 3, I present results from statistical modelling for prevalence estimation of PCOS in the general population. However, from the literature review, limited amount of information was identified with significant heterogeneity between studies. Since there are many uncertainties around the prevalence estimates in the previous chapter, I was motivated to find alternative methods to investigate the prevalence or other epidemiological outcomes of PCOS. Aside from this, the database study would allow me to explore the incidence, prevalence and prescribing patterns of the condition under the UK-specific context as well as to include more covariates (these may have influence on the epidemiological outcomes), which can hardly be achieved by looking at data provided by published studies.

4.2.1 UK primary care

In the UK, in general, most of the patients' experience is considered to be within primary care and the majority of the general practitioners (GPs) record patient information by computer. In the past, the main function of the information system in the general practices was to inform GPs or other clinical members to provide care on a day-to-day basis. The patient information was also used for registration and delivery of payments made to practices under a new GP contract.

The National Health Service (NHS) aims to provide care from cradle to grave through a universal coverage of services by GPs in the primary care and as early as in 1998, the Department of Health's strategy Information for Health (National Heatlh Service, 1998) attempted to establish lifelong electronic health records for all the patients. The data collected in routine clinical practices are population-based, extending from birth information to details of diagnoses recording as well as management and health outcomes. Therefore, by exploring primary care data, it is possible to grasp a better understanding of the disease history including access to healthcare by individual patients and the patterns of care (i.e. prescribing rates) in the general practices.

In 2006, the UK Department of Health has introduced a scheme termed the "Quality and Outcome Framework" (QOF), whose objective was to encourage the delivery of high quality primary medical services. Since the introduction of the QOF, GPs have been required to record a wider range of data electronically. For example, GPs are required to record clinical management following a diagnosis. Given that these data are also linked with payments to general practices, it is expected that the accuracy and quality of electronic health records can be improved gradually

in the primary care (Gnani and Majeed, 2006).

4.2.2 THIN database

History of the database

The primary care database used for this piece of analysis is *The Health Improvement Network* (THIN). The idea of developing a database was firstly brought up in mid-1970s by a GP, Dr. Alan Dean, who attempted to computerise the patient records and make these files paperless (Health Service Journal, 1999). He managed to build this program with the help from some IT staff and then began to sold it to other practices in 1979. By the end of the 1980s, an increasing proportion of GPs in the UK started to work in computerised practices and the need for large databases to conduct pharmaceuticoepidemiological research was raised (Hall, 1992).

In 1987, Dr. Alan Dean established a company — Value Added Information Medical Products (VAMP) and he aimed to include patient data from 950 practices. This first database was named the Vamp Research Databank (VRD) however the way that GPs used computers largely complicated its development. This is due to the fact that GPs generally needed to take three steps to convert the paper records to computerised records. For example, they firstly used the computer system to record prescriptions and then, to record details of consultations and finally, other relevant information such as secondary care information or medical history was recorded. Apparently, the analyses of the database were considered to be optimal by using data contributed from practices that completed all these three steps. As a consequence, free computers were offered to practices in order to incentivise more GPs to join this program.

In 1993, VAMP was sold to Reuters who later donated the VRD database to the Department of Health for the purpose of public health research. After this, the VRD database was renamed as the General Practice Research Database (GPRD, Ogdie et al., 2012) and regulated by the Office for National Statistics (ONS). The database was further moved to the Medicines and Healthcare products Regulatory Agency who is in charge of it nowadays. The name of the database was changed again to Clinical Practice Research Datalink (CPRD) in 2012.

In 1994, Dr. Dean set up a new company called EPIC to regain the license of GPRD and at the same time, VAMP was renamed as In Practice Systems Ltd (INPS) who bought back the software platform (named Vision) for data management from Reuters. In 2002, EPIC, INPS and Vision teamed up to establish The Health Improvement Network (THIN). It should be noted

that although the data collection process of THIN was started in 2003, the patient records can be retrieved back to 1980s for certain practices.

THIN at present

Nowadays, if a practice decides to join the THIN scheme, an initial process of full data collection is carried out. In order words, the practice needs to provide all the retrospective data available from the time point when the practice firstly began to use Vision or VAMP. The incremental data collected each year are then automatically downloaded and stored in an electronic form. This makes it less likely to incur disruptions to the daily activities of the practice while ensures the data security.

It is estimated that THIN collects data on 3 to 5 million of patients on an annual basis (Lewis et al., 2007). The latest data showed that the database included 559 practices covering over 11,350,933 patients (among which over 3.8 million are active patients who are eligible to be followed prospectively). It should be noted that the patient population accounted for 6.2% of the total population in the UK (IMSHealth, 2015). Evidence has suggested that in 2011, the number of patients on the NHS patient register record exceeded that of the population size in the UK estimated by the Office for National Statistics (Office for National Statistics, 2014). This could be due to the over and under coverage, e.g. patients registered in more than one area, delays in birth or death notification.

Structure of THIN database

The information of patients in the THIN database is recorded in a range of different files as shown in Table 4.1. There are seven files included in the database and some of the information may be recorded in two files, e.g. additional health data may record the reason for death related to a medical diagnosis indexed in the medical records. Due to the nature of the database structure, when extracting data from separate files, it is essential to consider all the files that may contain relevant information.

It should be noted that the symptoms and diagnoses are categorised using the so-called "Read code" system, which is a unique hierarchical recording system developed by Dr. James Read (Chisholm, 1990) and is widely used in the practices across UK. This classification codes disease and its associated history and symptoms, findings and signs of the examination, diagnostic

Table 4.1: Files that contain patient information in THIN database.

File name	Information recorded in the file
Patient records	Age, sex, the date of registration and leaving the practice
Medical records	Medical diagnosis and its associated date, the location of the event and free text option
Therapy records	Prescription (indication not recorded) with formulation and dosage information and the date issued, events leading to withdrawal
Additional Health Data (AHD)	Records of vaccinations, other information such as height, weight, smoking, birth, death, laboratory results
Postcode Variable Indicator (PVI)	The indices for postcode linked area-level socioeconomic, ethnicity and environmental information
Consultation records	Date, time and duration of consultation
Staff records	Information of staff who entered the data

and test procedures, drugs prescribed as well as some social information. All the Read codes are divided into chapters with a similar structure to the International Classification of Disease (ICD). For example, Read codes starting with 'C' indicate endocrine and metabolic diseases and "C11y200" is the Read code for impaired glucose tolerance. Over 100,000 Read codes are currently included and the same condition can often be described by a range of Read codes or the combination of different Read codes, e.g. 'Stein-Leventhal syndrome' is also used to define PCOS.

Case identification from the database

In general, the primary step for research based on databases is case identification. This involves extracting patient information from different files in the database to generate a big spreadsheet containing individual level patient data (i.e. each line corresponds to a unique patient). This spreadsheet can then be used to conduct more complicated analysis.

The data extraction process from THIN is graphically displayed in Figure 4.1 and can be described as follows:

- The first step is to identify a list of Read codes relevant for a specific disease by looping through the description fields of all the codes using 'key' words (Davé and Petersen, 2009). These 'key' words are either definition of a disease or the descriptions of symptoms related to a disease (see the rectangle with the terms 'Read codes' on the left side).

- The code list is then merged with the medical records of patients for case identification.

- After identifying all the cases, this file is merged with other files such as patient record (which contains the demographic information), therapy record (which contains the prescription data), additional health data (which contains other variables) and postcode variable indicator (PVI) records (which contains information on area-level social deprivation level) depending on the research question and type of analysis needed. The final study population file (see the bottom right rectangle) should contain all the disease and prescription indicators as well as covariates of interest to facilitate the further analysis of the dataset.
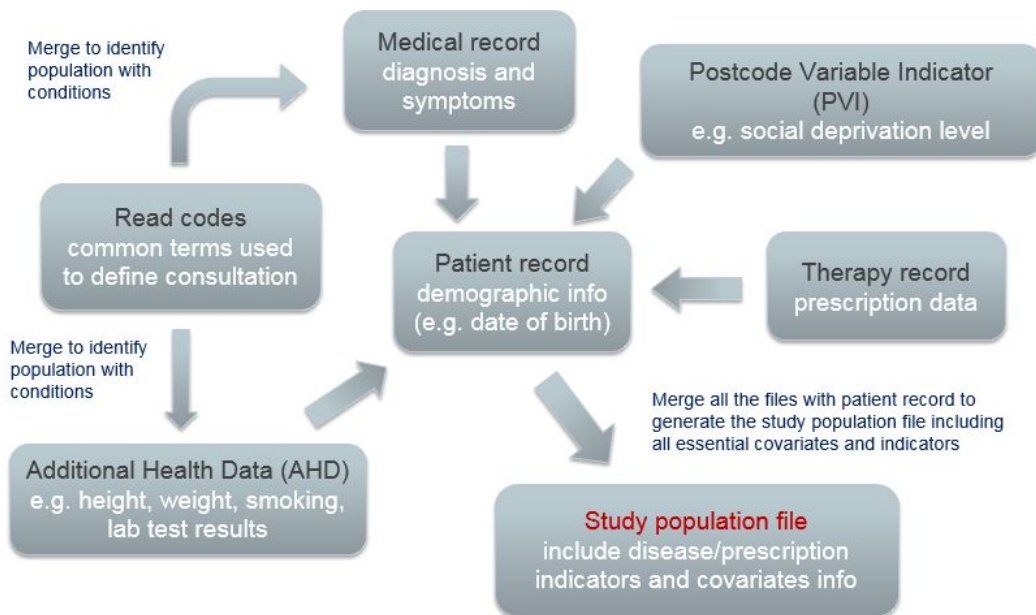
Figure 4.1: Overview of data extraction process from THIN. A list of Read codes defining a disease or its major symptoms is identified to merge with the medical records of patients for case identification. This file is then merged with all the other files (which contain other types of patient information) to generate the final study population file for further analysis.

95

4.3 Aim

The aim of this chapter was to firstly identify PCOS cases from the THIN database and then estimate the incidence and prevalence of PCOS in the UK general population. The second objective was to examine the impact of age, social deprivation level and time period on the incidence rates. Finally, I aimed at investigating the prescribing patterns of PCOS in the UK primary care.

4.4 Methods

4.4.1 Patient identification for PCOS

Specifically for PCOS, I primarily identified codes defining PCOS by looping through the description field including 'key' words such as 'polycystic ovary syndrome', 'endoscopic drilling of ovary', 'Stein-Leventhal syndrome' or the symptoms of PCOS (i.e. irregular menstrual dysfunction, hirsutism, polycystic ovary). The code list of diseases that mimic PCOS was also developed in the same way to be applied as exclusion criteria, including adrenal tumours, adrenal hyperplasia), Cushing's syndrome, Nelson's syndrome, pituitary disorders, prolactinoma. The reasons for exclusion are presented in Table 4.2. The resulting code list is provided in the Appendix E. This list was firstly developed based on the clinical guideline of PCOS and then was filtered and modified through the discussion with clinical experts.

The THIN patient files were firstly restricted to only females who permanently registered. The permanent patients refer to those who are residents in the practice area for more than 3 months. The identification process of the study population is graphically displayed in Figure 4.2.

Table 4.2: Diseases that mimic PCOS.

Disease	Reason to exclude
Cushing's syndrome	Make body produce excess cortisol hormone, leading to too much hair in women and abnormal menstrual periods
Nelson's syndrome	Caused by operations that used to treat Cushings syndrome
Prolactinemia	Over production of prolactin, leading to oligomenorrhea or amenorrhea, acne and excessive growth of facial and body hair
Adrenal disorders	Make body produce too little cortisol and too much male hormone, leading to severe and early acne (before teenage), facial hair in women and infrequent or absent menstrual periods
Pituitary disorders	Can cause less frequent or no menstrual periods
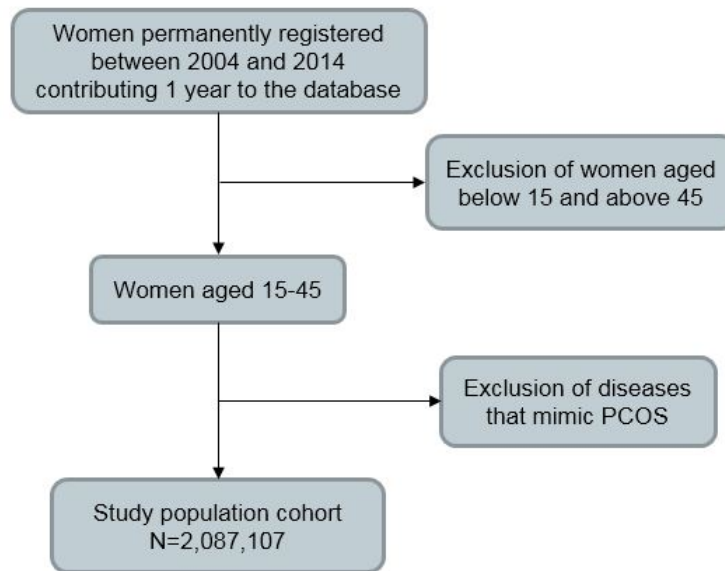
Figure 4.2: Identification of study population.

Two categories of PCOS cases were identified: diagnosed and probable cases. This is because as described in Chapter 2, large data gaps between the prevalence rates estimated by community studies and database studies were identified, suggesting that PCOS is a condition that is often under-diagnosed. Since a combination of symptoms can also be indicative of PCOS, I included probable cases in the analysis. It should be noted that the definition of probable cases was validated through consultation with clinicians.

I now introduce how patients in these two categories were defined. The diagnosed case group included PCOS patients with a record of either of the three Read codes: "Polycystic ovary syndrome" (Read code `C165.00`), "Stein-Leventhal syndrome" (`C164.12`) and "Endoscopic drilling of ovary" (`7E25300`) recorded in their medical or additional health records. The probable case group included patients with at least two Read codes from the following three groups recorded in their medical or additional health records: Group A (menstrual or ovarian dysfunction related PCOS features), Group B (androgen excess related PCOS features) and Group C (Polycystic ovaries). Moreover, the Read codes must come from at least two different groups, e.g. one from Group A and another from Group B or two from Group A and one from Group C. The requirement for the (minimum) time between the 2 recorded features (or any 2 out of 3 features) for probable cases must not exceed 3 years. This time period was considered to be long enough to capture half of the probable cases based on some pilot investigation of the

97

study population (refer to Figure 4.5 in Section 4.5.2). The longer the time interval between the two features, the less likely it is for a woman to be qualified as a case because one would suspect that the woman may have developed some other conditions rather than PCOS.

Group B codes were further divided into two subgroups (I and II). Subgroup I included Read codes for descriptive hyperandrogenic features (e.g. hirsutism, male pattern alopecia) whereas subgroup II included Read codes for a range of laboratory tests for androgen related markers. Subgroup I codes were considered if they were recorded in the medical or additional health records. Subgroup II codes were only included if they were recorded in the additional health records and have:

- A qualitative marker indicating that the test result was above the normal range ("`high`", "`very high`", "`significantly high`", "`above high reference limit`");

- A quantitative marker indicating that the test result was above the normal range (where the upper reference range for a quantitative result was not recorded in the additional health records, the mode of the upper ranges for that practice in that year was used; where both upper reference range and the mode of the upper ranges for that practice in that year were not recorded, the mode of the upper reference ranges for all practices in that year was used).

The following covariates were extracted, which may potentially be important factors affecting epidemiological outcomes of interest in the current analysis:

- **Age** — age in 2014 was used here and women were categorised into 5-year age bands, i.e. 15-19, 20-24, 25-29, 30-34, 35-39, 40-44 years.

- **Period** — the whole study period was divided into 3 short periods, i.e. 2004-2007, 2008-2011, 2012-2014.

- **Townsend score** — which describes the area-level social deprivation. The Townsend score is a census-based index incorporating four variables (i.e. unemployment, non-car ownership, non-home ownership, household crowding). Note that percentages rather than absolute values are computed for each of the variables. For example, the percentage of households without access to a car or van within an area (which is usually small geographical areas such as wards or local authorities across the whole country). These percentages

are then standardised to a Z-score. The formulas for standardisation are shown below (the first formula is for variables non-car ownership and non-home ownership and the second formula is for variables unemployment and household overcrowding where log transformation is considered necessary):

$$\text{Z scores} = \text{(percentage} - \text{mean of all percentages)}/ \text{SD of all percentages}$$

$$\text{Z scores} = \text{(log percentage} - \text{mean of log percentages)}/ \text{SD of log percentages}$$

These standardised scores are then added up to be the final Townsend score, which ranges from 1 to 5. Individuals with a score of 1 are considered to live in the least deprived areas whereas those with a score of 5 reside in the most deprived areas (Townsend et al., 1988). Similar to the wide accepted index of multiple deprivation (Department for Communities and Local Government, 2015), Townsend score can be used to measure the deprivation level of any geographical area where census data are available.

Since ethnicity information in THIN is incomplete and only reported by half of the patient population, this variable was not included in the current analysis.

4.4.2 Analysis plan for PCOS

I performed analysis of both prevalence and incidence for the study population from 2004 to 2014. The start date was selected as 2004 because the second diagnostic criterion of PCOS was established in 2003 and results from the literature review have suggested that very few prevalence studies have been conducted before 2003. Therefore, it may be inferred that the diagnostic criterion of PCOS was known by very few GPs and consequently hardly followed. In other words, the recording rates of PCOS are expected to be lower before 2003, which may bring in potential bias for the current analysis. In terms of the study end date, since the study was performed in 2015, the end date was defined as the last date of 2014.

The **incidence** of PCOS records in the period from 2004 to 2014 were calculated for diagnosed cases, probable cases and both diagnosed and probable cases combined. The denominator for the incidence calculation was the sum of all eligible follow-up time in the study population with the start of each woman's follow-up being defined as the latest of:

- 5,478 days after a woman's date of birth (i.e. 15 years of age);

- 1st January 2004;

- 365 days after a woman's registration date;

- The date when practices met the criteria for continuously acceptable computer usage (ACU), i.e. one medical record, one additional health record per patient per year, and at least two prescription per patient per year (Horsfall et al., 2013);

- The date when practices were determined to have acceptable mortality reporting (AMR), i.e. a time point at which the observed death rate for a practice reaches the standard predicted numbers of deaths derived from national statistics given the practice's demographics (Maguire et al., 2009).

The end date of each woman's follow-up was defined as the earliest of:

- 16,436 days after a woman's date of birth (i.e. 45 years of age);

- The date a woman transferred out of the practice;

- 31st December 2014;

- The date data were last collected from a woman's practice;

- The date of the first PCOS diagnosis (for the probable case group, the date the second PCOS feature is recorded was used as the date of first diagnosis of PCOS regardless of the time interval between the first and second PCOS feature).

If presented in formula format, the start and end date were then defined as:

$$\text{start date} \;=\; \max\left\{15 \text{ years, 1st Jan. 2004, registration date} + 1 \text{ year, ACU date, AMR date}\right\}$$
$$\text{end date} \;=\; \min\left\{45 \text{ years, 31st Dec. 2014, transfer out date, date of the first PCOS diagnosis}\right\}$$

The underlying reason for only including women in reproductive age range is that the major symptoms of PCOS (i.e. menses) only appear within this age range and it is uncertain whether a woman has developed PCOS before menarche or after menopause.

The incidence rates (raw data) were computed using the following formula:

$$\text{Incidence rates} \;=\; \frac{\text{Total number of events (new PCOS cases)}}{\text{Total number of person-year at risk from 2004 to 2014}}$$

The annual rates were then graphed to investigate the time trends of the diagnosis of PCOS (diagnosed cases, probable cases and both diagnosed and probable cases combined). Moreover, I graphed the incidence rates (also raw data) according to age bands to see the change of rates across different age groups. Multivariable Poisson regression models were used to evaluate recordings of all PCOS cases by age (in 5-year age band), social deprivation (quintiles of Townsend scores) and effects of year (given the changing definition of PCOS, the study period was split into 3 periods, i.e. 2004-2007, 2008-2011, 2012-2014) as well as to account for clustering effect of patients nested in different practices. A two-level Poisson regression model was considered here. Suppose there are $j = 1, 2, ..., N$ practices with each practice consisting of $i = 1, 2, ..., m_j$ patients. I then model:

$$y_{ij} \sim \text{Poisson}(\lambda_{ij})$$
$$\log(\lambda_{ij}) = \boldsymbol{x}_{ij}\boldsymbol{\beta} + \boldsymbol{u}_j$$

where y_{ij} are counts and $\lambda_{ij} = \exp(\boldsymbol{x}_{ij}\boldsymbol{\beta} + \boldsymbol{u}_j)$; the vector $\boldsymbol{x}_{ij}$, a $1 \times p$ row vector, represents the covariates for the fixed-effects part of the model with regression coefficients $\boldsymbol{\beta}$; $\boldsymbol{u}_j$ is the random effects components, which is a multivariate normal distribution with mean $\boldsymbol{0}$ and $q \times q$ variance matrix $\boldsymbol{\Sigma}$. The variance matrix is a diagonal matrix structured as $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \ldots & 0 \\ 0 & \sigma_2^2 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \sigma_q^2 \end{bmatrix}$.

For the sake of simplicity and in view of the large amount of data (which would dominate in the absence of particularly strong prior information on the model parameters), I have conducted this exploratory analysis under a standard Frequentist approach. Note that this piece of analysis will be fully expanded to a proper Bayesian model in Chapter 5 because the uncertainties associated with both parameters and models can easily be propagated to the rest of the decision model (e.g. Markov model) by using a Bayesian approach.

The period **prevalence** of PCOS records in calendar year 2014 was calculated for diagnosed cases, probable cases and both diagnosed and probable cases combined (see formula below).

$$\text{Prevalence} = \frac{\text{Total number of PCOS cases in 2014}}{\text{Total number of women eligible to be included in 2014}}$$

The denominator of the prevalence study consisted of any woman with at least 12 months of post-registration follow up, at least 6 months of which must have occurred in 2014. The numerator consisted of all women in this denominator who were identified as PCOS cases defined by the previous section. Secondary analysis in terms of post-registration period (i.e. 1 year, 2 years) and time registered within 2014 (i.e. 3, 6, 9 months) were also performed.

PCOS cases were then stratified by age (5-year age band) to investigate the age-specific prevalence of PCOS for the study population.

The followings were also graphed to explore the characteristics of the patient cohort: (i) the time distribution for a probable case to be formally indexed as a diagnosed case (diagnosed cases with two or more features recorded before the date of their confirmed diagnosis were looked at); (ii) the time between the index date of the first two diagnoses of PCOS features for probable cases; (iii) the distribution of time between the date of PCOS diagnosis for cases and the end of study period (31st Dec. 2014) stratified by age groups.

4.4.3 Analysis of drug treatment

Since there is variation in the symptoms of PCOS, it is expected that PCOS patients may receive a wide range of prescriptions depending on their clinical presentations. Consequently, different chapters of the British National Formulary were searched to identify drugs that are most relevant to PCOS. e.g. oral contraceptives, metformin. I referred to the recommendations of treatments in the PCOS guideline and other published studies on this aspect (Sheehan, 2004; Conway et al., 2014).

The **first** record of a list of relevant treatments to PCOS was extracted both before and after the diagnosis of PCOS. For each treatment included, the followings were computed:

- The proportion of PCOS patients who were on a certain type of drug before the diagnosis of PCOS;

- The proportion of PCOS patients who newly (i.e. not on this drug before the diagnosis of PCOS) initiate a certain type of drug following the diagnosis of PCOS.

Given the post-registration time of individual patients (1 year), the cumulative incidence plots of different drug categories were restricted to 2004 to 2012 (censored two year before the

end of study period in case those who started in 2012 do not have enough follow-up time as the post-registration time requirement is 1 year).

4.5   Results

4.5.1   Prevalence and incidence outcomes

In total, over 14,000 individuals (amongst 2,087,107 eligible women) aged 15-45 years old had a medical record indicative of potential PCOS between 2004 and 2014, which equated to 1.84 per 1000 person-year at risk (PYAR, 95% CI: 1.81-1.87). The rate for diagnosed PCOS cases was 0.93 per 1000 PYAR (95% CI: 0.91-0.96) whereas the rate for probable cases was 0.91 per 1000 PYAR (95% CI: 0.89-0.93).

The change of incidence rates over time and the age-specific rates are presented in Figure 4.5.1. The overall incidence of PCOS increased from 1.67 (95% CI: 1.58-1.77) per 1000 PYAR in 2004 to 2.00 (95% CI: 1.89-2.10) per 1000 PYAR in 2010, after which the rate remained relatively constant at approximately 2 per 1000 PYAR. The incidence rates after 2007 are significantly higher compared with that before 2007 for probable cases as shown in Table 4.3.

The incidence was the highest for those in the 20-24 age group (3.59 per 1000 PYAR, 95% CI: 3.47-3.70) whereas the age group 40-45 reported the lowest incidence at a rate of 0.62 (95% CI: 0.58-0.66) per 1000 PYAR. The age-specific trend of PCOS diagnosis was similar for both diagnosed and probable cases. After adjusting for effects of year, social deprivation and practice-level variability, significant differences still remained in the incidence of PCOS, i.e. the incidence rates estimated are significantly lower for all the other age groups compared with the reference group (Table 4.4).

In terms of social deprivation (measured by Townsend score), there was an increase in recording of PCOS for individuals living in more deprived areas. The incidence of PCOS for individuals who lived in the least deprived areas was 1.59 (95% CI: 1.53-1.65) per 1000 PYAR whereas those in most deprived areas reported a rate of 2.23 (95% CI: 2.15-2.32) per 1000 PYAR. This was statistically significant even after adjusting for effects of other covariates (i.e. age, year) and practice-level variability (Table 4.4).

The random effects variance estimates suggest that there are some differences in the incidence estimates across practices.

The overall prevalence of PCOS in 2014 was approximately 2.27% (95% CI: 2.23-2.31%), with diagnosed and probable cases reporting 1.34% and 0.93%, respectively. The age-specific prevalence was the highest for the 30-34 age group and decreased for older age groups. Prevalence estimates were not sensitive to the varying post-registration period and the minimal time registered in 2014, remaining consistently at about 2%.
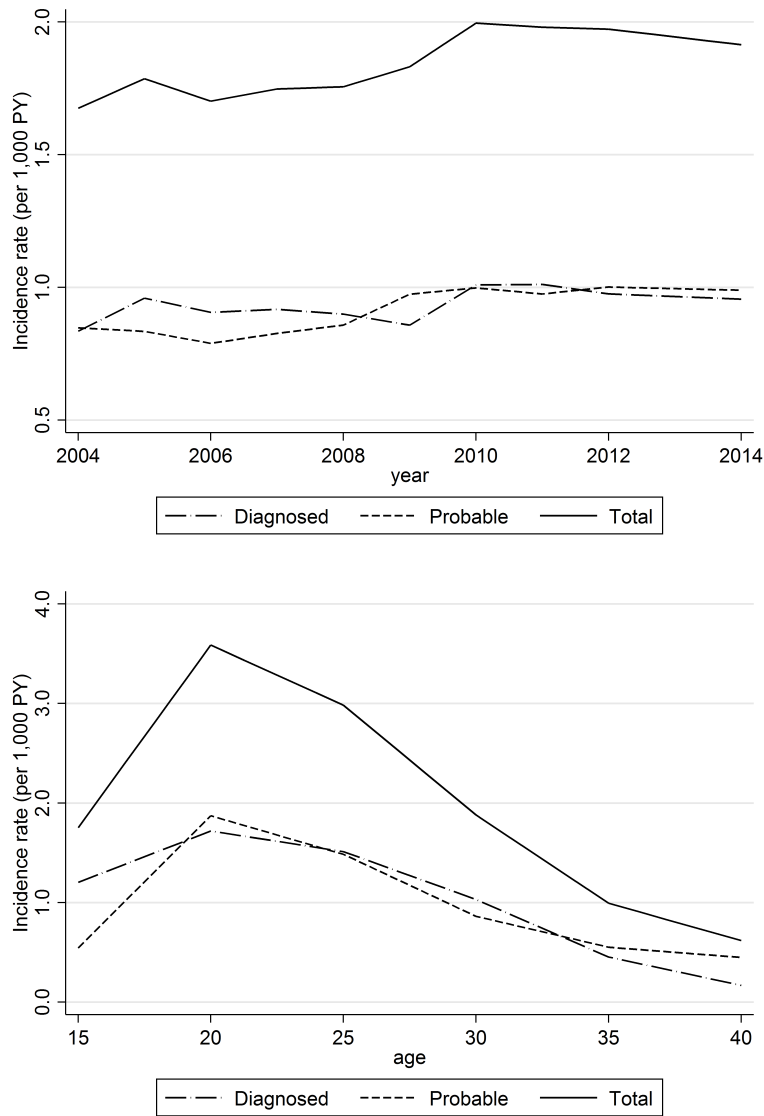
Figure 4.3: Time trends in PCOS diagnosis record (top) and age-specific PCOS diagnosis record (bottom) for diagnosed, probable and total cases.

Table 4.3: Recorded rates (raw rates without modelling) of PCOS by definition of PCOS, social and demographical characteristics.

	Diagnosed PCOS	Probable PCOS	Overall
	Rate per 1000 PYAR (95%CI)	Rate per 1000 PYAR (95%CI)	Rate per 1000 PYAR (95%CI)
Townsend quintile			
1	0.80 (0.76-0.84)	0.80 (0.76-0.84)	1.59 (1.53-1.65)
2	0.90 (0.85-0.95)	0.79 (0.75-0.84)	1.69 (1.62-1.75)
3	0.94 (0.90-0.99)	0.91 (0.86-0.96)	1.85 (1.78-1.92)
4	1.02 (0.97-1.07)	0.97 (0.92-1.02)	1.98 (1.91-2.05)
5	1.07 (1.01-1.13)	1.17 (1.11-1.23)	2.23 (2.15-2.32)
Age, years			
15-19	1.20 (1.14-1.27)	0.54 (0.50-0.59)	1.75 (1.67-1.83)
20-24	1.72 (1.64-1.80)	1.87 (1.79-1.96)	3.59 (3.47-3.70)

	Diagnosed PCOS	Probable PCOS	Overall
	Rate per 1000 PYAR (95%CI)	Rate per 1000 PYAR (95%CI)	Rate per 1000 PYAR (95%CI)
25-29	1.51	1.49	2.98
	(1.44-1.58)	(1.42-1.56)	(2.88-3.08)
30-34	1.03	0.86	1.88
	(0.98-1.09)	(0.81-0.91)	(1.81-1.96)
35-39	0.45	0.55	0.99
	(0.42-0.49)	(0.51-0.59)	(0.94-1.05)
40-45	0.17	0.45	0.62
	(0.15-0.19)	(0.42-0.48)	(0.58-0.66)
Period			
2004-2007	0.91	0.82	1.73
	(0.87-0.94)	(0.79-0.86)	(1.68-1.78)
2008-2011	0.94	0.95	1.89
	(0.91-0.98)	(0.92-0.99)	(1.84-1.94)
2012-2014	0.96	0.98	1.92
	(0.92-1.01)	(0.93-1.02)	(1.86-1.98)

Table 4.4: The incidence rate ratios (IRR) and the random-effects variance estimates obtained from multilevel Poisson regression model are provided.

	Diagnosed PCOS	Probable PCOS	Overall
	Adjusted IRR (95%CI)	Adjusted IRR (95%CI)	Adjusted IRR (95%CI)
Townsend quintile			
1	1	1	1
2	1.14 (1.06-1.23)	0.99 (0.92-1.07)	1.08 (1.02-1.14)
3	1.12 (1.04-1.21)	1.07 (0.99-1.15)	1.10 (1.05-1.16)
4	1.15 (1.06-1.24)	1.06 (0.98-1.15)	1.11 (1.05-1.17)
5	1.15 (1.05-1.25)	1.21 (1.11-1.32)	1.19 (1.11-1.26)
Age, years			
15-19	0.69 (0.64-0.74)	0.29 (0.27-0.32)	0.49 (0.46-0.51)
20-24	1	1	1
25-29	0.86 (0.80-0.91)	0.78 (0.73-0.83)	0.81 (0.78-0.85)

	Diagnosed PCOS	Probable PCOS	Overall
	Adjusted IRR (95%CI)	Adjusted IRR (95%CI)	Adjusted IRR (95%CI)
30-34	0.58	0.46	0.51
	(0.54-0.62)	(0.42-0.49)	(0.49-0.54)
35-39	0.26	0.30	0.27
	(0.24-0.28)	(0.27-0.32)	(0.26-0.29)
40-45	0.10	0.24	0.17
	(0.08-0.11)	(0.22-0.26)	(0.16-0.18)
Period			
2004-2007	1	1	1
2008-2011	1.00	1.13	1.06
	(0.95-1.06)	(1.07-1.19)	(1.02-1.10)
2012-2014	0.98	1.13	1.04
	(0.92-1.04)	(1.07-1.20)	(1.00-1.09)
Random-effects variance estimates			
	0.52	0.32	0.25
	(0.44-0.61)	(0.27-0.38)	(0.22-0.29)

4.5.2 Features of patients

The proportion of diagnosed cases who had a PCOS feature of menstrual dysfunction, hyper-andrgenism and polycystic ovaries (PCO) recorded was 28.4%, 39.2% and 2.8%, respectively (Table 4.5). The results suggested that 8.3% of the diagnosed cases used to be a probable case and the probability for a probable case to be formally indexed as a diagnosed case within 1 years was about 7% (Figure 4.4). On the other hand, menstrual dysfunction, hyperandrogenism and PCOS were captured in 88.8%, 88.2% and 23.2% of the probable cases, respectively.

Table 4.5: Number and proportion of diagnosed and probable cases with major PCOS features.

Features No. (%)	Diagnosed cases (n=7233)	Probable cases (n=7057)
Menstrual dysfunction	2055 (28.4)	6265 (88.8)
Hyperandrogenism	2836 (39.2)	6221 (88.2)
Polycystic overies	199 (2.8)	1636 (23.2)
≥2 features	597 (8.3)	7057 (100)
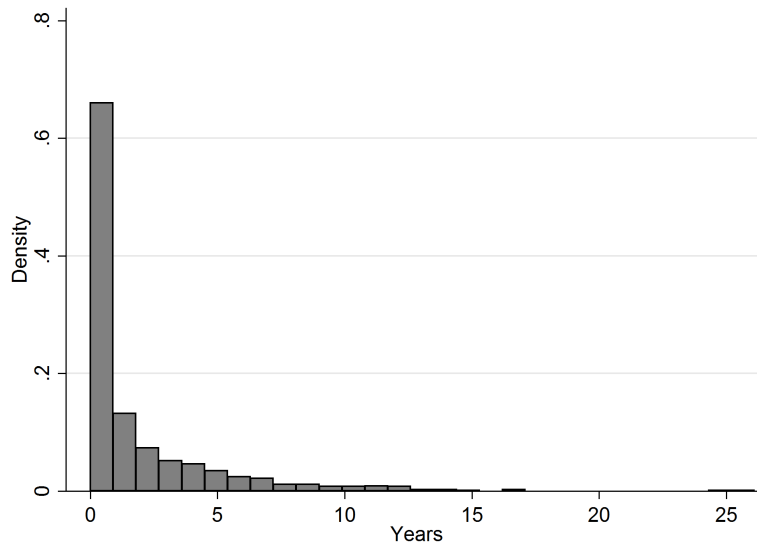
Figure 4.4: Histogram displaying the time distribution for a probable case to be formally indexed as a diagnosed case.

The time between the index date of the first two diagnoses of PCOS features for probable cases is presented in Figure 4.5. It suggests that approximately half of the probable cases have their second PCOS feature recorded within 3 years after the index date of their first PCOS

feature. This 3-year interval was used as the maximum inter-feature period for the identification of probable cases as discussed in the methods section.
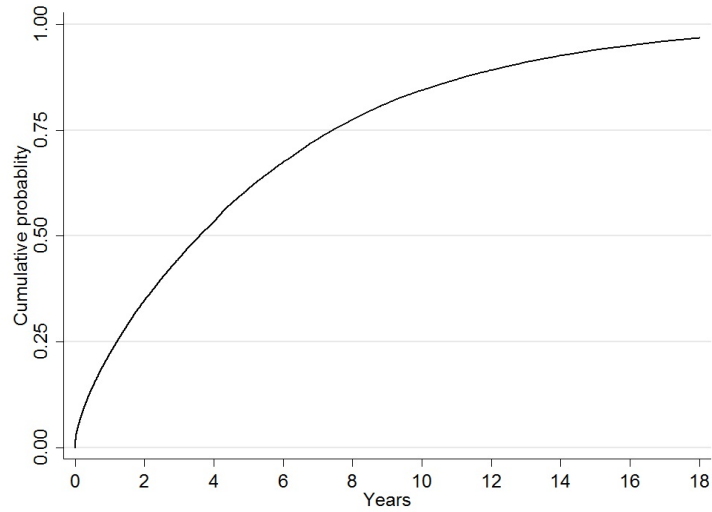
Figure 4.5: The time between the index date of the first two diagnoses of PCOS features for probable cases.

As presented in Figure 4.6, older women with PCOS were more likely to have their diagnosis recorded longer time ago (10-15 years ago) whereas for young women with PCOS, their PCOS recordings were unlikely to be indexed more than 10 years ago.
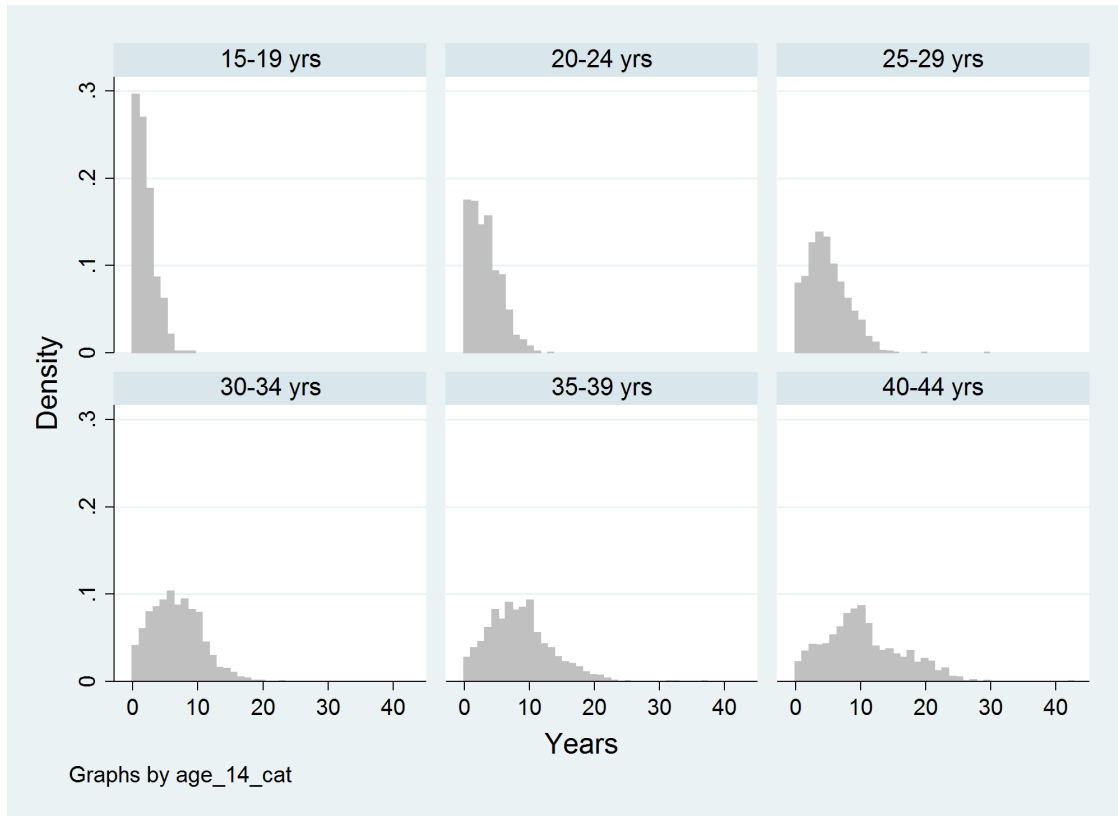
Figure 4.6: Distribution of time between the date of PCOS diagnosis and the end of study period (31st Dec. 2014) stratified by ageband. Older women with PCOS generally have their date of diagnosis recorded longer time ago compared with younger women with PCOS.

4.5.3 Prescribing patterns

As presented in Table 4.6, the results have demonstrated that before the diagnosis of PCOS, a large proportion (over 40%) of women were on combined oral contraceptives (COC), followed by acne-related drugs (about 30%) and then progestin oral contraceptives (POC, about 20%). Approximately 18% of the women had previously been prescribed with at least one of the remaining drugs, i.e. intrauterine devices, clomiphene, metformin, gonadotrophins, spironolactone, cyproterone, flutamide, eflornithine, weight control/loss drugs, lipid regulators. The cumulative percentage of patients who newly initiated acne-related drugs, COC and metformin was the most common drugs prescribed in the 24 months after a diagnosis of PCOS. The cumulative incidence plots show that while there is an initial increase in prescribing on or within a short period after the PCOS index date, this is greater for metformin, acne-related drugs compared with others (COCs, POCs). Prescription results stratified by case definition indicated that acne-related

111

drugs and POCs were more commonly prescribed for probable than diagnosed cases whereas COCs, metformin, clomiphene, cyproterone, eflorinithine and weight loss drugs were prescribed more commonly in diagnosed than probable cases (Figure 4.7).

Table 4.6: Number and percentage of PCOS women on relevant drugs for PCOS both prior to and following the diagnosis of PCOS.

Types of drugs No. (%)	Before	After 2004-2007	After 2008-2011	After 2011-2014
COC	12349 (40.22)	901 (17.01)	912 (18.87)	459 (17.88)
POC	5806 (18.91)	474 (6.45)	691 (10.22)	272 (7.80)
IUDs	733 (2.38)	44 (0.51)	64 (0.75)	25 (0.52)
Clomiphene	518 (1.69)	210 (2.46)	165 (1.92)	70 (1.43)
Metformin	1278 (4.16)	1212 (14.77)	1102 (13.25)	495 (10.52)
Gonadotrophins	435 (1.42)	29 (0.34)	13 (0.15)	13 (0.26)
Spironolactone	235 (0.77)	109 (1.26)	111 (1.28)	43 (0.87)
Cyproterone	91 (0.30)	28 (0.32)	19 (0.22)	3 (0.06)
Flutamide	6 (0.02)	4 (0.05)	1 (0.01)	0
Eflornithine	530 (1.79)	354 (4.10)	407 (4.81)	172 (3.58)
Weight control/loss drugs	1330 (4.33)	403 (4.84)	364 (4.43)	91 (1.95)
Lipid regulators	194 (0.63)	69 (0.79)	46 (0.53)	8 (0.16)
Acne-related drugs*	9000 (29.31)	1182 (19.52)	1234 (20.78)	582 (18.32)

COC: combined oral contraceptives; POC: progestin oral contraceptives; IUDs: intrauterine devices; *Acne-related drugs also include a range of topical treatments, e.g. benzoyl peroxide cream
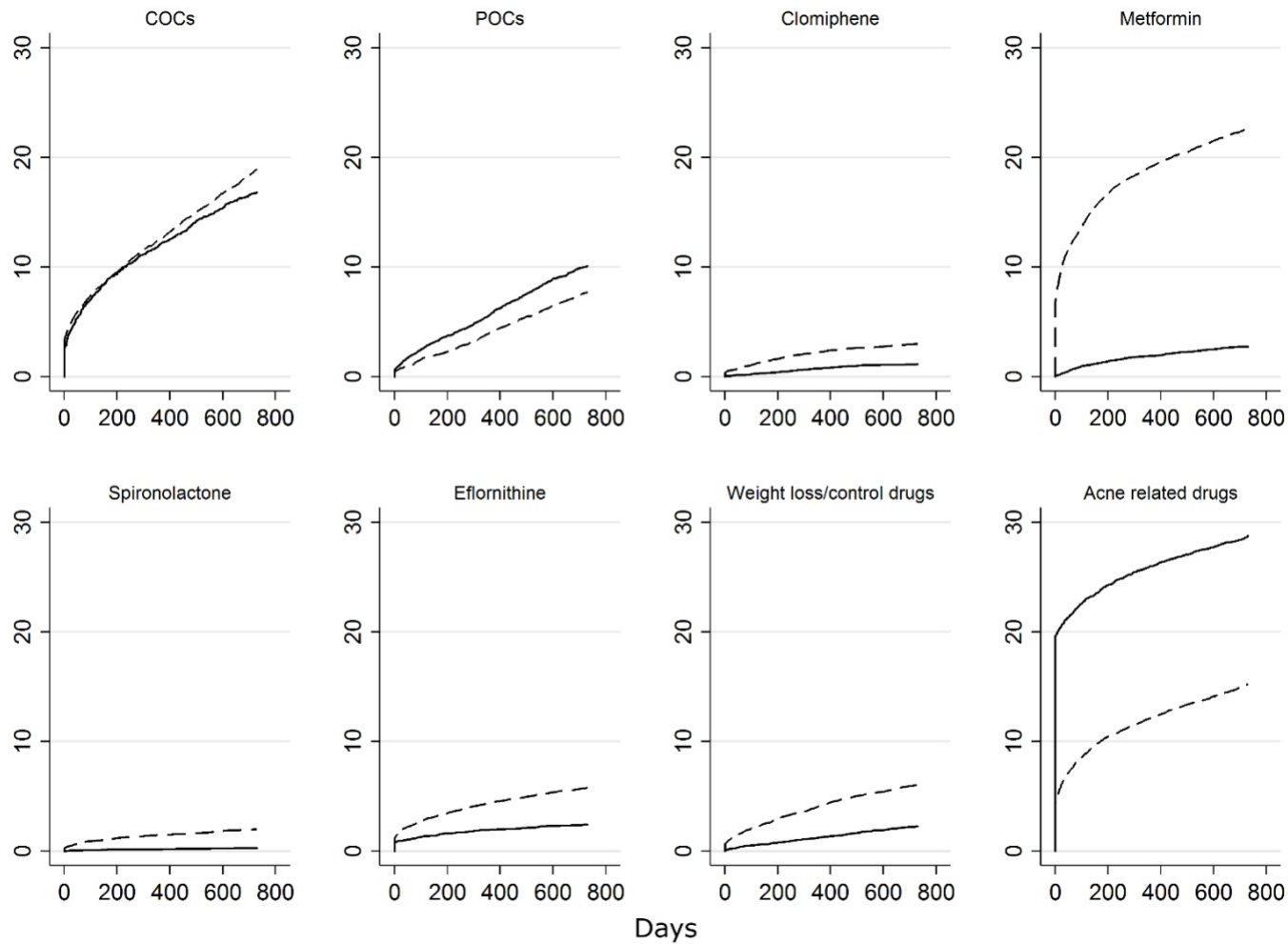
Figure 4.7: Cumulative incidence of women with a prescription for each drug type over the 24 months following their index date stratified according case definition (dashed line for diagnosed case, solid line for probable case). Results shown for the eight most commonly prescribed drugs. COC, combined oral contraceptive; POC, progestin oral contraceptive.

4.6 Discussion

4.6.1 Summary of main findings

For the current analysis, over 14,000 potential PCOS cases in women of reproductive age were identified in practices across UK during the 10-year period from 2004 to 2014. More than half of the cases identified have received a confirmed diagnosis of PCOS (diagnosed cases) whereas the remaining women only have features indicative of PCOS (probable cases). This corresponds to an incidence rate of 0.93 per 1000 PYAR (95% CI 0.91 to 0.96) and 0.91 per 1000 PYAR (95% CI 0.89 to 0.93) for the diagnosed and probable cases, respectively. The prevalence of PCOS in 2014 was estimated to be 2.27%. A wide range of treatments have been initiated in the primary care for women with PCOS and the prescribing pattern varies for diagnosed and probable cases. The results suggested that oral contraceptives have been frequently used by women with PCOS both prior to and after the diagnosis of PCOS was recorded. A greater proportion of diagnosed cases received metformin on the day, or after the diagnosis has been recorded while a greater proportion of probable cases received treatments tailored to the external symptoms of PCOS such as acne.

4.6.2 Comparison with literature

The prevalence of PCOS in the UK primary care in 2014 was comparable to that obtained from studies using other databases with a range from 0.56% to 2.22% (Lo et al., 2006a; Okoroh et al., 2012; Christensen et al., 2013; Sirmans et al., 2014a). However, the rates are significantly lower than those from community-based studies in Europe where rates generally ranged from 5.4% to 8% were observed using the NIH criterion (Michelmore et al., 1999; Diamanti-Kandarakis et al., 1999; Asunción et al., 2000; Sanchón et al., 2012).

The proportion of women with PCOS who had been using oral contraceptives prior to their PCOS diagnosis is comparable to other studies (Okoroh et al., 2012; Glintborg et al., 2015). The proportion of PCOS patients who initiated metformin following a diagnosis is also similar to a Danish study that reported the percentage of the PCOS population in their sample who have received metformin to be about 12% (Glintborg et al., 2015).

4.6.3 Interpretation

The incidence of PCOS increased slightly over the study period although no significant changes were observed in any consecutive year. This may suggest that there was increasing awareness of this syndrome after the gradual establishment of major criteria of PCOS. However, the fact that the database has been improved over the time could also be an underlying reason for the increasing rates.

Results from the multilevel Poisson regression model suggested that women who lived in more deprived areas had a larger incidence of PCOS compared with those living in the less deprived areas. This could be due to that obesity is more prevalent among women living in more deprived areas. Alternatively, women living in these areas may have more complications and therefore visit their GPs more frequently to have PCOS diagnosed and recorded.

The incidence of probable cases was comparable to that of diagnosed cases, indicating that a large number of women who present with features of PCOS (within a 3 year period) in primary care do not receive PCOS diagnosis subsequently. However, it is possible that some of these probable cases have been referred to the secondary care for a confirmed diagnosis whereas a diagnosis has not been recorded in the primary care. While a confirmed diagnosis may not be relevant for some of the probable cases, it is likely that some of the women with features of PCOS actually meet the criteria of PCOS and should receive further assessment. Failure to refer such women may result in that these women are not offered the proper care for the treatment of PCOS such as lifestyle modification or medications. Considering the relative low cost of screening/diagnosis and the high cost of care for the associated disorders (Azziz et al., 2005), further work is needed to inform women and health care providers about this condition.

There is considerable variation in the treatments prescribed to diagnosed and probable cases, for example, a greater proportion of diagnosed women received metformin as medication whereas many probable cases received medication tailored to the features they present with. This suggests that diagnosed and probable cases indeed received different care for this syndrome and some of the probable cases may not receive the most effective treatments. The wide variation in prescribing pattern could also be explained by the fact that the clinical presentations of PCOS not only varies by individual but also by age, i.e. younger women are more likely to initiate drugs to regulate their menses while older women may ask for drugs to prevent rapid conversion to diabetes.

The most common drugs prescribed for women with diagnosed PCOS were metformin and oral contraceptives, possibly reflecting the major long-term metabolic concerns of this syndrome as indicated by the PCOS consensuses. However, even among women with diagnosed PCOS, the treatments prescribed varied, indicating the lack of consensus for the ideal treatment of PCOS. Therefore, there is much potential for future research to investigate the most cost-effective treatments for this condition.

4.6.4 Strengths and limitations

To the best of our knowledge, this piece of analysis is the first to investigate the diagnosis and management of PCOS in the primary care setting in the UK. THIN contains patient information from more than 500 practices and our study population included more than 2,000,000 women. Therefore, the current analysis is advantageous in terms of sample size and because the data were contributed from practices across the UK, the sample was also considered to be broad representative for the UK population. The trends of PCOS recording was examined over a 10-year period, which has not been explored by any previous epidemiological studies where study population was often sampled at a single time point.

The results in the current study are largely likely to reflect the true burden of PCOS for the healthcare system since the data were coming from the routine clinical practice. However, notice that since only data considered relevant at a consultation are recorded, underdiagnosis was a major concern for this analysis. Consequently, it is not surprising to observe a low proportion (i.e. 8%) of diagnosed cases with 2 or more features of PCOS recorded because once a confirmed diagnosis has been made by a specialist, general practitioners are unlikely to record anything further. Attempts have been made to identify and include women with features of PCOS (interfeature within 3 years) as probable cases to address the under-reporting issue. This may introduce case misclassification since some of the probable cases may not be true cases. On the contrary, it could also be that some of the probable cases indeed have developed the condition but were not captured and recorded for some reasons.

Another issue is considered to be the incomplete information of ethnicity (less than 50% of the patients have this variable recorded in THIN) and the incidence rates can be influenced by the unobserved differences in ethnicity distributions across the other variables such as age and deprivation level.

The indication of drugs is not specified in the prescription files of each patient and therefore it is uncertain if the drugs prescribed following a diagnosis or even on the day of a record were indeed for the treatment of PCOS. The results suggested that approximately 30% of the PCOS patients prescribed metformin had a diagnosis of type 2 diabetes before their PCOS diagnosis. By excluding these patients, it is more certain that metformin was prescribed at least partly for the treatment of PCOS, with the proportion of patients initiated the drug lying somewhere between 8% (on the day of diagnosis) and 20% (cumulative proportion within 24 months). The advantage of using a 24-month window is that it allowed to capture prescriptions initiated by specialists in secondary care that are transferred back to primary care. This is supported by the cumulative incidence curves because the initiation of drugs that are more commonly prescribed in secondary care (e.g. spironolactone) was further away from the origin (the day of PCOS diagnosis).

In the next chapter, I will explore the healthcare burden of PCOS and the quality of life for PCOS population in the UK by synthesising evidence from previous chapters using modelling approach.

Chapter 5

MODELLING THE BURDEN OF PCOS IN THE UK

5.1 Chapter summary

The previous chapters suggested that PCOS is associated with a range of morbidities, however the major concern for PCOS patients raised by most endocrine societies is the elevated risk of type 2 diabetes. Based on the evidence identified in the previous chapters, it is likely that a substantial proportion of PCOS patients may develop type 2 diabetes later in life (e.g. after the age of 35 or 40), exerting large economic impact on the National Health Service (NHS) in the UK. Therefore, it is of great interest to project the prevalence/incidence of diabetes in the PCOS population and evaluate the potential burden of disease as well as the quality of life (QoL) of individual patients in the near future. In such way, one is likely to have a clear view of the public health impact of PCOS and can attempt to propose some ways to improve the current situation. This chapter introduces the economic evaluation of disease burden and the analysis of QoL associated with PCOS under the UK context. I firstly gives an overview of the methods commonly used for the health economic evaluation. After this, the specific methods for modelling the population dynamics of PCOS (i.e. a multistate Markov model) and the sources of data used to populate the model are described. I then present outcomes from simulations (based on a virtual PCOS cohort estimated from the UK census data) in terms of the number of patients ending up in different states of the model, the healthcare burden of disease and the QoL for the entire PCOS population in the UK over a relatively long period of follow-up. The results

of QoL were compared with the simulated outcomes for a virtual healthy cohort of the same population size. The main findings in this chapter have been converted into a paper submitted to `Human Reproduction` and is currently under revision.

5.2   Background

In clinical decision making, a decision tree can often model the prognosis of patients and help to guide the choice of management strategy for the condition (Elwyn et al., 2001). Figure 5.1 presents a possible decision tree for PCOS patients who come to see the clinician for treatment consultation. A woman with PCOS may already have developed diabetes and therefore, treatments for diabetes such as metformin should be prescribed to such patients. However, if the woman have not developed diabetes yet, the clinician may prescribe treatments tailored to the most prominent external symptoms of the patient. For example, as suggested in Figure 5.1, a non-diabetic but obese patient may receive orlistat and is asked to take a diet. Conversely, a non-diabetic and non-obese patient may only need drugs such as oral contraceptives to regulate menstrual dysfunction. It should be noted that although drugs for disease management are prescribed for all the non-diabetic patients, an adverse event may still occur, e.g. non-obese patient who receive oral contraceptives may still develop diabetes. For each clinical scenario, the associated cost and utility can be measured. For example, for obese patients who received orlistat and take a diet but still develop diabetes later in life, the overall cost is calculated as the cost spent on orlistat and a diet plus the cost of treating diabetes. Similarly, the utilities can be attached to patients under each clinical scenario, e.g. we would expect that patients who eventually develop diabetes have lower quality of life compared with those who do not develop diabetes.

However, there are a few shortcomings for this model. The most obvious one is that the time when an event occur is not clearly indicated. Moreover, an adverse event may occur more than once. For example, some women with PCOS may lose weight and have normal BMI after a period of diet management while become obese again. The decision tree can be very complicated to depict if there are too many branches and one may find it difficult to assign utilities for each of the outcomes, e.g. non-obese PCOS patients without diabetes.

Consider another scenario where there are only four possible health states: *Healthy*, *Disease*, *Recovery*, *Death*. If we use a decision tree to describe this, then a patient starts in healthy state,

119
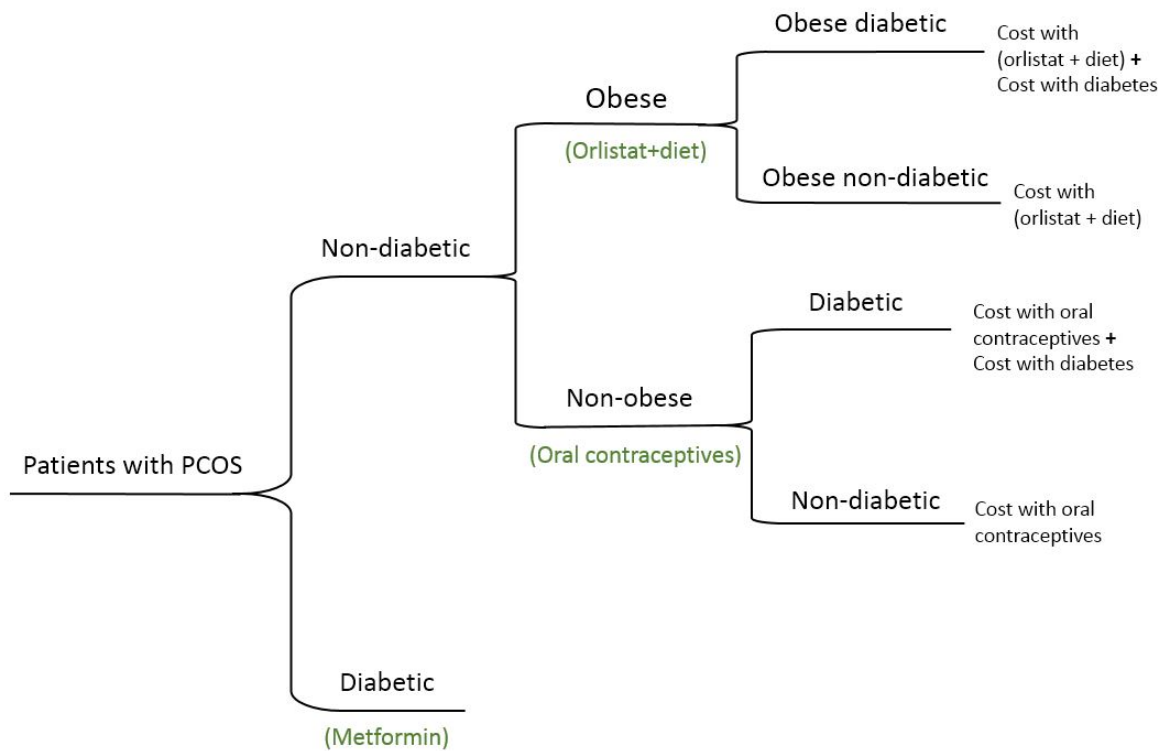
Figure 5.1: A possible decision tree for PCOS patients who come to see the clinician for drug consultation. The recommended prescriptions for patients with different phenotypes of the condition are indicated in green.

may develop the disease, stay healthy or die. However, apart from staying in death state (transit from either healthy or disease state) forever, the patient can transit from healthy to disease state (or back) recurrently. Consequently, the decision tree would get too many branches (see Figure 5.2) and become very complicated to analyse.

However, this situation can be much better depicted and analysed using Markov models (MM, Sonnenberg and Beck, 1993). In a MM, a finite set of states (assumed to be mutually exclusive and exhaustive) are used to represent possible health status of a progression disease. It is assumed that patients can transit from one state to another with some probabilities which may be based on patient characteristics such as age and sex or time point. For example, Figure 5.3 presents the structure of a MM including the four health states defined previously. Arrows indicate that individuals can move between the two states from the one where the arrow originates to the one where it ends. The absence of an arrow indicates that the movement between the two states is not possible. The healthy and disease state are recurrent states. The
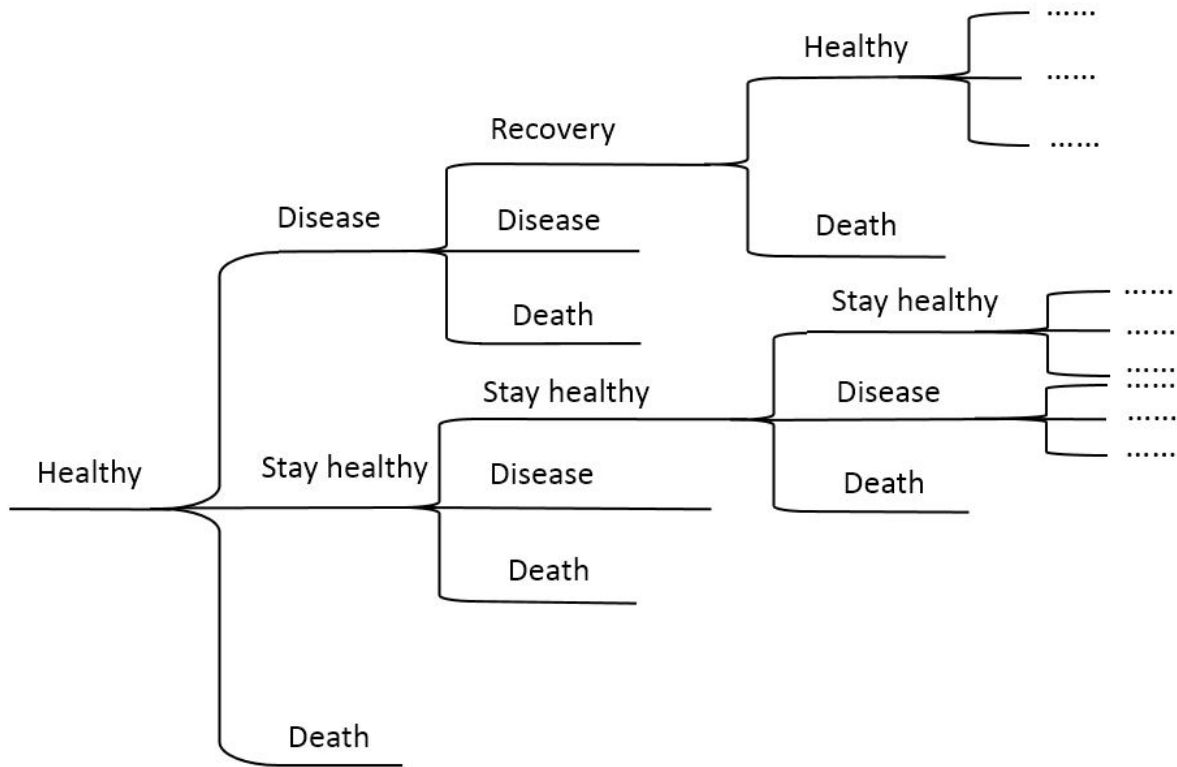
Figure 5.2: A decision tree for a scenario that only includes four possible health states: *Healthy*, *Disease*, *Recovery*, *Death*. A patient starts in healthy state may develop the disease, stay healthy or die. Apart from staying in death state (from either healthy or disease state) forever, the patient can transit from healthy to disease state (or reversely) recurrently.

death state is an absorbing state such that once individuals move there, they will remain there forever. In this way, one can easily get rid of the large number of branches of the decision tree. Given that MM provides a much more convenient and efficient way to model the likely outcomes of a disease, particularly for recurrent events occurring over time, they are widely used for health economic evaluation.

Once the structure of the model is specified (e.g. the number of health states), we need to define the time horizon (e.g. 15 years) and then estimate the transition probabilities. The time horizon (also called the follow-up period) can often be modelled by discrete cycles (e.g. months, years). This "virtual" follow-up period can be determined depending on the specific disease or outcome of interest examined and we would expect that if the follow-up period is long enough (e.g. 100 years), all the individuals coming into the model would end up in the absorbing state of death. After defining all these quantities, we can use the MM to simulate the likely outcomes
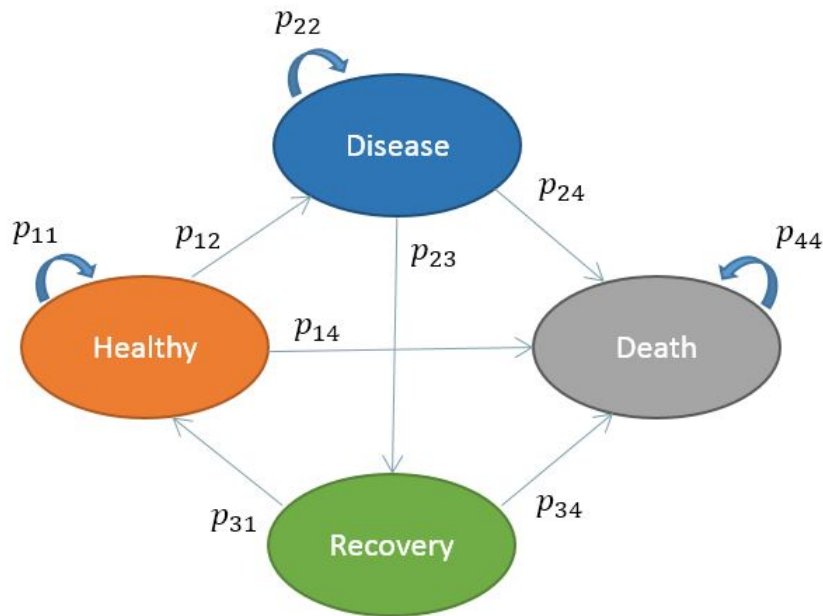
Figure 5.3: The structure of a basic Markov model involving four possible states. The circle represents each possible health state. Arrows indicate that individuals can move between the two states from the one where the arrow originates to the one where it ends. If an arrow connecting the two states is absent, then a transition is not possible. The blue thick arrows represent that individuals remain in the original state without transiting to another state in the model. The state of death is considered to be an absorbing state such that individual moving from all the states to there remain forever. The transition probabilities between the states are indicated as "p_{ij}" where $i, j = 1, 2, 3, 4$. Since death is an absorbing state, p_{44} is equal to 1.

for a "virtual" cohort over a period of time. The outcomes (e.g. the number of individuals in the disease state at endpoint) can then be summarised from these simulations. Note that it is possible to evaluate the impact of some given interventions. For example, suppose that a new drug is available on the market which largely increases the probability of recovery (this is equivalent to the assumption that p_{23} in Figure 5.3 suddenly increases). Moreover, by attaching costs and quality of life data to each state in the model, we can also perform economic evaluation and utility analysis.

In reality, uncertainties are often expected because sometimes it is difficult to collect experimental data to inform some of the parameters, e.g. the transition probabilities from PCOS to diabetes for menopausal women. For these parameters, assumptions based on limited evidence and/or clinical expert consultation are typically applied. A Bayesian approach can formally incorporate existing evidence from several sources to construct a probabilistic MM (Baio, 2012).

By using a simulation method, it is possible to characterise the uncertainties about all the parameters and further propagate these uncertainties to the MM. In fact, the Bayesian procedure produces a large number of simulations from the joint distribution of all the parameters. These can be used to simulate for a bunch of potential "futures" with respect to population dynamics as determined by the transitions between either two states of the MM.

Suppose that for a given discrete time MM, there are S states and the simulations are performed for K cycles. We use $\mathbf{m}_k = (m_{1k}, m_{2k}, \ldots, m_{Sk})$ to represent the total number of patients who enter into each state, where the element m_{sk} included is defined as the number of patients in state s at time k. Let P_k denote the matrix that includes all the transition probabilities between states over the discrete time (i.e. $k \in N = \{0, 1, 2, \ldots\}$). Then the element $p_{ss'k}$ arranged in P_k is defined as the probability for a patient to transit from state s to s' between the consecutive cycles $k-1$ and k. The total number of patients at a given time point for a given state can then be computed as follows:

$$\mathbf{m}_k \quad = \quad \mathbf{m}_{k-1} \times P_k.$$

In other words, the total number of patients distributed over the states at cycle k equals the total number of patients at cycle $k-1$ times the transition probabilities between cycles $k-1$ and k.

The associated costs and utilities for each state at a given time point can be computed by the multiplication of unit cost/utility and the absolute number of patients in that state. For example, assuming that the cost of treatments for an individual patient who remains in State 1 is constant over time (represented by c_1), then the overall cost for patients in this health state at time k is estimated as $m_{1k} \times c_1$. The utilities attached to each state at a given time can be estimated in the same way.

Note that the time horizon of a MM is often large (e.g. 25 years) whereas people are likely to value more for the benefits that can be received closer to the present time compared with those that they will get in the later future. As a consequence, it is essential to apply discounting when measuring the costs and utilities in the future (Torgerson and Raftery, 1999). Discounting differentiates outcomes that occur at different time point and accounts for people's time preference by devaluing the costs and benefits that will be seen in the future. It is generally accepted that if the time horizon for an economic evaluation is greater than 1 year, then a discounting factor

$(1 + d)^k$ needs to be applied for the costs or other outcomes measured where d is the discount rate and k is the time when the costs/outcomes occur (Drummond et al., 2015). It may also be of interest to compute the present value (PV) of costs or other outcomes that occur over the entire time horizon examined. For example, the PV of the total costs over the follow-up period examined for a MM can be specified as

$$\text{PV} = \sum_{k=0}^{K}\sum_{s=1}^{S} \frac{c_k \times m_{sk}}{(1 + d)^k},$$

where c_k represents the costs of the treatment at time k and K is defined as the number of time points included in the analysis. This is equivalent to summing up the discounted costs at each time point for all the states included in the model. Similarly, this formula also applies for the calculation of the PV of benefits. Note that NICE suggests a discount rate of 3.5% (i.e. $d = 0.035$) for all the costs and outcome measurements (National Institute for Clinical Excellence, 2004).

To conclude, the general steps required to conduct a health economic evaluation using the MM include:

- Specify the model structure including the number of possible states through the disease progression;

- Define the time horizon of the model;

- Construct the matrix including transition probabilities between either two of the possible states at each time point evaluated;

- Define the cohort matrix which includes the number of patients in each possible state at each time point evaluated for the purpose of simulation;

- Run the MM model (multiplication of the cohort matrix and matrix of transition probabilities) to get a large set of simulated "futures";

- Summarise results from simulations and record the economic and utility outcomes associated with each state at each time point;

- Discounting should be applied if the time horizon of the MM exceeds one year.

5.3 Aim

The objective of this chapter was to firstly estimate the prevalence of type 2 diabetes among PCOS population in the UK over a 25-year follow-up period from 2014 to 2039. This would further allow me to establish the economic burden related to healthcare for PCOS and evaluate the quality of life for the entire PCOS population in this country over the period examined.

5.4 Modelling approach

5.4.1 Overview of the model

For PCOS, it is assumed that there are four possible states for the disease progression (as shown in Figure 5.4) and the descriptions for each state are described below:

- State 1 (*Probable PCOS*): Women with PCOS in this state are assumed to have developed PCOS but remain clinically undiagnosed. From the previous chapters, we have seen that there is a large number of women with symptoms indicative of PCOS but have not received a confirmed diagnosis.

- State 2 (*Diagnosed PCOS*): Women with PCOS in State 2 have been clinically diagnosed and this population consists of prevalent cases and incidence cases (newly emergent cases each year).

- State 3 (*PCOS with diabetes*): This state includes PCOS patients who have developed diabetes. The number of patients in this state would increase if the accumulative number of patients who gradually develop diabetes (coming from both State 1 and 2) exceeds those who have died (transit into State 4).

- State 4 (*Death*): This state includes individuals who are dead from all states.

The assumptions for this proposed model include:

- Each year, a certain proportion of probable cases convert into diagnosed cases with a transition probability p_{12}. Women who remain undiagnosed after the age of 45 are assumed to be unable to receive a diagnosis for this syndrome since the most prominent external presentation (i.e. menses) disappears after reproductive age.
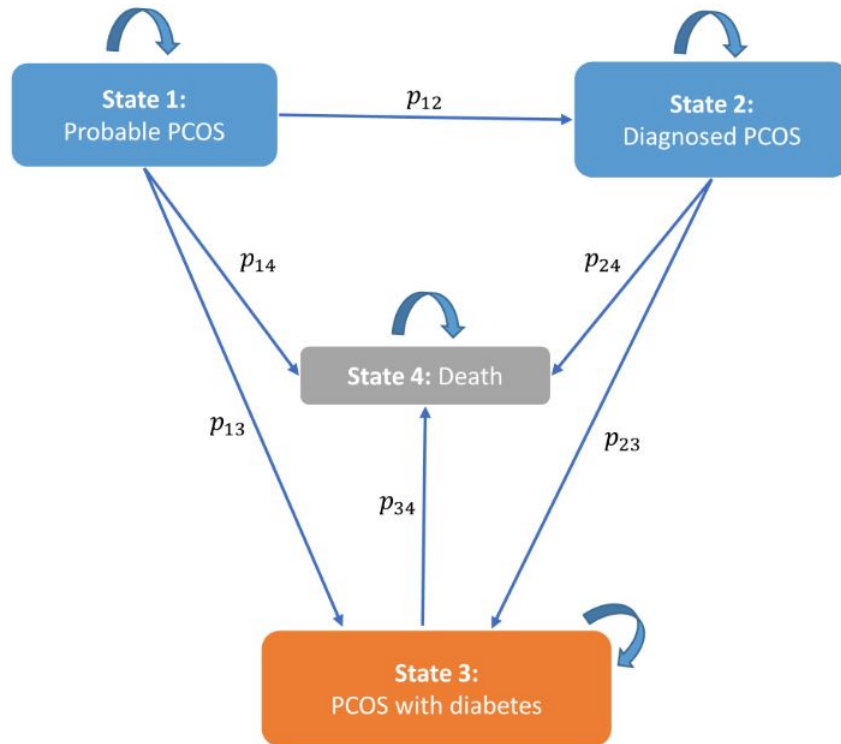
Figure 5.4: The structure of Markov model for modelling the population dynamics of PCOS. It is assumed that there are four possible states for the disease progression of PCOS: *Probable PCOS* (State 1), *Diagnosed PCOS* (State 2), *PCOS with diabetes* (State 3) and *Death* (State 4). Rounded rectangles represent a single health state. Arrows indicate that women can move across two states (from the one at which the arrow originates, to the one where it ends). If the arrow is absent between the two states, then the assumption is that a particular transition is not possible. The state of death is an absorbing state and individuals from any state can move there and stay forever.

- The total number of cases in State 2 in a calendar year is then assumed to be equal to the number of newly emergent cases plus the prevalent cases (defined as cases that have already been diagnosed in the previous year).

- For women who have developed PCOS (either probable or diagnosed cases), some proportion will transit into State 3.

- Once a patient with PCOS have developed diabetes, she is assumed to remain in State 3 until death and can never recover given the nature of diabetes.

- PCOS itself does not increase mortality but mortality for PCOS patients who have devel-

126

oped diabetes differs significantly from that for individuals in other states. This is a strong assumption considering that cardiovascular diseases (CVD) can largely increase the mortality rates and PCOS patients are found to at elevated risk of CVD compared with the general population as suggested in Chapter 2. However, a review by Legro et al. pointed out that while existing data suggested that PCOS is likely to adversely affect the cardiovascular risk profile, its impact is limited for pre-menopausal age group (Legro, 2003). This is supported by the findings from a study conducted in the UK where the CHD-related mortality among 786 women with PCOS (evaluated after 30 years of diagnosis) was not found to be significantly higher than the national rates (Pierpoint et al., 1998). Therefore, the current analysis only considered diabetes as the main cause of elevated mortality for PCOS population.

The list of parameters in Figure 5.4 with their definition and sources of 'real-world' evidence are presented in Table 5.1.

Table 5.1: Model inputs

Parameter	Definition	Source
p_{12}	Transition probability from probable PCOS to diagnosed PCOS	Estimated from THIN
p_{13}	Transition probability from probable PCOS to diabetes	Estimated from THIN
p_{24}	Transition probability from diagnosed PCOS to diabetes	Estimated from THIN and published literature
p_{14}, p_{24}	Transition probabilities from probable and diagnosed PCOS to death	Assumption based (assume that PCOS does not increase mortality and therefore is estimated using rates in general population)
p_{34}	Transition probability from diabetes to death	Based on mortality rates in published studies

5.4.2 Estimating the transition probabilities

I started by estimating the transition probabilities between different disease states (State 1, 2 and 3) in the Markov model for PCOS cohort. A Poisson regression model fitted using a Bayesian approach and its hierarchical version (to account for practice-level variability) were used to obtain the incidence rates. Recall that in Chapter 4, the Frequentist method was used to model the incidence rates of PCOS and in this Chapter, the model was fully expanded to

a Bayesian version. This is particularly relevant as the Bayesian framework allows to easily propagate uncertainty in the transition probability to the ultimate estimates through the MM. This enables a better characterisation and quantification of the burden of the disease. The incidence rates were converted into transition probabilities. The formula for conversion and its associated assumptions will be discussed later on.

Age group specific incidence rates were estimated due to the fact that: (i) results in the previous chapter suggested that younger women are more likely to be diagnosed and most women with PCOS received a diagnosis before the age of 35; (ii) age is a factor strongly associated with the onset of type 2 diabetes (Sharma et al., 2016).

The standard Poisson regression model used to estimate the age group specific incidence rates was specified below:

$$y_i \quad \sim \quad \text{Poisson}(\lambda_i) \tag{5.1}$$

$$\log(\lambda_i) \quad = \quad \beta_0 + \boldsymbol{X}_i\boldsymbol{\beta} + \log(t_i), \tag{5.2}$$

where y_i is the event indicator representing a Poisson process for individual patient during the entire follow-up period. The parameter β_0 is the baseline incidence rate and $\boldsymbol{\beta}$ is the incremental effects of other variables on the baseline rate. For the current analysis, age was included as a categorical variable with 6 levels and β_0 is the incidence rate for individuals aged 40-44. The parameters $\beta_1, \beta_2, \ldots, \beta_5$ represent the effect of age on the incidence rates in age group 15-19, 20-24, 25-29, 30-34, 35-39 compared with the reference group. Note that t_i is the time of follow-up and was included as a log offset in the linear predictor. This model was applied to estimate the parameters λ_{12}, λ_{23} and λ_{13}.

For the estimation of λ_{12} and λ_{13} (incidence rates for probable cases to receive a confirmed diagnosis of PCOS and diabetes, respectively), minimally informative priors were used, e.g. $\beta_0 \sim \text{Normal}(0, 100)$. This is due to the fact that no studies have been identified that provide some insights into such transition rates in women with symptoms indicative of PCOS. Therefore, little is known (even after discussion with the clinician in the team) to allow me to make strong assumptions for statistical inference. In terms of estimating λ_{23} (incidence rates for diagnosed cases to convert to diabetes), different versions of prior were attempted due to the availability of published studies. This is likely to help minimise the impact of bias (e.g. referral bias,

incompleteness of patient information, recording errors of the database, etc.) that may exist in the data I used and the published studies.

Three scenarios were considered.

- Scenario 1: minimally informative priors for baseline incidence rate and incidence rate ratio in comparison with other age categories were included:

$$\beta_0, \beta_1, ..., \beta_5 \quad \sim \quad \text{Normal}(0, 100).$$

As all the parameters are on a log scale, the prior $\text{Normal}(0, 100)$ is very vague. In this scenario, no formal knowledge on the range of rates was included and consequently, the results would be driven only by the data.

- Scenario 2: informative priors for baseline incidence rate and incidence rate ratio were included based on external evidence from published literature. Firstly I referred to the incidence of diabetes in the general female population aged 40-44, denoted by β_{pop} on the log scale. It is suggested that the incidence rate for this age group is 3 per 1000 person-year (PY) (Sharma et al., 2016). This can be translated into a mean on the log scale of $\ln(0.003) = -5.809$. Thus the parameter β_{pop} was modelled using a Normal distribution centred at -5.809 with some variability that reflect the uncertainties of this estimate:

$$\beta_{pop} \quad \sim \quad \text{Normal}(-5.809, 0.01).$$

Figure 5.5 is a graphical representation of the distributional assumption made for the parameter β_{pop}. We can see that the mean is centred at 3 per 1000 PY with a small 95% interval of (2.95, 3.05). This is consistent with what is reported in the study by Sharma et al. (2016).

The relative risk of diabetes comparing PCOS patients and the general population was then defined and represented by ρ. Therefore, the incidence of diabetes in PCOS patients in the reference group (defined as β_0) can be calculated as.

$$\beta_0 \quad = \quad \beta_{pop} \times \rho.$$
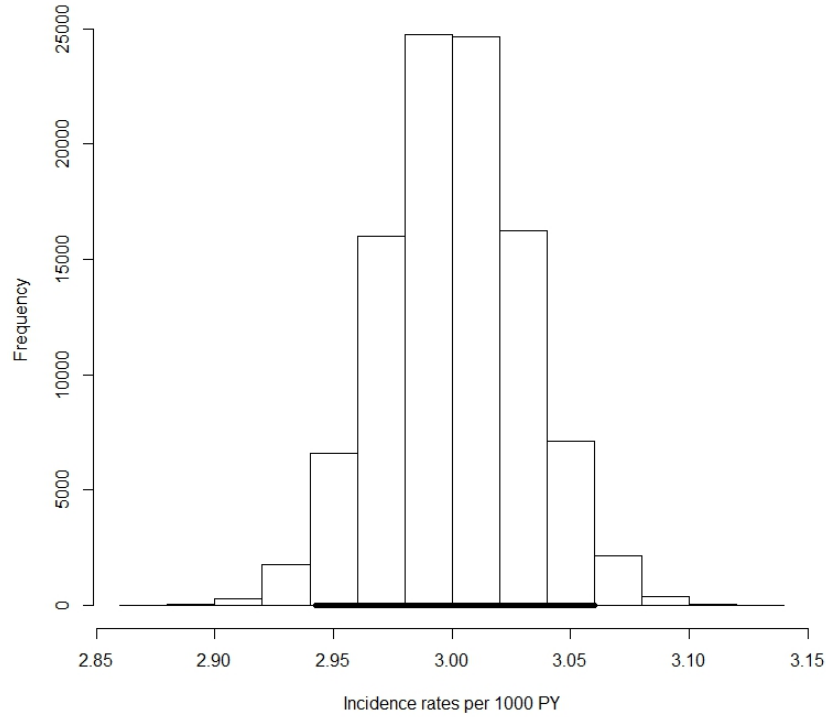
Figure 5.5: Graphical representation of the prior distribution of β_{pop} on the natural scale, which reflects the assumption that the incidence of diabetes in the general female population aged 40-44 is 3 per 1000 PY with a small variance (the horizonal bar in bold represents the 95% interval: 2.95-3.05 per 1000 PY). This reflects the estimate in the published study by Sharma et al. (2016).

Note that there is evidence that the relative risk of developing diabetes for women with PCOS are 3.02 (95%: 2.73-3.33) compared with the general female population (Morgan et al., 2012). Therefore, the incidence of diabetes in the PCOS patients aged 40-44 is expected to be around 9 per 1000 PY (3 per 1000 PY times the relative risk of 3), with some variability (95% interval: 8.1-10.2 per 1000 PY). After converting this rate to the log scale, it corresponds to a value of -4.71 with a 95% interval of $(-4.82$ to $-4.59)$. In other words, the product of β_{pop} and ρ (represented by β_0) is expected to follow a distribution shown as below:

$$\beta_{pop} \times \rho \;=\; \beta_0 \sim \text{Normal}(-4.71, 0.06).$$

The resulting distribution of β_0 on the natural scale is presented in Figure 5.6. We can see that the mean is centred at 9 per 1000 PY. The 95% interval also largely corresponds to the rates obtained in the previous calculation.
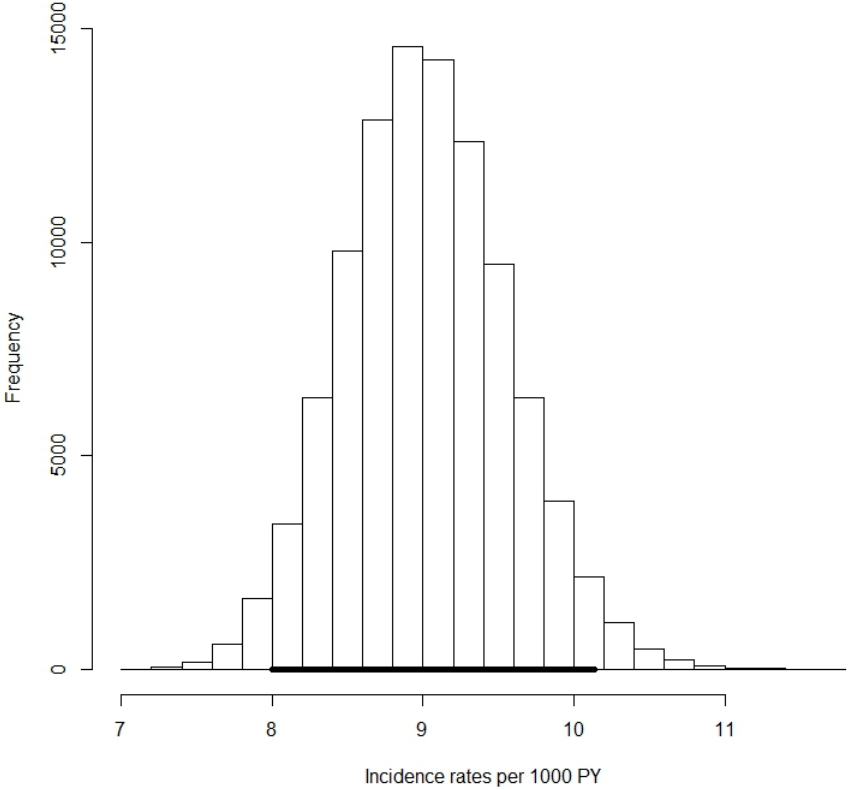
Figure 5.6: Graphical representation of the prior distribution of β_0 on the natural scale, which reflects the assumption that the incidence of diabetes in the PCOS population aged 40-44 is 9 per 1000 PY with a small variance (the horizonal bar in bold represents the 95% interval: 8.1-10.2 per 1000 PY). This is produced based on estimates in the published studies.

I further referred to the relative risk of diabetes (represented by $\beta_1, \beta_2, ..., \beta_5$ on the log scale) across different age groups compared with the reference age group in the study by Sharma et al. (2016). The prior distributions for $\beta_1, \beta_2, ..., \beta_5$ were included to encode the assumption that, compared with the reference group, women aged 15-19, 20-29 and 30-39 are 0.09, 0.37 and 0.63 times less likely to develop diabetes. These relative risks were converted to the log scale. For each parameter, the log mean is consistent with the estimate on the natural scale and a relatively small variance was included to reflect the

uncertainties of these estimates as reported by Sharma et al. (2016):

$$\beta_1 \sim \text{Normal}(-2.408, 0.001)$$

$$\beta_2, \beta_3 \sim \text{Normal}(-0.994, 0.001)$$

$$\beta_4, \beta_5 \sim \text{Normal}(-0.462, 0.001).$$

- Scenario 3: similarly, the priors included in this model were informed by external evidence but in a slightly different manner. A case-control matched study using the General Practice Research Data (GPRD) in the UK was identified that examined the incidence rate of type 2 diabetes in PCOS population from 1990 to 2010 (Morgan et al., 2012). The structure of GPRD is similar to that of THIN. For example, GPRD also collects primary care data across practices in the UK with a higher coverage (e.g. covering 7% of the total population in the UK) and 50-60% of the practices contributing data to THIN also provide data to GPRD. Note that the two databases also share the same coding system. Therefore, results from this study were considered to be relevant to inform the model I used and help to limit bias that may exist both in the THIN data and the previous study.

Equation 5.2 suggests that β_0 is the baseline incidence rate (on a log scale) for the reference group (i.e. women aged 40-44) and $\beta_1, \beta_2, ..., \beta_5$ represent the relative risk of diabetes across age groups compared with the baseline group. Therefore, the incidence rate on the natural scale for women in the comparison groups can be computed as $\exp(\beta_0 + \beta_h)$ where $h = 1, 2, 3, 4, 5$ represent women in age group 15-19, 20-24, 25-29, 30-34 and 35-39, respectively. The overall incidence rate β_{prev} then can be calculated by proportional weighting using the following formula:

$$\exp(\beta_0 + \beta_1)w_1 + \exp(\beta_0 + \beta_2)w_2 + ... + \exp(\beta_0 + \beta_5)w_5 + \exp(\beta_0)w_6 = \exp(\beta_{prev}),$$

where $w_1, w_2, ..., w_5$ are the proportion of women in each age group estimated from the THIN cohort, e.g. the percentage of women aged 15-19 in the study population included from THIN. Consequently, the assumption here is that the age distribution of the cohort from THIN is the same as that of the GPRD study. This was considered to be reasonable since the source of data and the structure of GPRD is quite similar to the THIN database.

By rearranging the above formula, the following equation was obtained that can be used to compute the baseline rate:

$$\beta_0 = \beta_{prev} - \log\Big\{\sum_{n=1}^{5}[w_h \times \exp(\beta_h)] + w_6\Big\}.$$

The model was completed by including informative prior distributions to encode assumptions that reflect estimates in the published studies. For example, the overall rate of diabetes (represented by β_{prev} and is on log scale) was estimated to be 5.7 per 1000 PY in the PCOS population in the GPRD study, with a 95% interval of (5.17, 6.29). This was converted to the log scale:

$$\beta_{prev} \sim \text{Normal}(-5.167, 0.05).$$

The same assumptions in Scenario 2 were still applied for women in the remaining age groups: women aged 15-19, 20-29 and 30-39 are 0.09, 0.37 and 0.63, respectively, less likely to develop diabetes compared with women aged 40-44. This is represented by the following distributions on the log scale:

$$\beta_1 \sim \text{Normal}(-2.408, 0.001)$$
$$\beta_2, \beta_3 \sim \text{Normal}(-0.994, 0.001)$$
$$\beta_4, \beta_5 \sim \text{Normal}(-0.462, 0.001).$$

It should be noted that for Scenario 2 and 3, the variances of all the prior distributions included are small and are consistent with the variability of the estimates provided in the published studies by Sharma et al. (2016) and Morgan et al. (2012). In addition, sensitivity analyses were performed to assess the impact of increasing the variance of some of the parameters (e.g. increase the variance from 0.001 to 1 for relative risk of diabetes across different age groups).

Apart from the above standard model, a multilevel Poisson regression model was also fitted to account for the variances that may exist across different practices. The hierarchical version

of the model can be specified as below:

$$y_{ij} \sim \text{Poisson}(\lambda_{ij})$$
$$\log(\lambda_{ij}) = \beta_0 + \boldsymbol{\beta X}_{ij} + \log(t_{ij}) + u_j,$$

where y_{ij} was defined as the event indicator representing a Poisson process for the ith individual in the jth practice; t_{ij} is the time of follow-up for the ith individual in the jth practice; u_j is the random components of the model indicating variability between practices and the following prior distribution was included for this parameter:

$$u_j \sim \text{Normal}(0, \sigma^2)$$
$$\sigma \sim \text{Uniform}(0, 10).$$

This prior is quite vague and the underlying assumption is that the difference in the incidence rates across practices in the UK is centred at zero whereas is associated with large variability.

The priors included for the regression coefficients (i.e. $\boldsymbol{\beta}$) are exactly the same as those discussed previously for the standard Poisson regression model because the purpose for this hierarchical model was to assess whether there is variation in the incidence rates across practices.

The model convergence was assessed using the Gelman-Rubin statistics and the statistics were provided for all the model parameters. The Gelman-Rubin statistic is designed specifically to assess the output of MCMC algorithm by looking at differences between multiple Markov chains (Gelman and Rubin, 1992). By comparing the estimated between- and within-chains variances for each of the model parameters, the convergence can be examined and large differences between these two variances would indicate issue of non-convergence. Note that there is a numerical cut-off value of 1.1 for Gelman-Rubin statistic, values below which indicate the Monte Carlo Markov model procedure converges to the target posterior distribution.

The incidence rates estimated from the Poisson regression model (and its hierarchical version) were then converted to annual transition probabilities using the following formula (Gidwani, 2014):

$$p_{ij} = 1 - e^{-\lambda_{ij}}, i, j = 1, 2, 3.$$

The underlying assumption is that the incidence rate is constant for a given age group over a Markov cycle of 1 year. This was considered reasonable as PCOS and type 2 diabetes are chronic disease and therefore, the population-level incidence rate is unlikely to change drastically over a short period of time such as 1 year.

Mortality rates in the general population were referred to UK life table (Office for National Statistics, 2016) whereas those in the diabetic patients were based on empirical studies on diabetes, i.e. female patients with type 2 diabetes are 2.13 times more likely to die compared with the general female population (Mulnier et al., 2006). These rates were directly applied as the transition probabilities from disease states to death (p_{14}, p_{24} and p_{34}) without using any modelling approach.

Note that a healthy cohort was simulated to work as the comparison group for the PCOS cohort with respect to the reduction in quality of life over the follow-up period (which will be discussed later in Section 5.4.5). Similarly, a Markov model with three health states (*healthy*, *diabetes* and *death*) was specified and the transition probabilities between either two of the three states were estimated. The incidence rates of diabetes in the general population suggested by Sharma et al. (2016) were used to compute transition probabilities between the healthy and diabetic state. Likewise, mortality rates in the general population and the diabetic patients were referred to UK life table and Mulnier et al. (2006), respectively.

5.4.3 Cohort simulation

A virtual PCOS cohort estimated using the prevalence rates in 2014 (refer to Chapter 4) and UK census data from the Office of National Statistics (Office for National Statistics, 2014) were used to simulate the population dynamics of PCOS from 2014 to 2039. The transition probabilities estimated in the previous section were applied. Two scenarios were considered:

- Closed cohort model: women aged 15-44 were included and the number of patients who start in each state at the baseline year (2014) was assumed to be consistent with previous estimates from THIN, i.e. 18.2% of the PCOS patients had a prior diagnosis of type 2 diabetes and therefore, these women start in State 3 (PCOS with diabetes). Another assumption is that there is no death for the entire cohort simulated at the baseline year. Given that a limited number of deaths occur in women aged below 45 (i.e. mortality rates smaller than 0.13%), this assumption is not expected to have massive impact on the

135

overall simulation.

- Open cohort model: while others remain the same as the closed cohort model, women aged 1-14 at the baseline year were also simulated in a way such that this cohort enters the study population gradually when they reach the age of 15. The number of probable and diagnosed cases was calculated based on the prevalence rates estimated from THIN for the age group 15-19. Therefore, women aged 1-14 were assumed to be diagnosed at the same rates as the youngest age group once they are eligible to be included in the study population.

Likewise, for the virtual healthy cohort, both closed and open cohort model were simulated. Some of these women start in diabetic state and the number of women with diabetes was assumed to be consistent with the report by Diabetes UK (2015) where age-specific prevalence of diabetes is provided.

5.4.4 Economic analysis

In this section, the methods used to evaluate the healthcare-related economic burden of PCOS is introduced. This is mainly based on summing up the costs of a range of possible treatments to PCOS and care for patients who have developed diabetes. The treatments selected for the analysis were those that tailored to the major symptoms of PCOS (i.e. menstrual dysfunction, obesity, pre-diabetes, acne), with a proportion of patients receiving the treatments greater than 2% after the diagnosis of PCOS (refer to Table 4.6). Otherwise, the treatments were considered to be rarely prescribed to the patients and therefore, are likely to contribute only a small amount of costs to the overall burden of the condition. The total costs of an individual treatment were computed as the multiplication of the unit price, the daily dose, treatment duration and the proportion of patients who require that treatment. The recommended dose and duration of treatment for each drug considered are displayed in Table 5.2. The unit price of treatments was referred to the British National Formulary (BNF, 2014), conference abstracts and published reports by the NHS. The prior distributions included for the calculation of drugs costs are presented in Table 5.3. The range of drug dose for each treatment considered forms the boundary of the prior distribution, e.g. suppose that m represents the daily dose for metformin and consequently, m would follow a uniform distribution: $m \sim \text{Uniform}(1828, 1914)$. The total

costs of metformin can then be computed as £0.19 (the unit price) times m and the duration of the treatment as well as the total number of patients who require this treatment.

Table 5.2: The recommended dose and treatment duration for each drug considered.

Treatment	Dose and instruction
Oral contraceptives	One tablet daily for 21 days and repeat for each menstrual cycle until menopause
Metformin	500-1000mg daily for the first week and then 1000-1500mg daily for the second week and 1500-2000mg daily if tolerated thereafter for 3-6 month
Eflornithine	Apply twice daily and should be discontinued in the absence of improvement after treatment for 4 months.
Weight loss/control drugs	Orlistat: 120mg for maximum 3 times daily and continue treatment beyond 12 weeks only if weight loss since start of treatment exceeds 5%.
Acne drugs	Most are topical cream or gel,including erythromycin, benzoyl peroxide, tretinoin and isotretinoin. Apply 1-2 times daily and review at 8 weeks. Treatment may take up to 6 months or beyond depending on severity.

Table 5.3: The distributional assumptions of mean dose and duration of treatment for drug considered.

Item	Mean	Distribution	Note/source
Drug dose			
Oral contraceptives	1 tablet daily	Constant	BNF
Metformin	1828-1914mg daily	Uniform	Daily average for 3-6 months (BNF)
Eflornithine	4 cream tubes	Constant	Expert interviews
Orlistat	120-360mg daily	Uniform	Daily average for 3 months (BNF)
Acne drugs	0.5-2g daily	Uniform	Daily average for 2-6 months (National Health Service, 2015)
Drug costs			
Oral contraceptives	£124.33 per year	Constant	3-year average costs applied (Praet and D'Oca, 2014)
Metformin	£0.19 per gram	Constant	BNF
Eflornithine	£56.87 per cream tube	Constant	BNF
Orlistat	£3.14 per gram	Constant	BNF
Acne drugs	£0.047-0.398 per gram	Uniform	National Health Service (2015)
Diabetic care	£6400 per year	Constant	Published study (Hex et al., 2012)

The cumulative proportion of patients who receive a relevant treatment within one year after their diagnosis was estimated from THIN using the methods introduced in Chapter 3. The

137

overall healthcare burden was discounted at a rate of 3.5% as suggested by the National Institute for health and Clinical Excellence (NICE) (National Institute for Clinical Excellence, 2004).

5.4.5 Utility analysis

Quality of life (QoL) data from published studies were used for this analysis. It is assumed that the QoL of both probable and diagnosed cases are the same based on estimates reported by Coffey et al. (2006). The assumption is likely to be reasonable since many probable cases may be true cases but failed to receive a confirmed diagnosis due to the imperfection of the current screening system for this condition. The QoL of diabetic patients were used to approximate that of PCOS patients with diabetes because it was not able to identify any study reporting relevant information for this specific cohort. Although EQ-5D is recommended by the NICE to facilitate comparisons across economic evaluations (National Institute for Clinical Excellence, 2004), I only identified studies that report the QoL estimates for PCOS patients based on the 36-Item short survey form (SF-36). Therefore, the methods introduced by Ara and Brazier (2008) was applied as the conversion method between the two score measurements. In a nutshell, this article introduces a method to derive the mean EQ-5D score of a cohort using mean statistics of the eight dimension scores of SF-36 of the same cohort. Specifically speaking, the EQ-5D score is regressed on the eight dimension scores of the SF-36 form (e.g. physical function, mental health, social function) to obtain coefficients for each dimensional considered. Although the estimation is based on individual level patient data (sample data) but predictive abilities are assessed using the summary statistics from both subgroups in the sample data and some published studies. The regression coefficients (for each health component examined by SF-36) obtained from the model in this study was applied for conversion in the current analysis.

As mentioned previously, the population dynamics of a healthy cohort of the same population size as the PCOS cohort was also simulated. The purpose was to examine the reduction in quality of life (QoL) comparing the PCOS cohort and the healthy cohort over the follow-up period. This time, the MM only includes 3 states: *healthy* (State 1), *diabetes* (State 2), *death* (State 3). Individuals start in State 1 can transit to State 2 and 3 and once individuals develop diabetes, they were assumed to remain in State 2 or die (transit into State 3). I referred to QoL suggested by Kind et al. (1999) and the QoL for diabetic patients was still applied for healthy women who develop diabetes later in life.

A discounting rate of 3.5% was also used for all the utility analyses.

5.5 Results

5.5.1 Characteristics of patient cohort selected from THIN

A total number of 14,135 eligible patients with PCOS selected from THIN were used for estimating the incidence rates from one disease state to another. Among these patients, 776 (5.5%) had a diagnosis of diabetes before the date when the diagnosis of PCOS was recorded. The recording of an HbA1c test was reported in 2.4% of PCOS patients within one year after they received their diagnosis of PCOS, which increased to slightly above 4% after 3 years of diagnosis.

5.5.2 Incidence rates estimated from the Poisson model

The incidence rates estimated from both the standard and multilevel Poisson regression model are presented in Table 5.4. In general, the rates are smaller for the hierarchical model (bottom line in the table) and associated with a narrower 95% credible interval (CrI) compared with the standard model (top line in the table). The plots for rate comparison are presented in Figure 5.7 and 5.8.

Table 5.4: Incidence rates estimated from the Poisson model using Bayesian approach. The top line in each row represents results from the standard model and the bottom line represents those from the hierarchical model.

Age, years	Incidence rates, per 1000 person years (95% Credible Intervel, CrI)				
	Probable PCOS to diagnosed PCOS	Probable PCOS to diabetes	Diagnosed PCOS to diabetes		
			Scenario 1	Scenario 2	Scenario 3
15-19	121 (82-163)	26 (9-28)	27 (12-41)	3 (3-3)	3 (2-3)
	94 (63-132)	17 (6-35)	18 (9-31)	2 (1-2)	2 (1-2)
20-24	58 (45-70)	11 (3-13)	22 (15-26)	13 (12-14)	10 (9-12)
	48 (37-61)	5 (3-9)	15 (11-21)	6 (5-8)	7 (6-9)
25-29	24 (19-28)	8 (3-8)	23 (18-25)	13 (12-15)	11 (9-13)
	18 (14-22)	4 (2-6)	16 (12-20)	7 (5-8)	7 (6-9)
30-34	17 (12-18)	6 (3-7)	30 (25-33)	22 (20-24)	18 (15-20)
	12 (9-15)	4 (2-6)	22 (18-26)	11 (9-13)	12 (10-14)
35-39	10 (7-11)	8 (5-10)	28 (22-32)	22 (19-24)	17 (14-19)
	7 (5-10)	5 (3-8)	20 (16-25)	11 (9-13)	11 (9-13)
40-44	6 (3-7)	7 (3-8)	32 (25-36)	33 (30-36)	26 (24-29)
	3 (2-5)	4 (2-6)	24 (19-29)	17 (14-20)	17 (14-21)

The incidence rates from probable to diagnosed PCOS obtained from both models suggest that younger women more likely to receive a confirmed diagnosis of PCOS compared with older women. For example, the rates for women aged 15-19 are as high as 121 per 1000 PY whereas are below 20 per 1000 PY for women aged 30 and above as suggested by the standard model. In terms of probable cases converting to diabetes, both of the models suggest that the younger women are associated with higher rates in contrast to older women. The estimated incidence rates of diagnosed cases converting to diabetes are presented by scenarios (based on different assumptions as discussed in the methods section). When minimally informative prior was included (Scenario 1), the rates remain constant across different age groups for both of the models. However, when informative prior (Scenario 2 and 3) was included, variation arises in the rates across age groups and the rates increase from young to older age groups, i.e. from 3 per 1000 PY in women aged 15-19 to around 30 per 1000 PY for women aged 40-44 for the standard model. The results obtained from inflating the variances (from 0.001 to 1) in the prior distributions for the relative risk of diabetes (across different age groups) are generally in line with Scenario 1 with slightly lower rates for the younger age groups compared with Scenario 1. For example, the incidence rates were estimated to be 19 per 1000 PY (95% CrI: 9-33 per 1000 PY) for the youngest age group whereas older women (aged 30 and above) are associated with higher rates at around 28 per 1000 PY (95% CrI: 22-33 per 1000 PY). The model convergence was reached for all the analyses and the plots of diagnostic statistics are presented in Appendix G.
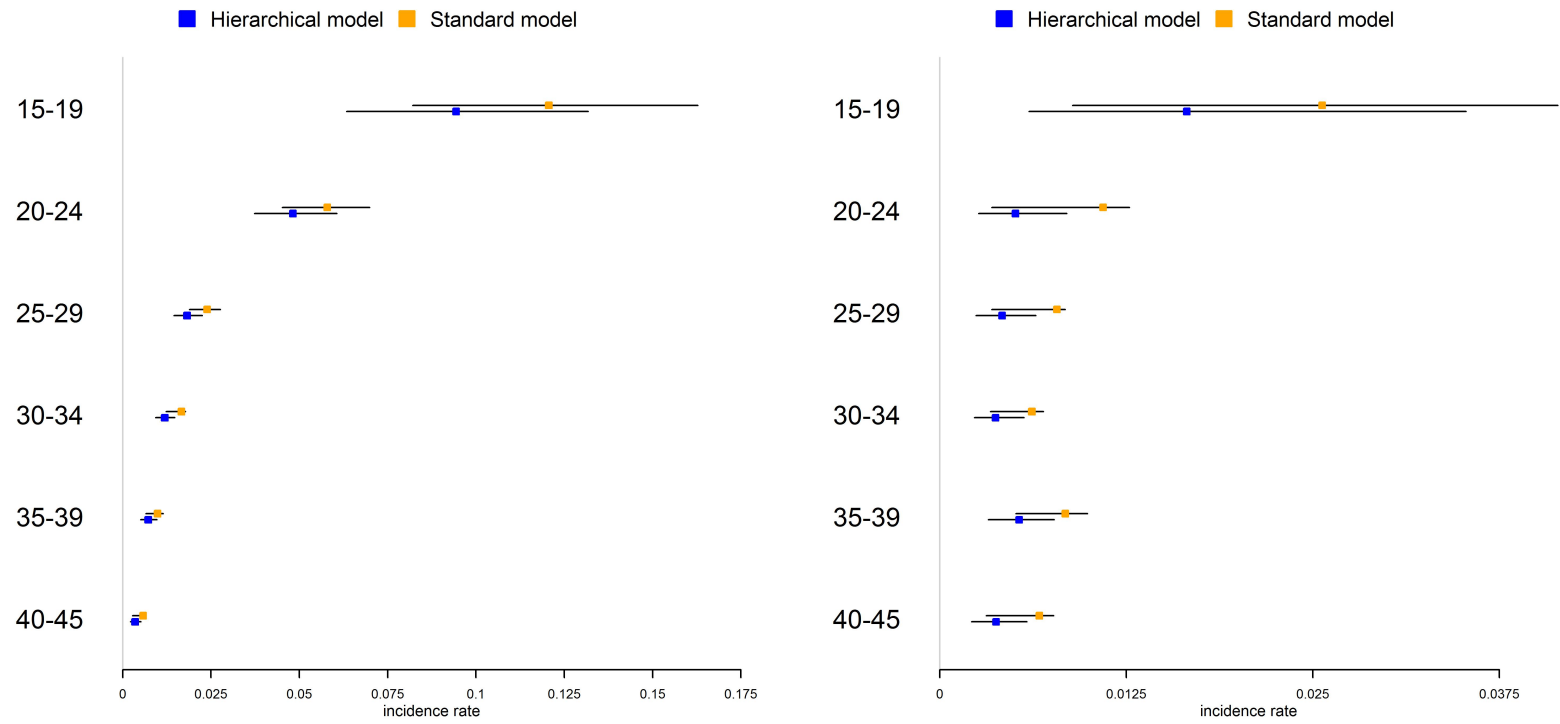
Figure 5.7: The plots are for model comparison in terms of age-specific incidence rates estimated under different scenarios: (i) transition from probable to diagnosed PCOS (left plot); (ii) transition from probable PCOS to diabetes (right plot).
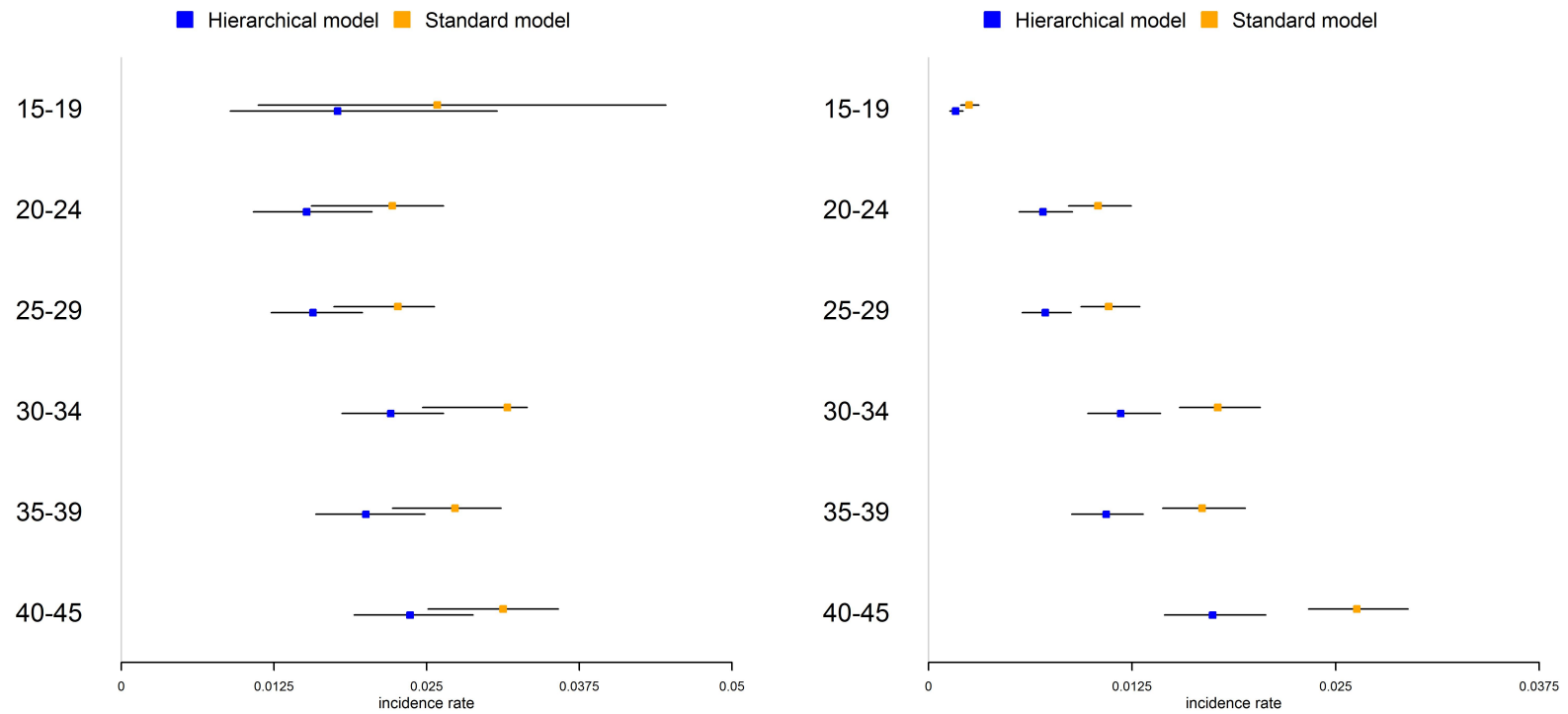
Figure 5.8: The plots are for model comparison in terms of age-specific incidence rates estimated under different scenarios: (i) transition from diagnosed PCOS to diabetes by including minimally informative prior in the model (left plot); (ii) transition from diagnosed PCOS to diabetes by including informative prior (Scenario 3) in the model (right plot).

5.5.3 Results from population dynamics simulation

The age distribution of the PCOS population estimated from the UK census data is presented in the Table 5.5. There is a higher proportion of women in the older age groups (i.e. age above 25) compared with the younger age groups.

Table 5.5: Age distribution of PCOS population in the UK for baseline year (2014). The number of patients who aged 1-14 in 2014 was estimated based on the assumption that they will be diagnosed at the same rate as women aged 15-19 once they reach the age of 15.

Age in 2014, years	Population size (%)
1-4	9,270 (2.91)
5-9	11,021 (3.46)
10-14	10,056 (3.16)
15-19	10,957 (3.44)
20-24	40,166 (12.62)
25-29	69,061 (21.70)
30-34	69,873 (21.96)
35-39	52,801 (16.59)
40-44	45,029 (14,15)

The projected number of patients ending up in each state from 2014 and onwards at a consecutive 5-year interval is presented in Table 5.6. It should be noted that the results in this table are based on the most conservative estimates in Table 5.4 (Scenario 3). Results from both closed cohort and open cohort model are presented. The figures on the top and bottom line in each row represents the number of patients estimated from the standard Poisson regression model and the hierarchical version of the model, respectively.

For the standard model, at the baseline year, 12.3% of the PCOS patients had a diagnosis of diabetes. The closed cohort model indicates that proportion of PCOS patients who are likely to convert to type 2 diabetes is 26.3% (95% Credible Interval, CrI: 25.2-27.6%) after 25 years of follow-up whereas the corresponding proportion estimated from the open cohort model is 26.5% (95% CrI: 25.4-27.8%). The rainbow plot (Figure 5.5.3) graphically displays the distribution of patients over states (same legend for both plots). It should be noted that these results are generally robust even when different transition probabilities from diagnosed PCOS to diabetes (Scenario 1, 2 and 3) were applied. For example, for the base case scenario, the proportion of patients who end up in the diabetic state by the end of follow-up was estimated to be 30.4% (95% interval: 29.2-31.9%).

The results obtained from the hierarchical model are slightly smaller but do not differ much

from those using the standard model. For example, the closed cohort model estimates that 22.0% (95% CrI: 20.7-23.5%) of the patients are likely to end up in the diabetic states after the follow-up period and corresponding estimates from the open cohort model are 22.1% (95% CrI: 20.8-23.7%).

The results from simulating the healthy cohort suggest that the prevalence of type 2 diabetes in this population increases from 1.5% in the baseline year 2014 to 8.3% by the end of follow-up.

Table 5.6: The predicted distribution of cases over states in the follow-up period (estimates for both closed and open cohort model presented). The figures on the top line in each row were simulated based on estimates from standard Poisson regression model and those on the bottom line were simulated using estimates from the hierarchical version of the model.

| Year | Probable PCOS | | Diagnosed PCOS | | PCOS with diabetes | | Death | |
	Closed	Open	Closed	Open	Closed	Open	Closed	Open
2014	139,659	139,659	112,609	112,609	35,620	35,620	0	0
	139,659	139,659	112,609	112,609	35,620	35,620	0	0
2019	122,628	126,588	115,750	120,318	48,704	50,229	806	809
	127,235	132,126	115,411	119,098	44,442	45,916	800	803
2024	114,236	120,830	112,206	123,060	59,805	63,423	1,640	1,653
	120,143	129,542	114,062	122,293	52,064	55,500	1,619	1,630
2029	109,599	117,135	107,267	124,128	68,529	74,450	2,494	2,520
	116,192	128,340	111,135	123,846	58,111	63,573	2,450	2,475
2034	107,356	112,744	103,220	120,975	73,957	81,111	3,356	3,403
	114,212	124,390	108,501	122,319	61,889	68,194	3,286	3,330
2039	106,486	111,096	101,073	118,232	76,110	84,617	4,220	4,287
	113,378	122,380	107,023	121,152	63,364	70,515	4,122	4,186
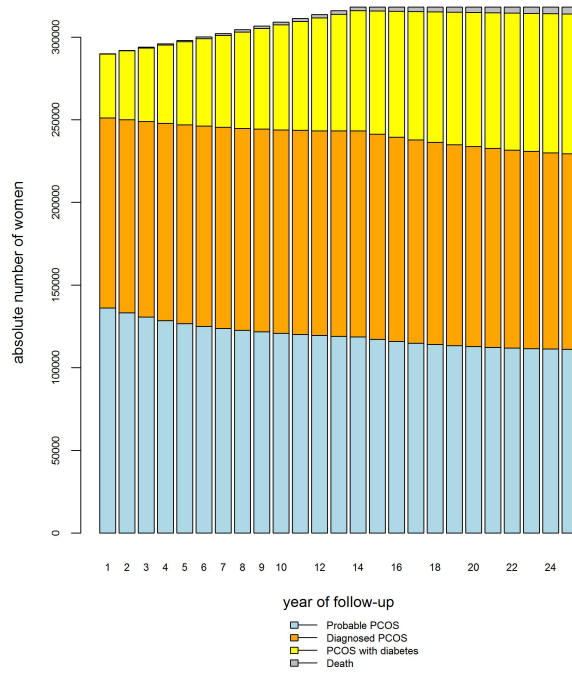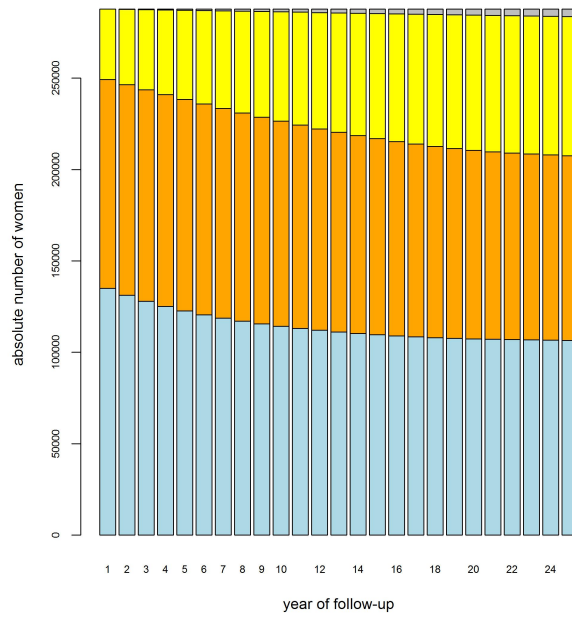
Figure 5.9: Distribution of women over states in the follow-up: (i) closed cohort model (on the top); (ii) open cohort model (on the bottom). The graphs are based on simulations using estimates from the standard model.

5.5.4 Pharmaceutical outcomes

The cumulative proportion of patients who receive a relevant prescription within one year after their diagnosis is shown in Table 5.7. A higher proportion of diagnosed cases received oral contraceptives (either combined or progestogen, 23.4%), metformin (17.8%) and eflornithine (4.6%) compared with probable cases. In contrast, a notable higher proportion of probable cases received acne drugs (24.9%), which more than doubles that of diagnosed cases. This is consistent with results shown in Table 4.6.

Table 5.7: Cumulative proportion of cases receiving relevant prescription for PCOS by case definition (between 2012 and 2014).

Treatment	Percentage of patients receiving prescription	
	Probable cases	Diagnosed cases
Combined oral contraceptives	12.8%	17.5%
Progestogen oral contraceptives	4.8%	5.9%
Metformin	1.7%	17.8%
Eflornithine	1.8%	4.6%
Weight loss/control drugs	0.63%	2.3%
Acne drugs	24.9%	11.1%

5.5.5 Economic burden

The economic burden of PCOS was estimated to be £237 million (95% CrI: £237-238 million) in 2014 and it continues to increase, with a present value of overall economic burden of over £7 billion (95% CrI: £6.8-7.3 billion, these are conservative estimates from the standard model). The overall burden of the disease estimated from the hierarchical model is slightly lower (£6.2 billion). Figure 5.10 presents the trend of the country level healthcare burden of PCOS over the 25-year follow-up (results from standard model). The average costs for an individual patient per year were estimated to be within a range of £723 and £950 during the 25-year follow-up (e.g. £723 in 2014 and increase to a discounted value of £950 in 2023). This corresponds to an overall average costs per patient on an annual basis of £876 for the entire follow-up period. A significant proportion (over 96%) of healthcare burden for PCOS population is attributed to treating diabetes.
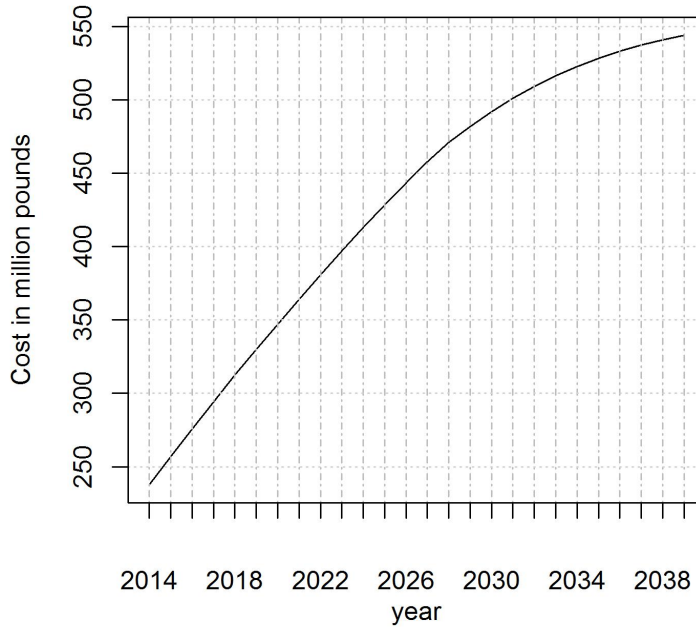
Figure 5.10: The estimated healthcare-related economic burden of PCOS over the follow-up period (based on simulations using estimates from the standard model).

5.5.6 Quality of life

The quality of life (QoL) measured by EQ-5D for women with PCOS (both diagnosed and probable cases) was estimated to be at a mean score of 0.76. The EQ-5D score is 0.7 (SD 0.3) for PCOS patients with type 2 diabetes. The results from the model suggested that over the period of follow-up, there is a reduction in QoL from 0.75 (95% CrI: 0.67-0.79) to 0.31 (95% CrI: 0.24-0.34) for the entire PCOS population in the UK (similar results were obtained from the hierarchical model). This is significantly lower compared with the results simulated for the healthy cohort, as represented by the shaded area in grey formed by the two lines in different colours in Figure 5.11. The magnitude of quality deficits is within a range of 0.04-0.1 over the follow-up period. Note that since discounting was applied in both PCOS and healthy cohort, the quality deficits represent the loss of quality of life measured by EQ-5D score for PCOS cohort compared with the healthy population. The results presented here were obtained from the open cohort model and similar results were obtained from the closed cohort model.
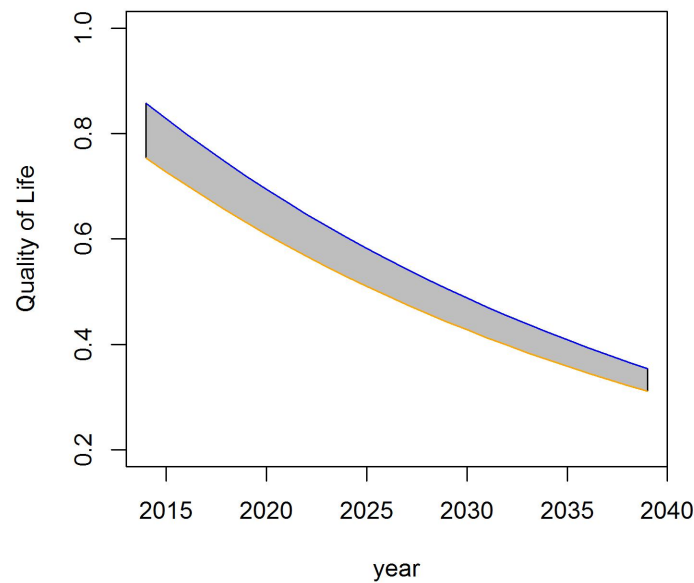
Figure 5.11: Quality of life (QoL) simulated for the PCOS cohort (yellow line) and a healthy cohort (blue line) in the UK over the follow-up period (2014-2039). The shaded area in grey representes the loss of QoL of PCOS cohort compared with the healthy cohort. Results are based on simulations using estimates from standard model (open cohort scenario).

5.6 Discussion

5.6.1 Summary of main findings

In this chapter, I modelled the population dynamics of PCOS, the healthcare-related costs and the quality of life for the entire patient cohort in the UK. The indication from this simulation-based analysis is that approximately 26% of the PCOS patients are likely to develop type 2 diabetes over the next 25 years starting from 2014. This remains robust for transition rates estimated based on different assumptions. The quality of life of PCOS patients is generally lower compared with the general population throughout the follow-up period examined. The costs to the NHS for this condition were estimated to be at least £237 million on an annual basis.

5.6.2 Interpretation

The results from the Poisson regression model suggested that only a very small proportion of potential cases (the incidence rate is as high as 120 per 1000 PY for the youngest age group) is expected to receive a confirmed diagnosis. This points out the issue of lack of public awareness of this syndrome as indicated by previous chapters and also reflects the nature of primary care recording: GPs are likely to record symptoms rather than making a confirmed diagnosis when they are less certain about patients' condition; the recording of some of the patients who are referred to secondary care may be lost if care is not transferred back to the GPs. However, high conversion rates from PCOS to diabetes were estimated from our model. The failure of screening (to make a confirmed diagnosis of PCOS) and the rapid conversion from PCOS to diabetes result in the fact that over a quarter of the PCOS population was projected to develop diabetes by the end of the follow-up. This is likely to incur large amount of costs to the NHS (an estimated present value of the overall economic burden of at least £6.2 billions) and significantly impact the quality of life for individual patients (the magnitude of quality deficits ranges from 0.04 to 0.1 compared with the general population). All these urge the need for a more efficient screening approach for PCOS given that an earlier diagnosis would allow interventions to reduce the impact of the condition in this cohort (approximately three quarters of a million women in the UK).

Screening with an HbA1c test has been found cost effective for other high risk populations at diabetes, as the test is inexpensive (£4.04) (National Institute for Clinical Excellence, 2011)

and facilitates intervention to prevent diabetes. The term HbA1c is used to refer to glycated haemoglobin, which develops when haemoglobin (a protein in the red blood cells that carries oxygen throughout one's body) joins with glucose in the blood and becomes 'glycated'. By measuring glycated haemoglobin (HbA1c), the clinicians will be able to know the average blood sugar level over a period of time. However, although an HbA1c test is recommended by the guideline of PCOS (Legro et al., 2013), this is hardly followed at present: only 2.4% of the PCOS patients eligible to be included from the THIN database for analysis had a recording of HbA1c test within one year after the date of their diagnosis of PCOS, which increases to slightly over 4% after 3 years of diagnosis (recall results in Section 5.5.1). Note that the HbA1c test is already incorporated in the NHS health check and the National Diabetes Prevention Programme (National Health Service). These programmes target those aged 40-74 with BMI⩾30 (or BMI⩾27.5 for South Asian population) in whom the incidence of diabetes ranges from 14.3 to 23.8 per 1000 PY (Tillin et al., 2015). Therefore, apart from improved screening for PCOS, our finding of conversion rates of diabetes in women with PCOS also suggests a possible role for screening PCOS patients for diabetes.

It is possible that expanding the current targeted population of these programmes to include women with PCOS or symptoms indicative of this condition may be cost effective. For example, this may help improve the detection of symptoms indicative of diabetes (e.g. pre-diabetes as suggested by results from an HbA1c test ranging from 42 to 47 mmol/mol) and consequently, more patients with PCOS can benefit from early intervention to prevent rapid conversion to diabetes. From a country perspective, this is likely to produce massive amount of savings in healthcare costs associated with PCOS.

5.6.3 Strength and limitation

This is the first study that applied modelling approach to simulate the population dynamics of PCOS using the real world evidence. Modelling is often considered to be cost-effective to explore what-if scenarios and allow data from multiple sources to partially inform the model to examine more complicated research questions, which is unlikely to be achieved by an individual study.

The individual patient data (IPD) from THIN were used to estimate some of the transition probabilities between possible states included in the Markov model. As the IPD are from routine clinical practice, the estimated transition rates are likely to reflect the true conversion between

the disease states. However, referral bias may exist which meant that women with severe PCOS (e.g. obesity, prediabetes) are more likely to visit general practitioners (GPs) and receive a diagnosis consequently. For these women, the duration between the consecutive consultations also tends to be shorter. All these factors may introduce bias into our estimates. I attempted to address this issue by using the Bayesian approach to allow the IPD to be complemented by external evidence from other comparable observational studies conducted in similar context (Morgan et al., 2012). While there may be bias in both our data and published studies, the formal machinery of the Bayesian approach is likely to make a compromise and provide more realistic estimates based on more sources of information.

The Poisson regression model and its hierarchical version were applied to estimate the transition probabilities and the results obtained from the two models largely agree with each other. The estimates from the hierarchical model are generally more stable (with smaller variances) due to the nature of the hierarchical structure. Specifically speaking, individual estimates that are based on less information can borrow strengths from other estimates to be pooled closer to the group mean and this is called the "shrinkage" property (Gelman et al., 2014). Both models controlled for the age variable and the hierarchical version of the model further accounted for the variability across practices.

Nevertheless, other factors may need to be considered in the Poisson regression model. There is evidence that obesity may be a factor affecting the transition from PCOS to diabetes (Legro et al., 1999; Zhao and Qiao, 2013). For example, obesity can worsen PCOS and thus result in a faster conversion to diabetes (Sam, 2007). However, since part of the risk of developing PCOS is obesity, it is not possible to separate the effect of obesity and PCOS on the higher conversion rates from PCOS to diabetes. The previous chapters also suggested that the prevalence of PCOS varies across different ethnic groups and a study by Zhao and Qiao (2013) indicated that Southeast Asian women with PCOS are at higher risk of diabetes compared with PCOS patients with other ethnic background. Therefore, higher conversion rates may be expected for certain ethnic groups and this would further lead to higher prevalence of diabetes in patients with certain ethnic backgrounds by the end of follow-up. Nonetheless, this may not have strong influence on our estimates given that over 85% of the population in UK are White Caucasian (Office for National Statistics, 2011) and THIN is considered to be a true reflection of the mixture of various ethnic groups across different regions of the country. Another reason for not including obesity

status and ethnicity into the current model is the incompleteness of data on these variables from THIN database, i.e. the information of weight/BMI and ethnicity were only recorded in less than 40% and less than 50% of the patients, respectively. The proportion of patients with both weight/BMI and ethnicity information reported would be even lower ($\leq$40%). Further analysis considering these aspects needs to be warranted.

A further concern regarding the methods applied in this piece of analysis is that modelling is often a simplification of the complex situation in reality. The current Markov model only included four states however it is possible that other states may exist. For example, a state of PCOS with pre-diabetes may be reasonable given that pre-diabetes is an independent state often considered in modelling the population dynamics of diabetes (Thurecht et al., 2007; Bernier, 2012). Note that the state of resolved PCOS was not included due to the fact that PCOS is believed to be generically inherited and is largely incurable (Bernier, 2012).

As for the economic analysis, the list of drugs considered as the major burden of care for PCOS is supported by a recent European survey where a group of specialists in PCOS were interviewed. The results from this survey suggested that obesity and diabetes were the major concern by over 60% of the interviewees. Metformin was selected as the top line treatment by a third of the interviewees, followed by lifestyle modification (25%) and oral contraceptives (22%) (Conway et al., 2014).

However, the estimate for the overall healthcare costs of PCOS in the UK was considered to be conservative due to several reasons. The primary reason is that the current analysis did not include the costs associated with other comorbidities of PCOS such as cardiovascular diseases (CVD), obstetrical complications and infertility due to lack of supporting evidence. For example, Azziz et al. (2005) suggested that the annual cost of infertility care for PCOS patients in the US is $533 million each year. While this may contribute a certain proportion to the overall disease burden, it is less plausible to apply strong assumptions for the economic evaluation of infertility given the fact that not all women desire pregnancy. Some studies reported that women with PCOS are at elevated risks of CVD compared with the general population (Wild et al., 2000; Dahlgren et al., 1992). Nevertheless, the review by Legro (2003) argued that the premenopausal age group did not appear to be at higher risk of presenting clinical signs of CVD. As a consequence, limited data with respect to these aspects are currently available to draw definite conclusions. It is also suggested that women with PCOS are more likely to experience

depression (Barry et al., 2011) and antidepressants and counselling or cognitive behavioural therapy for psychological disorders may be costly. Besides, it is often the case that prescribing rates in primary care are under-reported due to various reasons (e.g. patient records are not transferred back from the secondary care). I attempted to address this issue by extracting the first drug record within 12 months following a diagnosis of PCOS and this period was considered to be relatively long to examine the treatments initiated in the secondary care, which made it less likely to bias the estimate for overall healthcare costs.

Chapter 6

SUMMARY AND CONCLUSIONS

6.1 Chapter summary

This thesis explored the prevalence of PCOS from both a global and country-specific perspective and established the healthcare related economic burden of the disease in the UK. It addressed these objectives coherently through both empirical studies and individual level patient data from primary care database in the UK, using a wide range of suitable methodologies for each of the specific research questions. Each piece of analysis mirrored the aims and objectives of the thesis, as represented by the flow of the chapters. This chapter starts by drawing together the main findings from previous chapters. I then discuss the methodological strengths and weaknesses and finally present the implications of the work in this thesis for practice and further research.

6.2 Main findings

There are several major findings from the analyses in this thesis. Firstly, there is a large gap between the prevalence rates estimated from community studies and database studies (which is often based on primary care database or medical insurance database), highlighting the issue of under-diagnosis of PCOS. The under-diagnosis could be due to the wide variation in the clinical presentations of the condition across individual patients and different populations, resulting in difficulties for the clinicians to make a confirmed diagnosis. Another possible reason is the lack of awareness of this syndrome by both the public and the healthcare workers given that the major criteria of PCOS were established relatively recently. For example, the first criterion

of PCOS (the NIH criterion) was established in 1990 and it may take years for clinicians to start following the guideline for diagnosis. This is reflected by the limited number of studies conducted in 1990s and the low recording rates in early 21st century.

The prevalence of PCOS varies across different ethnic groups and even for the same ethnic group, there is significant variation in the prevalence estimates under different diagnostic standards. This raised the concern for the accuracy of current measurement techniques and the applicability of current diagnostic criteria to populations with different ethnic backgrounds. Empirical studies suggested that women with PCOS are at significantly higher risk of a wide range of morbidities such as type 2 diabetes and infertility. Therefore, the failure of screening may lead to serious health consequences for PCOS patients, which severely affect the daily activities and quality of life for women with PCOS. On the other hand, while under-diagnosis remains a big issue, over-diagnosis of the condition may exert negative psychological impact on women with PCOS, which can worsen the major symptoms of PCOS such as menstrual dysfunction and thus increase the chance for potential cases to be qualified as true cases. This further emphasises the importance of diagnostic accuracy of this condition.

The prescribing patterns of PCOS in the primary care supported the previous argument because a wide range of treatments tend to be initiated for patients following a diagnosis of PCOS. This also reflects the fact that currently, there is lack of most effective treatment for PCOS and patients generally receive treatments tailored to the external symptoms they present with.

Given there is limited supporting evidence on the causal relationship between a range of cardiometabolic diseases other than type 2 diabetes provided by the long-term follow-up studies, a full economic evaluation by including costs of care for all the morbidities potentially induced by PCOS was challenged. The current economic analysis estimated the costs of treating external symptoms presented by patients and the care required by those who have converted to type 2 diabetes over a course of 25-year follow-up. The results suggested that the disease burden associated with care for PCOS population in the UK is significant and the costs of treating type 2 diabetes account for an overwhelming proportion of the overall burden.

6.3   Methodological strengths and weaknesses

The work presented in this thesis had a number of strengths and some acknowledgeable weaknesses. Firstly, I systematically reviewed the published literature to collect evidence relevant

155

to the main research questions and at the same time, to identify the aspects of disease that had not been well explored (e.g. prevalence of PCOS in populations with different ethnic backgrounds). Alongside the data collection process, the generalised models for evidence synthesis using Bayesian approach were developed through the creation of an R package. The Bayesian methods allow the observed data to be complemented by some formal representation of prior belief based on empirical evidence or pilot studies when the available data are limited. The model templates provided by the package can be applied to a wide range of disease areas and therefore, are considered to be an important methodological highlight derived from the first part of this thesis. However, it should be noted that the current models designed for evidence synthesis fitted to observational studies or randomised trials can only account for confounding factors at study level, which is likely to be inadequate. The heterogeneity in the way that information is provided by individual studies also largely limited our ability to include some potential factors. For example, the body mass index of patients are often observed to be provided in different ways, e.g. mean BMI for the study population, range of BMI for subgroups. The number of confounding factors to be considered in meta-regression models is limited because model convergence may be an issue when a large number of parameters are included.

Apart from literature review, I also conducted analysis using real-world evidence from a primary care database in the UK, which largely formed the second part of this thesis. The database analysis is advantageous in terms of sample size and is considered to be broadly representative for the UK population. As the routine clinical data are collected, the results reflected the reality in primary care. However, due to the nature of data recording in 'real' world setting of UK primary care, under-reporting is expected because the general practitioners are likely to record only the main diagnosis in the coded patient records but unlikely to code all the features related to a diagnosis or all the other patient information. Moreover, due to the issue of missing data, complete information cannot be obtained on variables such as ethnicity or drug indication, which may potentially influence the epidemiological and pharmaceutical outcomes in this analysis.

Based on the aggregate data identified from empirical studies and the individual level patient data extracted from THIN, I evaluated the economic burden and quality of life for PCOS population in the UK over a relatively long period of follow-up by using a simulation-based method. The modelling approach allowed me to explore what-if scenarios and synthesise evidence from multiple sources to inform the model for examining more complicated research questions, which

is unlikely to be achieved by an individual study. Nonetheless, modelling is almost always a simplification of the complex real world situation and many factors were not able to be considered at the moment due to lack of supporting evidence, e.g. the costs of treating infertility and cardiovascular diseases. This may lead to an underestimation of the true healthcare burden of the disease for this country.

6.4   Implications of the findings

The findings in this thesis have significant public health impact and important implications for practice and further research. The work presented is quite novel and add much to the current state of science for this disease area. For example, by the time when the literature review was conducted, there had been no studies attempting to investigate the prevalence of PCOS in different ethnic groups by combining information from single observational studies using a meta-analysis approach. Similarly, modelling approach was firstly introduced to simulate the population dynamics of PCOS based on evidence from multiple sources. The specific implications with respect to each of the main findings are discussed below.

Firstly, the significant difference in the prevalence rates estimated using medical recordings and through systematic screening in designed studies suggested that the public awareness of PCOS still remains to be improved. Health campaigns can be a potential way to allow the public and general practitioners to learn more aspects of PCOS. Early screening may also be recommended given the relatively low costs of care for PCOS itself compared with the high costs of diabetic care required after disease progression. However, considering the frequent issue of under- and over-diagnosis of PCOS, further work is still expected to enhance the accuracy of current screening approach for PCOS. This could possibly be achieved through the establishment of a more detailed ethnicity-specific diagnostic criteria, as suggested by the large variation in prevalence rates and metabolic features of the condition across different ethnic groups.

The results from the current evaluation for the prevalence of type 2 diabetes in the PCOS cohort indicated that PCOS or possibly menstrual dysfunction (major symptom of PCOS) may be a marker for the detection of type 2 diabetes as a high proportion of women with PCOS was estimated to develop diabetes by the end of the 25-year follow-up. Therefore, the current targeted population of the Diabetes Prevention Programme may be recommended to be expanded to include women with PCOS or symptoms suggestive of PCOS. This is likely

to massively improve the detection of symptoms indicative of diabetes in PCOS patients and prevent rapid conversion to diabetes for this selected population. In this way, significant amount of healthcare costs associated with PCOS can be saved for the NHS in the UK.

The wide variation in the prescribing patterns of PCOS in the primary care suggested the issue of lack of most effective treatment for this condition currently, urging the need for further research to explore this area. While there are many trials comparing the efficacy of different treatments on ameliorating the metabolic presentations of PCOS patients, longitudinal studies are recommended to investigate the long-term effect of various treatments in terms of reducing the risk of conversion from PCOS to diabetes as well as the improvement in quality of life for individual patients. Data from this type of longitudinal studies are expected for conducting further cost-effectiveness analysis based on the current model structure, e.g. the public health impact of improved screening offered to PCOS patients. It is expected that if more effective screening is available, a larger number of potential cases is likely to receive a confirmed diagnosis and proper care thereafter. Although this increases the number of diagnosed cases, the actual number of cases transiting into diabetic state tend to be reduced given that appropriate treatments are initiated following a diagnosis. The economic burden can then be re-evaluated to see the effect of the new intervention from a country perspective. Similarly, the public health impact of other interventions or the combination use of a range of interventions can be modelled in the same way.

Bibliography

Adi, E. D. and Tank, P. (2010). Irving stein, michael leventhal and a slice of endocrine history. *Journal of obstetrics and gynaecology of India*, 60(2):121.

Aguiree, F., Brown, A., Cho, N. H., Dahlquist, G., Dodd, S., Dunning, T., Hirst, M., Hwang, C., Magliano, D., Patterson, C., et al. (2013). Idf diabetes atlas.

Al Khaduri, M., Al Farsi, Y., Al Najjar, T. A. A., and Gowri, V. (2014). Hospital-based prevalence of polycystic ovarian syndrome among omani women. *Middle East Fertility Society Journal*, 19(2):135–138.

Alberti, K. G. M., Zimmet, P., and Shaw, J. (2005). The metabolic syndrome-a new worldwide definition. *The Lancet*, 366(9491):1059.

Alberti, K. G. M., Zimmet, P., and Shaw, J. (2007). International diabetes federation: a consensus on type 2 diabetes prevention. *Diabetic Medicine*, 24(5):451–463.

Alberti, K. G. M. M. and Zimmet, P. f. (1998). Definition, diagnosis and classification of diabetes mellitus and its complications. part 1: diagnosis and classification of diabetes mellitus. provisional report of a who consultation. *Diabetic medicine*, 15(7):539–553.

Albu, A., Radian, S., Fica, S., and Barbu, C. G. (2015). Biochemical hyperandrogenism is associated with metabolic syndrome independently of adiposity and insulin resistance in romanian polycystic ovary syndrome patients. *Endocrine*, 48(2):696–704.

Álvarez-Blasco, F., Botella-Carretero, J. I., San Millán, J. L., and Escobar-Morreale, H. F. (2006). Prevalence and characteristics of the polycystic ovary syndrome in overweight and obese women. *Archives of internal medicine*, 166(19):2081–2086.

Alvarez-Blasco, F., Martínez-García, M. Á., Luque-Ramírez, M., Parraza, N., San Millan, J. L., and Escobar-Morreale, H. F. (2009). Role of haptoglobin in polycystic ovary syndrome (pcos), obesity and disorders of glucose tolerance in premenopausal women. *PLoS One*, 4(5):e5606.

Amato, M. C., Verghi, M., Galluzzo, A., and Giordano, C. (2011). The oligomenorrhoic phenotypes of polycystic ovary syndrome are characterized by a high visceral adiposity index: a likely condition of cardiometabolic risk. *Human Reproduction*, page der088.

American Diabetes Association (2010). Diagnosis and classification of diabetes mellitus. *Diabetes care*, 33(Supplement 1):S62–S69.

Anderson, S. A., Barry, J. A., and Hardiman, P. J. (2014). Risk of coronary heart disease and risk of stroke in women with polycystic ovary syndrome: A systematic review and meta-analysis. *International journal of cardiology*, 176(2):486–487.

Apridonidze, T., Essah, P. A., Iuorno, M. J., and Nestler, J. E. (2005). Prevalence and characteristics of the metabolic syndrome in women with polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, 90(4):1929–1935.

Ara, R. and Brazier, J. (2008). Deriving an algorithm to convert the eight mean sf-36 dimension scores into a mean eq-5d preference-based score from published studies (where patient level data are not available). *Value in Health*, 11(7):1131–1143.

Asgharnia, M., Mirblook, F., and Soltani, M. A. (2011). The prevalence of polycystic ovary syndrome (pcos) in high school students in rasht in 2009 according to nih criteria. *International journal of fertility & sterility*, 4(4):156.

Asunción, M., Calvo, R. M., San Millán, J. L., Sancho, J., Avila, S., and Escobar-Morreale, H. F. (2000). A prospective study of the prevalence of the polycystic ovary syndrome in unselected caucasian women from spain 1. *The Journal of Clinical Endocrinology & Metabolism*, 85(7):2434–2438.

Attaoua, R., El Mkadem, S. A., Radian, S., Fica, S., Hanzu, F., Albu, A., Gheorghiu, M., Coculescu, M., and Grigorescu, F. (2008). Fto gene associates to metabolic syndrome in women with polycystic ovary syndrome. *Biochemical and biophysical research communications*, 373(2):230–234.

Attlee, A., Nusralla, A., Eqbal, R., Said, H., Hashim, M., and Obaid, R. S. (2014). Polycystic ovary syndrome in university students: occurrence and associated factors. *International journal of fertility & sterility*, 8(3):261.

Azziz, R., Carmina, E., Dewailly, D., Diamanti-Kandarakis, E., Escobar-Morreale, H. F., Futterweit, W., Janssen, O. E., Legro, R. S., Norman, R. J., Taylor, A. E., et al. (2006). Criteria for defining polycystic ovary syndrome as a predominantly hyperandrogenic syndrome: an androgen excess society guideline. *The Journal of Clinical Endocrinology & Metabolism*, 91(11):4237–4245.

Azziz, R., Marin, C., Hoq, L., Badamgarav, E., and Song, P. (2005). Health care-related economic burden of the polycystic ovary syndrome during the reproductive life span. *The Journal of Clinical Endocrinology & Metabolism*, 90(8):4650–4658.

Azziz, R., Woods, K. S., Reyna, R., Key, T. J., Knochenhauer, E. S., and Yildiz, B. O. (2004). The prevalence and features of the polycystic ovary syndrome in an unselected population. *The Journal of Clinical Endocrinology & Metabolism*, 89(6):2745–2749.

Baio, G. (2012). *Bayesian methods in health economics*. CRC Press.

Balen, A. H., Conway, G., Homburg, R., and Legro, R. (2005). *Polycystic ovary syndrome: a guide to clinical management*. CRC Press.

Barry, J. A., Azizia, M. M., and Hardiman, P. J. (2014). Risk of endometrial, ovarian and breast cancer in women with polycystic ovary syndrome: a systematic review and meta-analysis. *Human reproduction update*, 20(5):748–758.

Barry, J. A., Kuczmierczyk, A. R., and Hardiman, P. J. (2011). Anxiety and depression in polycystic ovary syndrome: a systematic review and meta-analysis. *Human Reproduction*, 26(9):2442–2451.

Barthelmess, E. K. and Naz, R. K. (2014). Polycystic ovary syndrome: current status and future perspective. *Frontiers in bioscience (Elite edition)*, 6:104.

Berger, J. O. (1990). Robust bayesian analysis: sensitivity to the prior. *Journal of statistical planning and inference*, 25(3):303–328.

Bernier, D. (2012). Polycystic ovary syndrome: Pathogenesis, health consequences, and treatment of pcos in relation to insulin resistance.

Bhattacharya, S. M. and Jha, A. (2011). Prevalence and risk of metabolic syndrome in adolescent indian girls with polycystic ovary syndrome using the 2009 joint interim criteria. *Journal of Obstetrics and Gynaecology research*, 37(10):1303–1307.

Bloom, M. S., Schisterman, E. F., and Hediger, M. L. (2006). Selecting controls is not selecting normals: design and analysis issues for studying the etiology of polycystic ovary syndrome. *Fertility and sterility*, 86(1):1–12.

BNF (2014). *British National Formulary.* https://www.bnf.org/.

Bode, D., Seehusen, D. A., and Baird, D. (2012). Hirsutism in women. *American family physician*, 85(4):373–380.

Borenstein, M., Hedges, L. V., Higgins, J., and Rothstein, H. R. (2009). *Introduction to Meta-analysis.* Wiley Online Library.

Boudreaux, M. Y., Talbott, E. O., Kip, K. E., Brooks, M. M., and Witchel, S. F. (2006). Risk of t2dm and impaired fasting glucose among pcos subjects: results of an 8-year follow-up. *Current diabetes reports*, 6(1):77–83.

Boyle, J. A., Cunningham, J., O'Dea, K., Dunbar, T., and Norman, R. J. (2012). Prevalence of polycystic ovary syndrome in a sample of indigenous women in darwin, australia. *Med J Aust*, 196(1):62–6.

Brooks, S., Gelman, A., Jones, G., and Meng, X.-L. (2011). *Handbook of markov chain monte carlo.* CRC press.

Caliskan, E., Kilic, T., Bodur, H., and Zeteroglu, S. (2007). The frequency of metabolic syndrome in women with polycystic ovaries at reproductive age and comparison of different diagnostic criteria for metabolic syndrome. *J Turkish German Gynecol Assoc Artemis*, 8:402–407.

Çalışkan, E., Kılıç, T., Bodur, H., and Zeteroğlu, Ş. (2007). The frequency of metabolic syndrome in women with polycystic ovaries at reproductive age and comparison of different diagnostic criteria for metabolic syndrome. *Journal of the Turkish-German Gynecological Association*, 8(4).

Carmina, E., Napoli, N., Longo, R., Rini, G., and Lobo, R. (2006). Metabolic syndrome in polycystic ovary syndrome (pcos): lower prevalence in southern italy than in the usa and the influence of criteria for the diagnosis of pcos. *European Journal of Endocrinology*, 154(1):141–145.

Celik, C., Abali, R., Bastu, E., Tasdemir, N., Tasdemir, U. G., and Gul, A. (2013). Assessment of impaired glucose tolerance prevalence with hemoglobin a1c and oral glucose tolerance test in 252 turkish women with polycystic ovary syndrome: a prospective, controlled study. *Human Reproduction*, page det002.

Celik, C., Tasdemir, N., Abali, R., Bastu, E., and Yilmaz, M. (2014). Progression to impaired glucose tolerance or type 2 diabetes mellitus in polycystic ovary syndrome: a controlled follow-up study. *Fertility and sterility*, 101(4):1123–1128.

Chan, W. P. A., Ngo, D. T., Sverdlov, A. L., Rajendran, S., Stafford, I., Heresztyn, T., Chirkov, Y. Y., and Horowitz, J. D. (2013). Premature aging of cardiovascular/platelet function in polycystic ovarian syndrome. *The American journal of medicine*, 126(7):640–e1.

Chen, C.-I., Hsu, M.-I., Lin, S.-H., Chang, Y.-C. I., Hsu, C.-S., and Tzeng, C.-R. (2015). Adiponectin and leptin in overweight/obese and lean women with polycystic ovary syndrome. *Gynecological Endocrinology*, 31(4):264–268.

Chen, X., Yang, D., Mo, Y., Li, L., Chen, Y., and Huang, Y. (2008). Prevalence of polycystic ovary syndrome in unselected women from southern china. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 139(1):59–64.

Cheung, A. P. (2001). Ultrasound and menstrual history in predicting endometrial hyperplasia in polycystic ovary syndrome. *Obstetrics & Gynecology*, 98(2):325–331.

Cheung, L., Ma, R., Lam, P., Lok, I., Haines, C., So, W., Tong, P., Cockram, C., Chow, C., and Goggins, W. (2008). Cardiovascular risks and metabolic syndrome in hong kong chinese women with polycystic ovary syndrome. *Human reproduction*, 23(6):1431–1438.

Chisholm, J. (1990). The read clinical classification. *BMJ: British Medical Journal*, 300(6732):1092.

Christensen, S. B., Black, M. H., Smith, N., Martinez, M. M., Jacobsen, S. J., Porter, A. H., and Koebnick, C. (2013). Prevalence of polycystic ovary syndrome in adolescents. *Fertility and sterility*, 100(2):470–477.

Ciampelli, M., Fulghesu, A., Cucinelli, F., Pavone, V., Caruso, A., Mancuso, S., and Lanzone, A. (1997). Heterogeneity in beta cell activity, hepatic insulin clearance and peripheral insulin sensitivity in women with polycystic ovary syndrome. *Human reproduction (Oxford, England)*, 12(9):1897–1901.

Ciampelli, M., Fulghesu, A. M., Murgia, F., Guido, M., Cucinelli, F., Apa, R., Caruso, A., and Lanzone, A. (1998). Acute insulin response to intravenous glucagon in polycystic ovary syndrome. *Human Reproduction*, 13(4):847–851.

Cibula, D., Cifkova, R., Fanta, M., Poledne, R., Zivny, J., and Skibova, J. (2000). Increased risk of non-insulin dependent diabetes mellitus, arterial hypertension and coronary artery disease in perimenopausal women with a history of the polycystic ovary syndrome. *Human Reproduction*, 15(4):785–789.

Clark, A., Thornley, B., Tomlinson, L., Galletley, C., and Norman, R. (1998). Weight loss in obese infertile women results in improvement in reproductive outcome for all forms of fertility treatment. *Human Reproduction*, 13(6):1502–1505.

Clayton, W., Lipton, M., Elford, J., Rustin, M., and Sherr, L. (2005). A randomized controlled trial of laser treatment among hirsute women with polycystic ovary syndrome. *British Journal of Dermatology*, 152(5):986–992.

Coffey, S., Bano, G., and Mason, H. D. (2006). Health-related quality of life in women with polycystic ovary syndrome: a comparison with the general population using the polycystic ovary syndrome questionnaire (pcosq) and the short form-36 (sf-36). *Gynecological Endocrinology*, 22(2):80–86.

Collaboration, C. et al. (2011). Review manager (revman)[computer program]. *Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration*.

Conway, G., Dewailly, D., Diamanti-Kandarakis, E., Escobar-Morreale, H. F., Franks, S., Gambineri, A., Kelestimur, F., Macut, D., Micic, D., Pasquali, R., et al. (2014). European survey

of diagnosis and management of the polycystic ovary syndrome: results of the ese pcos special interest group's questionnaire. *European Journal of Endocrinology*, 171(4):489–498.

Cook, S., Weitzman, M., Auinger, P., Nguyen, M., and Dietz, W. H. (2003). Prevalence of a metabolic syndrome phenotype in adolescents: findings from the third national health and nutrition examination survey, 1988-1994. *Archives of pediatrics & adolescent medicine*, 157(8):821–827.

Coviello, A. D., Legro, R. S., and Dunaif, A. (2006). Adolescent girls with polycystic ovary syndrome have an increased risk of the metabolic syndrome associated with increasing androgen levels independent of obesity and insulin resistance. *The Journal of Clinical Endocrinology & Metabolism*, 91(2):492–497.

Cussons, A. J., Watts, G. F., Burke, V., Shaw, J. E., Zimmet, P. Z., and Stuckey, B. G. (2008). Cardiometabolic risk in polycystic ovary syndrome: a comparison of different approaches to defining the metabolic syndrome. *Human Reproduction*, 23(10):2352–2358.

Daan, N. M., Louwers, Y. V., Koster, M. P., Eijkemans, M. J., de Rijke, Y. B., Lentjes, E. W., Fauser, B. C., and Laven, J. S. (2014). Cardiovascular and metabolic profiles amongst different polycystic ovary syndrome phenotypes: who is really at risk? *Fertility and sterility*, 102(5):1444–1451.

Dahlgren, E., Janson, P., Johansson, S., Lapidus, L., and Oden, A. (1992). Polycystic ovary syndrome and risk for myocardial infarction: evaluated from a risk factor model based on a prospective population study of women. *Acta obstetricia et gynecologica Scandinavica*, 71(8):599–604.

Dalamaga, M., Papadavid, E., Basios, G., Vaggopoulos, V., Rigopoulos, D., Kassanos, D., and Trakakis, E. (2013). Ovarian saha syndrome is associated with a more insulin-resistant profile and represents an independent risk factor for glucose abnormalities in women with polycystic ovary syndrome: A prospective controlled study. *Journal of the American Academy of Dermatology*, 69(6):922–930.

Dargham, S. R., Ahmed, L., Kilpatrick, E. S., and Atkin, S. L. (2017). The prevalence and metabolic characteristics of polycystic ovary syndrome in the qatari population. *PloS one*, 12(7):e0181467.

Davé, S. and Petersen, I. (2009). Creating medical and drug code lists to identify cases in primary care databases. *Pharmacoepidemiology and drug safety*, 18(8):704–707.

de Ferranti, S. D., Gauvreau, K., Ludwig, D. S., Neufeld, E. J., Newburger, J. W., and Rifai, N. (2004). Prevalence of the metabolic syndrome in american adolescents. *Circulation*, 110(16):2494–2497.

Department for Communities and Local Government (2015). *The English Index of Multiple Deprivation (IMD) 2015 Infographic.* https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/464431/English_Index_of_Multiple_Deprivation_2015_-_Infographic.pdf.

Diabetes UK (2015). *DIABETES: FACTS AND STATS.* https://www.mrc.ac.uk/documents/pdf/diabetes-uk-facts-and-stats-june-2015/.

Diamanti-Kandarakis, E., Kouli, C. R., Bergiele, A. T., Filandra, F. A., Tsianateli, T. C., Spina, G. G., Zapanti, E. D., and Bartzis, M. I. (1999). A survey of the polycystic ovary syndrome in the greek island of lesbos: hormonal and metabolic profile. *The Journal of Clinical Endocrinology & Metabolism*, 84(11):4006–4011.

Diamanti-Kandarakis, E., Piperi, C., Kalofoutis, A., and Creatsas, G. (2005). Increased levels of serum advanced glycation end-products in women with polycystic ovary syndrome. *Clinical endocrinology*, 62(1):37–43.

Dias, S., Sutton, A. J., Welton, N. J., and Ades, A. (2013). Evidence synthesis for decision making 3 heterogeneitysubgroups, meta-regression, bias, and bias-adjustment. *Medical Decision Making*, 33(5):618–640.

Ding, T. and Baio, G. (2016). *Package bmeta.* https://cran.r-project.org/web/packages/bmeta/bmeta.pdf.

Ding, T., Baio, G., Hardiman, P. J., Petersen, I., and Sammon, C. (2016). Diagnosis and management of polycystic ovary syndrome in the uk (2004–2014): a retrospective cohort study. *BMJ open*, 6(7):e012461.

Ding, T., Hardiman, P., Petersen, I., Wang, F., Qu, F., and Baio, G. (2017). The prevalence

of polycystic ovary syndrome in reproductiveaged women of different ethnicity: a systematic review and meta-analysis.

Dokras, A., Bochner, M., Hollinrake, E., Markham, S., VanVoorhis, B., and Jagasia, D. H. (2005). Screening women with polycystic ovary syndrome for metabolic syndrome. *Obstetrics & Gynecology*, 106(1):131–137.

Domecq, J. P., Prutsky, G., Mullan, R. J., Hazem, A., Sundaresh, V., Elamin, M. B., Phung, O. J., Wang, A., Hoeger, K., Pasquali, R., et al. (2013). Lifestyle modification programs in polycystic ovary syndrome: systematic review and meta-analysis. *The Journal of Clinical Endocrinology & Metabolism*, 98(12):4655–4663.

Dos Reis, R., Foss, M., de Moura, M. D., Ferriani, R., and Silva de Sá, M. (1995). Insulin secretion in obese and non-obese women with polycystic ovary syndrome and its relationship with hyperandrogenism. *Gynecological Endocrinology*, 9(1):45–50.

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., and Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes.* Oxford university press.

Dunaif, A., Segal, K. R., Futterweit, W., and Dobrjansky, A. (1989). Profound peripheral insulin resistance, independent of obesity, in polycystic ovary syndrome. *Diabetes*, 38(9):1165–1174.

Dunaif, A., Wu, X., Lee, A., and Diamanti-Kandarakis, E. (2001). Defects in insulin receptor signaling in vivo in the polycystic ovary syndrome (pcos). *American Journal of Physiology-Endocrinology And Metabolism*, 281(2):E392–E399.

Echiburú, B., Pérez-Bravo, F., Maliqueo, M., Sánchez, F., Crisosto, N., and Sir-Petermann, T. (2008). Polymorphism t c (- 34 base pairs) of gene cyp17 promoter in women with polycystic ovary syndrome is associated with increased body weight and insulin resistance: a preliminary study. *Metabolism*, 57(12):1765–1771.

Ehrmann, D. A., Liljenquist, D. R., Kasza, K., Azziz, R., Legro, R. S., and Ghazzi, M. N. (2006). Prevalence and predictors of the metabolic syndrome in women with polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, 91(1):48–53.

Elwyn, G., Edwards, A., Eccles, M., and Rovner, D. (2001). Decision analysis in patient care. *The Lancet*, 358(9281):571–574.

Escobar-Morreale, H. F. and San Millán, J. L. (2007). Abdominal adiposity and the polycystic ovary syndrome. *Trends in Endocrinology & Metabolism*, 18(7):266–272.

Esmaeilzadeh, S., Delavar, M. A., Amiri, M., Khafri, S., and Pasha, N. G. (2014). Polycystic ovary syndrome in iranian adolescents. *International journal of adolescent medicine and health*, 26(4):559–565.

Faloia, E., Canibus, P., Gatti, C., Frezza, F., Santangelo, M., Garrapa, G., and Boscaro, M. (2004). Body composition, fat distribution and metabolic characteristics in lean and obese women with polycystic ovary syndrome. *Journal of endocrinological investigation*, 27(5):424–429.

Faria, F. R. d., Gusmão, L. S., Faria, E. R. d., Gonçalves, V. S. S., Cecon, R. S., Franceschini, S. d. C. C., and Priore, S. E. (2013). Polycystic ovary syndrome and intervening factors in adolescents from 15 to 18 years old. *Revista da Associação Médica Brasileira*, 59(4):341–346.

Fauser, B. C., Tarlatzis, B. C., Rebar, R. W., Legro, R. S., Balen, A. H., Lobo, R., Carmina, E., Chang, J., Yildiz, B. O., Laven, J. S., et al. (2012). Consensus on womens health aspects of polycystic ovary syndrome (pcos): the amsterdam eshre/asrm-sponsored 3rd pcos consensus workshop group. *Fertility and sterility*, 97(1):28–38.

Flyckt, R. L. and Goldberg, J. M. (2011). Laparoscopic ovarian drilling for clomiphene-resistant polycystic ovary syndrome. *Seminars in reproductive medicine*, 29(2):138–146.

Fulghesu, A., Magnini, R., Portoghese, E., Angioni, S., Minerba, L., and Melis, G. B. (2010). Obesity-related lipid profile and altered insulin incretion in adolescents with polycystic ovary syndrome. *Journal of Adolescent Health*, 46(5):474–481.

Gabrielli, L. and Aquino, E. (2012). Polycystic ovary syndrome in salvador, brazil: a prevalence study in primary healthcare. *Reprod Biol Endocrinol*, 10:96.

Gambineri, A., Repaci, A., Patton, L., Grassi, I., Pocognoli, P., Cognigni, G. E., Pasqui, F., Pagotto, U., and Pasquali, R. (2009). Prominent role of low hdl-cholesterol in explaining the high prevalence of the metabolic syndrome in polycystic ovary syndrome. *Nutrition, Metabolism and Cardiovascular Diseases*, 19(11):797–804.

Gamerman, D. and Lopes, H. F. (2006). *Markov chain Monte Carlo: stochastic simulation for Bayesian inference.* CRC Press.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2014). *Bayesian data analysis*, volume 2. Chapman & Hall/CRC Boca Raton, FL, USA.

Gelman, A. et al. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by browne and draper). *Bayesian analysis*, 1(3):515–534.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, pages 457–472.

Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):721–741.

Gidwani, R. (2014). *Deriving Transition Probabilities for Decision Models.* `https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/819-notes.pdf`.

Gilks, W., Richardson, S., and Spiegelhalter, D. (1996). Markov chain monte carlo in practice, ser. interdisciplinary statistics series.

Gill, H., Tiwari, P., and Dabadghao, P. (2012). Prevalence of polycystic ovary syndrome in young women from north india: A community-based study. *Indian journal of endocrinology and metabolism*, 16(Suppl 2):S389.

Gjønnaess, H. (1994). Ovarian electrocautery in the treatment of women with polycystii ovary syndrome (pcos): Factors affecting the results. *Acta obstetricia et gynecologica Scandinavica*, 73(5):407–412.

Glintborg, D., Rubin, K. H., Nybo, M., Abrahamsen, B., and Andersen, M. (2015). Morbidity and medicine prescriptions in a nationwide danish population of patients diagnosed with polycystic ovary syndrome. *European Journal of Endocrinology*, 172(5):627–638.

Glueck, C., Papanna, R., Wang, P., Goldenberg, N., and Sieve-Smith, L. (2003). Incidence and treatment of metabolic syndrome in newly referred women with confirmed polycystic ovarian syndrome. *Metabolism*, 52(7):908–915.

Gnani, S. and Majeed, A. (2006). *A user's guide to data collected in primary care in England.* ERPHO.

Godoy-Matos, A. F., Vaisman, F., Pedrosa, A. P., Farias, M. L., Mendonça, L. M. C., and Pinheiro, M. F. M. (2009). Central-to-peripheral fat ratio, but not peripheral body fat, is related to insulin resistance and androgen markers in polycystic ovary syndrome. *Gynecological Endocrinology*, 25(12):793–798.

Goodarzi, M. O., Quiñones, M. J., Azziz, R., Rotter, J. I., Hsueh, W. A., and Yang, H. (2005). Polycystic ovary syndrome in mexican-americans: prevalence and association with the severity of insulin resistance. *Fertility and sterility*, 84(3):766–769.

Grundy, S. M., Cleeman, J. I., Daniels, S. R., Donato, K. A., Eckel, R. H., Franklin, B. A., Gordon, D. J., Krauss, R. M., Savage, P. J., Smith, S. C., et al. (2005). Diagnosis and management of the metabolic syndrome. *Circulation*, 112(17):2735–2752.

Gulcelik, N. E., Aral, Y., Serter, R., and Koc, G. (2008). Association of hypoadiponectinemia with metabolic syndrome in patients with polycystic ovary syndrome. *Journal of the National Medical Association*, 100(1):64–68.

Guleria, A., Syal, S., Kapoor, A., Kumar, S., Tiwari, P., and Dabadghao, P. (2014). Cardiovascular disease risk in young indian women with polycystic ovary syndrome. *Gynecological Endocrinology*, 30(1):26–29.

Haidich, A. (2010). Meta-analysis in medical research. *Hippokratia*, 14(Suppl 1):29.

Hall, G. (1992). Pharmacoepidemiology using a uk database of primary care records. *Pharmacoepidemiology and Drug Safety*, 1(1):33–37.

Hardiman, P., Pillay, O. S., and Atiomo, W. (2003). Polycystic ovary syndrome and endometrial carcinoma. *The lancet*, 361(9371):1810–1812.

Hart, R., Doherty, D. A., Mori, T., Huang, R.-C., Norman, R. J., Franks, S., Sloboda, D., Beilin, L., and Hickey, M. (2011). Extent of metabolic risk in adolescent girls with features of polycystic ovary syndrome. *Fertility and sterility*, 95(7):2347–2353.

Hashemipour, M., Faghihimani, S., Zolfaghary, B., Hovsepian, S., Ahmadi, F., and Haghighi, S. (2004). Prevalence of polycystic ovary syndrome in girls aged 14–18 years in isfahan, iran. *Hormone Research in Paediatrics*, 62(6):278–282.

Health Service Journal (1999). Vamp comes alive. *Health Service Journal*.

Hex, N., Bartlett, C., Wright, D., Taylor, M., and Varley, D. (2012). Estimating the current and future costs of type 1 and type 2 diabetes in the uk, including direct health costs and indirect societal and productivity costs. *Diabetic Medicine*, 29(7):855–862.

Higgins, J., Thompson, S. G., and Spiegelhalter, D. J. (2009). A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(1):137–159.

Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). Bayesian model averaging: a tutorial. *Statistical science*, pages 382–401.

Holte, J., Bergh, T., Berne, C., Wide, L., and Lithell, H. (1995). Restored insulin sensitivity but persistently increased early insulin secretion after weight loss in obese women with polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, 80(9):2586–2593.

Homburg, R., Berkowitz, D., Levy, T., Feldberg, D., Ashkenazi, J., and Ben-Rafael, Z. (1993). In vitro fertilization and embryo transfer for the treatment of infertility associated with polycystic ovary syndrome. *Fertility and sterility*, 60(5):858–863.

Horsfall, L., Walters, K., and Petersen, I. (2013). Identifying periods of acceptable computer usage in primary care research databases. *Pharmacoepidemiology and drug safety*, 22(1):64–69.

Hosseinpanah, F., Barzin, M., Erfani, H., Serahati, S., Ramezani Tehrani, F., and Azizi, F. (2014). Lipid accumulation product and insulin resistance in iranian pcos prevalence study. *Clinical endocrinology*, 81(1):52–57.

HU, Z.-p., PANG, T.-s., Ying, W., Jie, Q., and LI, M.-z. (2014). Study on the clinical and endocrine characteristics of polycystic ovary syndrome with different ovarian morphology. *Journal of Reproduction and Contraception*, 25(3):133–145.

Huang, J., Ni, R., Chen, X., Huang, L., Mo, Y., and Yang, D. (2010). Metabolic abnormalities in adolescents with polycystic ovary syndrome in south china. *Reproductive Biology and Endocrinology*, 8(1):142.

Huber-Buchholz, M.-M., Carey, D., and Norman, R. (1999). Restoration of reproductive potential by lifestyle modification in obese polycystic ovary syndrome: Role of insulin sensitivity and luteinizing hormone 1. *The Journal of Clinical Endocrinology & Metabolism*, 84(4):1470–1474.

Hudecova, M., Holte, J., Olovsson, M., Larsson, A., Berne, C., and Poromaa, I. S. (2011a). Diabetes and impaired glucose tolerance in patients with polycystic ovary syndromea long term follow-up. *Human reproduction*, page der065.

Hudecova, M., Holte, J., Olovsson, M., Larsson, A., Berne, C., and Sundstrom-Poromaa, I. (2011b). Prevalence of the metabolic syndrome in women with a previous diagnosis of polycystic ovary syndrome: long-term follow-up. *Fertility and sterility*, 96(5):1271–1274.

IDF (2013). Idf diabetes atlas. *Brussels: International Diabetes Federation*.

IMSHealth (2015). Data content.

Jacob, E. and Jayakumar, K. (2012). On half-cauchy distribution and process. *International Journal of Statistika and Mathematika*, 3(2):77–81.

Jiao, J., Fang, Y., Wang, T., Wang, Z., Zhou, M., and Wang, X. (2013). Epidemiologic investigation of polycystic ovarian syndrome (pcos) in han ethnic women of reproductive age in liaoning province, china. *Clinical and experimental obstetrics & gynecology*, 41(3):304–309.

Joham, A., Ranasinha, S., Zoungas, S., Moran, L., and Teede, H. (2013). Gestational diabetes and type 2 diabetes in reproductive-aged women with polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, 99(3):E447–E452.

Joshi, B., Mukherjee, S., Patil, A., Purandare, A., Chauhan, S., and Vaidya, R. (2014). A cross-sectional study of polycystic ovarian syndrome among adolescent and young girls in mumbai, india. *Indian journal of endocrinology and metabolism*, 18(3):317.

Karaer, A., Cavkaytar, S., Mert, I., Buyukkagnici, U., and Batioglu, S. (2010). Cardiovascular risk factors in polycystic ovary syndrome. *Journal of Obstetrics and Gynaecology*, 30(4):387–392.

Karoli, R., Fatima, J., Chandra, A., Gupta, U., Islam, F.-u., Singh, G., et al. (2013). Prevalence of hepatic steatosis in women with polycystic ovary syndrome. *Journal of human reproductive sciences*, 6(1):9.

Kawai, T., Ng, M. C., Hayes, M. G., Yoshiuchi, I., Tsuchiya, T., Robertson, H., Cox, N. J., Polonsky, K. S., Bell, G. I., and Ehrmann, D. A. (2009). Variation in the perilipin gene (plin) affects glucose and lipid metabolism in non-hispanic white women with and without polycystic ovary syndrome. *Diabetes research and clinical practice*, 86(3):186–192.

Kim, H.-J., Bjonnes, A., Saxena, R., and Welt, C. (2014). Evaluating risk variants for korean women with polycystic ovary syndrome in women of european ethnicity. *Fertility and Sterility*, 102(3):e266–e267.

Kind, P., Hardman, G., Macran, S., et al. (1999). *UK population norms for EQ-5D*, volume 172. Centre for Health Economics, University of York York.

Knochenhauer, E., Key, T., Kahsar-Miller, M., Waggoner, W., Boots, L., and Azziz, R. (1998). Prevalence of the polycystic ovary syndrome in unselected black and white women of the southeastern united states: A prospective study 1. *The Journal of Clinical Endocrinology & Metabolism*, 83(9):3078–3082.

Kollipaka, R., Arounassalame, B., and Lakshminarayanan, S. (2013). Does psychosocial stress influence menstrual abnormalities in medical students? *Journal of Obstetrics and Gynaecology*, 33(5):489–493.

Kumarapeli, V., Seneviratne, R. d. A., Wijeyaratne, C., Yapa, R., and Dodampahala, S. (2008). A simple screening approach for assessing community prevalence and phenotype of polycystic ovary syndrome in a semiurban population in sri lanka. *American journal of epidemiology*, 168(3):321–328.

Lankarani, M., Valizadeh, N., Heshmat, R., Peimani, M., and Sohrabvand, F. (2009). Evaluation of insulin resistance and metabolic syndrome in patients with polycystic ovary syndrome. *Gynecological Endocrinology*, 25(8):504–507.

Lauritsen, M., Bentzen, J., Pinborg, A., Loft, A., Forman, J. L., Thuesen, L., Cohen, A., Hougaard, D., and Andersen, A. N. (2014). The prevalence of polycystic ovary syndrome

in a normal population according to the rotterdam criteria versus revised criteria including anti-müllerian hormone. *Human Reproduction*, 29(4):791–801.

Lee, H., Oh, J.-Y., Sung, Y.-A., Chung, H., and Cho, W. Y. (2009). The prevalence and risk factors for glucose intolerance in young korean women with polycystic ovary syndrome. *Endocrine*, 36(2):326–332.

Legro, R. S. (2003). Polycystic ovary syndrome and cardiovascular disease: a premature association? *Endocrine reviews*, 24(3):302–312.

Legro, R. S. (2012). Obesity and pcos: implications for diagnosis and treatment. *Seminars in reproductive medicine*, 30(6):496.

Legro, R. S., Arslanian, S. A., Ehrmann, D. A., Hoeger, K. M., Murad, M. H., Pasquali, R., and Welt, C. K. (2013). Diagnosis and treatment of polycystic ovary syndrome: an endocrine society clinical practice guideline. *The Journal of Clinical Endocrinology & Metabolism*, 98(12):4565–4592.

Legro, R. S., Castracane, V. D., and Kauffman, R. P. (2004). Detecting insulin resistance in polycystic ovary syndrome: purposes and pitfalls. *Obstetrical & gynecological survey*, 59(2):141–154.

Legro, R. S., Gnatuk, C. L., Kunselman, A. R., and Dunaif, A. (2005). Changes in glucose tolerance over time in women with polycystic ovary syndrome: a controlled study. *The Journal of Clinical Endocrinology & Metabolism*, 90(6):3236–3242.

Legro, R. S., Kunselman, A. R., Dodson, W. C., and Dunaif, A. (1999). Prevalence and predictors of risk for type 2 diabetes mellitus and impaired glucose tolerance in polycystic ovary syndrome: a prospective, controlled study in 254 affected women 1. *The journal of clinical endocrinology & metabolism*, 84(1):165–169.

Legro, R. S., Kunselman, A. R., and Dunaif, A. (2001). Prevalence and predictors of dyslipidemia in women with polycystic ovary syndrome. *The American journal of medicine*, 111(8):607–613.

Leibel, N. I., Baumann, E. E., Kocherginsky, M., and Rosenfield, R. L. (2006). Relationship of adolescent polycystic ovary syndrome to parental metabolic syndrome. *The Journal of Clinical Endocrinology & Metabolism*, 91(4):1275–1283.

Lewis, J. D., Schinnar, R., Bilker, W. B., Wang, X., and Strom, B. L. (2007). Validation studies of the health improvement network (thin) database for pharmacoepidemiology research. *Pharmacoepidemiology and drug safety*, 16(4):393–401.

Li, L., Chen, X., He, Z., Zhao, X., Huang, L., and Yang, D. (2012). Clinical and metabolic features of polycystic ovary syndrome among chinese adolescents. *Journal of pediatric and adolescent gynecology*, 25(6):390–395.

Li, R., Zhang, Q., Yang, D., Li, S., Lu, S., Wu, X., Wei, Z., Song, X., Wang, X., Fu, S., et al. (2013a). Prevalence of polycystic ovary syndrome in women in china: a large community-based study. *Human Reproduction*, page det262.

Li, R., Zhang, Q., Yang, D., Li, S., Lu, S., Wu, X., Wei, Z., Song, X., Wang, X., Fu, S., et al. (2013b). Prevalence of polycystic ovary syndrome in women in china: a large community-based study. *Human reproduction*, page det262.

Li, Y., Li, Y., Ng, E. H. Y., Stener-Victorin, E., Hou, L., Wu, T., Han, F., and Wu, X. (2011a). Polycystic ovary syndrome is associated with negatively variable impacts on domains of health-related quality of life: evidence from a meta-analysis. *Fertility and sterility*, 96(2):452–458.

Li, Y., Li, Y., Ng, E. H. Y., Stener-Victorin, E., Hou, L., Wu, T., Han, F., and Wu, X. (2011b). Polycystic ovary syndrome is associated with negatively variable impacts on domains of health-related quality of life: evidence from a meta-analysis. *Fertility and sterility*, 96(2):452–458.

LIANG, S.-J., LIOU, T.-H., LIN, H.-W., HSU, C.-S., TZENG, C.-R., and HSU, M.-I. (2012). Obesity is the predominant predictor of impaired glucose tolerance and metabolic disturbance in polycystic ovary syndrome. *Acta obstetricia et gynecologica Scandinavica*, 91(10):1167–1172.

Lim, S., Davies, M., Norman, R., and Moran, L. (2012). Overweight, obesity and central obesity in women with polycystic ovary syndrome: a systematic review and meta-analysis. *Human reproduction update*, page dms030.

Lindholm, Å., Andersson, L., Eliasson, M., Bixo, M., and Sundström-Poromaa, I. (2008). Prevalence of symptoms associated with polycystic ovary syndrome. *International Journal of Gynecology & Obstetrics*, 102(1):39–43.

Livadas, S. and Diamanti-Kandarakis, E. (2013). Polycystic ovary syndrome: definitions, phenotypes and diagnostic approach. *Frontiers of Hormone Research*, 40:1–21.

Lo, J. C., Feigenbaum, S. L., Yang, J., Pressman, A. R., Selby, J. V., and Go, A. S. (2006a). Epidemiology and adverse cardiovascular risk profile of diagnosed polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, 91(4):1357–1363.

Lo, J. C., Feigenbaum, S. L., Yang, J., Pressman, A. R., Selby, J. V., and Go, A. S. (2006b). Epidemiology and adverse cardiovascular risk profile of diagnosed polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, 91(4):1357–1363.

Louwers, Y., Lao, O., and Kayser, M. (2013). Inferred genetic ancestry versus reported ethnicity in polycystic ovary syndrome (pcos). In *HUMAN REPRODUCTION*, volume 28, pages 349–349. OXFORD UNIV PRESS GREAT CLARENDON ST, OXFORD OX2 6DP, ENGLAND.

Lowe, P., Kovacs, G., and Howlett, D. (2005). Incidence of polycystic ovaries and polycystic ovary syndrome amongst women in melbourne, australia. *Australian and New Zealand journal of obstetrics and gynaecology*, 45(1):17–19.

Lumley, T. (2009). rmeta: Meta-analysis. *R package version*, 2.

Lunn, D., Jackson, C., Best, N., Thomas, A., and Spiegelhalter, D. (2012). *The BUGS book: A practical introduction to Bayesian analysis*. CRC press.

Luque-Ramírez, M., Alpañés, M., and Escobar-Morreale, H. F. (2010). The determinants of insulin sensitivity, β-cell function, and glucose tolerance are different in patients with polycystic ovary syndrome than in women who do not have hyperandrogenism. *Fertility and sterility*, 94(6):2214–2221.

Ma, Y.-M., Li, R., Qiao, J., Zhang, X.-w., Wang, S.-Y., Zhang, Q.-f., Li, L., Tu, B.-B., and Zhang, X. (2010). Characteristics of abnormal menstrual cycle and polycystic ovary syndrome in community and hospital populations. *Chinese medical journal*, 123(16):2185–2189.

Maguire, A., Blak, B. T., and Thompson, M. (2009). The importance of defining periods of complete mortality reporting for research using automated data from primary care. *Pharmacoepidemiology and drug safety*, 18(1):76–83.

March, W. A., Moore, V. M., Willson, K. J., Phillips, D. I., Norman, R. J., and Davies, M. J. (2010). The prevalence of polycystic ovary syndrome in a community sample assessed under contrasting diagnostic criteria. *Human reproduction*, 25(2):544–551.

Márquez, J. L., Pacheco, A., Valdés, P., and Salazar, L. A. (2008). Association between capn10 ucsnp-43 gene polymorphism and polycystic ovary syndrome in chilean women. *Clinica Chimica Acta*, 398(1):5–9.

Mehrabian, F., Khani, B., Kelishadi, R., and Ghanbari, E. (2011). The prevalence of polycystic ovary syndrome in iranian women based on different diagnostic criteria. *Endokrynologia Polska*, 62(3):238–242.

Melo, A. S., Vieira, C. S., Romano, L. G. M., Ferriani, R. A., and Navarro, P. A. (2011). The frequency of metabolic syndrome is higher among pcos brazilian women with menstrual irregularity plus hyperandrogenism. *Reproductive Sciences*, 18(12):1230–1236.

Meyer, C., McGrath, B. P., and Teede, H. J. (2007). Effects of medical therapy on insulin resistance and the cardiovascular system in polycystic ovary syndrome. *Diabetes Care*, 30(3):471–478.

Michelmore, K., Balen, A., Dunger, D., and Vessey, M. (1999). Polycystic ovaries and associated clinical and biochemical features in young women. *Clinical endocrinology*, 51(6):779–786.

Moghetti, P. and Toscano, V. (2006). Treatment of hirsutism and acne in hyperandrogenism. *Best Practice & Research Clinical Endocrinology & Metabolism*, 20(2):221–234.

Moini, A. and Eslami, B. (2009). Familial associations between polycystic ovarian syndrome and common diseases. *Journal of assisted reproduction and genetics*, 26(2-3):123–127.

Moran, C., Tena, G., Moran, S., Ruiz, P., Reyna, R., and Duque, X. (2010a). Prevalence of polycystic ovary syndrome and related disorders in mexican women. *Gynecologic and obstetric investigation*, 69(4):274–280.

Moran, L. and Teede, H. (2009). Metabolic features of the reproductive phenotypes of polycystic ovary syndrome. *Human reproduction update*.

Moran, L. J., Misso, M. L., Wild, R. A., and Norman, R. J. (2010b). Impaired glucose tolerance, type 2 diabetes and metabolic syndrome in polycystic ovary syndrome: a systematic review and meta-analysis. *Human Reproduction Update*, 16(4):347–363.

Moran, L. J., Pasquali, R., Teede, H. J., Hoeger, K. M., and Norman, R. J. (2009). Treatment of obesity in polycystic ovary syndrome: a position statement of the androgen excess and polycystic ovary syndrome society. *Fertility and sterility*, 92(6):1966–1982.

Morgan, C. L., Jenkins-Jones, S., Currie, C. J., and Rees, D. A. (2012). Evaluation of adverse outcome in young women with polycystic ovary syndrome versus matched, reference controls: a retrospective, observational study. *The Journal of Clinical Endocrinology & Metabolism*, 97(9):3251–3260.

Mulnier, H., Seaman, H., Raleigh, V., Soedamah-Muthu, S., Colhoun, H., and Lawrenson, R. (2006). Mortality in people with type 2 diabetes in the uk. *Diabetic Medicine*, 23(5):516–521.

Munn, Z., Moola, S., Riitano, D., and Lisy, K. (2014). The development of a critical appraisal tool for use in systematic reviews addressing questions of prevalence. *International journal of health policy and management*, 3(3):123.

Musmar, S., Afaneh, A., and Mo'alla, H. (2013). Epidemiology of polycystic ovary syndrome: a cross sectional study of university students at an-najah national university-palestine. *Reprod Biol Endocrinol*, 11:47.

Nahuis, M., Kose, N., Bayram, N., Van Dessel, H., Braat, D., Hamilton, C., Hompes, P., Bossuyt, P., Mol, B., Van Der Veen, F., et al. (2011). Long-term outcomes in women with polycystic ovary syndrome initially randomized to receive laparoscopic electrocautery of the ovaries or ovulation induction with gonadotrophins. *Human reproduction*, 26(7):1899–1904.

Nandalike, K., Agarwal, C., Strauss, T., Coupey, S. M., Isasi, C. R., Sin, S., and Arens, R. (2012). Sleep and cardiometabolic function in obese adolescent girls with polycystic ovary syndrome. *Sleep medicine*, 13(10):1307–1312.

National Diabetes Data Group (1979). Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *Diabetes*, 28(12):1039–1057.

National Health Service (2015). *Guidelines for the management of acne (from 12 years of age).* http://www.lambethccg.nhs.uk/news-and-publications/meeting-papers/lambeth-borough-prescribing-committee/Lambeth%20Borough%20Prescribing%20Committee/Clinical%20Guidelines/Management%20of%20Acne%20Guidelines%20Jan%202015.pdf.

National Heatlh Service (1998). *Information for health: an information strategy for the modern NHS 1998-2005: a national strategy for local implementation.* NHS Executive.

National Institute for Clinical Excellence (2004). Guide to the methods of technology appraisal. *London: NICE*, 12.

National Institute for Clinical Excellence (2011). *Hyperglycaemia in acute coronary syndromes: Costing statement, Implementing NICE guidance.* https://www.nice.org.uk/guidance/cg130/resources/costing-statement-183367117.

Nidhi, R., Padmalatha, V., Nagarathna, R., and Amritanshu, R. (2011). Prevalence of polycystic ovarian syndrome in indian adolescents. *Journal of pediatric and adolescent gynecology*, 24(4):223–227.

Nur, M. M., Newman, I. M., and Siqueira, L. M. (2009). Glucose metabolism in overweight hispanic adolescents with and without polycystic ovary syndrome. *Pediatrics*, 124(3):e496–e502.

Office for National Statistics (2011). *Ethnicity and National Identity in England and Wales: 2011.* https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/articles/ethnicityandnationalidentityinenglandandwales/2012-12-11.

Office for National Statistics (2014). *Annual Mid-year Population Estimates: 2014.* https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/articles/ethnicityandnationalidentityinenglandandwales/2012-12-11.

Office for National Statistics (2016). *National life tables, UK: 20132015.* https://www.

ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/
lifeexpectancies/bulletins/nationallifetablesunitedkingdom/20132015.

Ogdie, A., Langan, S., Parkinson, J., Dattani, H., Kostev, K., and Gelfand, J. M. (2012). Medical record databases. *Pharmacoepidemiology. 5th ed. Oxford: Wiley-Blackwell*, pages 224–43.

Okoroh, E. M., Hooper, W. C., Atrash, H. K., Yusuf, H. R., and Boulet, S. L. (2012). Prevalence of polycystic ovary syndrome among the privately insured, united states, 2003-2008. *American journal of obstetrics and gynecology*, 207(4):299–e1.

O'Meara, N., Blackman, J., Ehrmann, D., Barnes, R., Jaspan, J., Rosenfield, R., and Polonsky, K. (1993). Defects in beta-cell function in functional ovarian hyperandrogenism. *The Journal of Clinical Endocrinology & Metabolism*, 76(5):1241–1247.

Ovalle, F. and Azziz, R. (2002). Insulin resistance, polycystic ovary syndrome, and type 2 diabetes mellitus. *Fertility and sterility*, 77(6):1095–1105.

Pall, M., Azziz, R., Beires, J., and Pignatelli, D. (2010). The phenotype of hirsute women: a comparison of polycystic ovary syndrome and 21-hydroxylase–deficient nonclassic adrenal hyperplasia. *Fertility and sterility*, 94(2):684–689.

Panidis, D., Tziomalos, K., Macut, D., Kandaraki, E. A., Tsourdi, E. A., Papadakis, E., and Katsikis, I. (2013). Age-and body mass index-related differences in the prevalence of metabolic syndrome in women with polycystic ovary syndrome. *Gynecological Endocrinology*, 29(10):926–930.

Phy, J. L., Conover, C. A., Abbott, D. H., Zschunke, M. A., Walker, D. L., Session, D. R., Tummon, I. S., Thornhill, A. R., Lesnick, T. G., and Dumesic, D. A. (2004). Insulin and messenger ribonucleic acid expression of insulin receptor isoforms in ovarian follicles from nonhirsute ovulatory women and polycystic ovary syndrome patients. *The Journal of Clinical Endocrinology & Metabolism*, 89(7):3561–3566.

Pierpoint, T., McKeigue, P., Isaacs, A., Wild, S., and Jacobs, H. (1998). Mortality of women with polycystic ovary syndrome at long-term follow-up. *Journal of clinical epidemiology*, 51(7):581–586.

Piltonen, T., Koivunen, R., Perheentupa, A., Morin-Papunen, L., Ruokonen, A., and Tapanainen, J. S. (2004). Ovarian age-related responsiveness to human chorionic gonadotropin in women with polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, 89(8):3769–3775.

Plummer, M. et al. (2003). Jags: A program for analysis of bayesian graphical models using gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing*, volume 124, page 125. Technische Universit at Wien.

Praet, C. and D'Oca, K. (2014). Cost-benefit model of varying nexplanon and other long-acting reversible contraceptive (larc) methods: Uptake compared to the oral contraceptive pill: Uk perspective. *Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research*, 17(7):A508–A508.

Qin, J. Z., Pang, L. H., Li, M. J., Fan, X. J., Huang, R. D., and Chen, H. Y. (2013). Obstetric complications in women with polycystic ovary syndrome: a systematic review and meta-analysis. *Reprod Biol Endocrinol*, 11:56.

Rahmanpour, H., Jamal, L., Mousavinasab, S. N., Esmailzadeh, A., and Azarkhish, K. (2012). Association between polycystic ovarian syndrome, overweight, and metabolic syndrome in adolescents. *Journal of pediatric and adolescent gynecology*, 25(3):208–212.

Rajkhowa, M., Talbot, J., Jones, P., and Clayton, R. (1996). Polymorphism of glycogen synthetase gene in polycystic ovary syndrome. *Clinical endocrinology*, 44(1):85–90.

Ramos, R. B. and Spritzer, P. M. (2015). Fto gene variants are not associated with polycystic ovary syndrome in women from southern brazil. *Gene*, 560(1):25–29.

Rashidi, H., Tehrani, F. R., Khomami, M. B., Tohidi, M., and Azizi, F. (2014). To what extent does the use of the rotterdam criteria affect the prevalence of polycystic ovary syndrome? a community-based study from the southwest of iran. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 174:100–105.

Rizzo, M., Longo, R., Guastella, E., Rini, G., and Carmina, E. (2011). Assessing cardiovascular risk in mediterranean women with polycystic ovary syndrome. *Journal of endocrinological investigation*, 34(6):422–426.

Roe, A. H., Prochaska, E., Smith, M., Sammel, M., and Dokras, A. (2013). Using the androgen excess–pcos society criteria to diagnose polycystic ovary syndrome and the risk of metabolic syndrome in adolescents. *The Journal of pediatrics*, 162(5):937–941.

Rotterdam, E. and ASRM-Sponsored, P. (2004a). Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertility and sterility*, 81(1):19.

Rotterdam, E. and ASRM-Sponsored, P. (2004b). Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (pcos). *Human Reproduction (Oxford, England)*, 19(1):41.

Sackett, D. L., Rosenberg, W. M., Gray, J., Haynes, R. B., and Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. *BMJ: British Medical Journal*, 312(7023):71.

Sam, S. (2007). Obesity and polycystic ovary syndrome. *Obesity management*, 3(2):69–73.

Sanchón, R., Gambineri, A., Alpañés, M., Martínez-García, M. Á., Pasquali, R., and Escobar-Morreale, H. F. (2012). Prevalence of functional disorders of androgen excess in unselected premenopausal women: a study in blood donors. *Human reproduction*, 27(4):1209–1216.

Sawathiparnich, P., Weerakulwattana, L., Santiprabhob, J., and Likitmaskul, S. (2005). Obese adolescent girls with polycystic ovary syndrome (pcos) have more severe insulin resistance measured by homa-ir score than obese girls without pcos. *J Med Assoc Thai*, 88(Suppl 8):S33–S37.

Schmidt, J., Brännström, M., Landin-Wilhelmsen, K., and Dahlgren, E. (2011). Reproductive hormone levels and anthropometry in postmenopausal women with polycystic ovary syndrome (pcos): a 21-year follow-up study of women diagnosed with pcos around 50 years ago and their age-matched controls. *The Journal of Clinical Endocrinology & Metabolism*, 96(7):2178–2185.

Schwarzer, G. (2007). Meta: An r package for meta-analysis. *SpherWave: An R Package for Analyzing Scattered Spherical Data by Spherical Wavelets*, page 40.

Seow, K.-M., Juan, C.-C., Hsu, Y.-P., Hwang, J.-L., Huang, L.-W., and Ho, L.-T. (2007). Amelioration of insulin resistance in women with pcos via reduced insulin receptor substrate-

1 ser312 phosphorylation following laparoscopic ovarian electrocautery. *Human reproduction*, 22(4):1003–1010.

Shalom-Paz, E., Holzer, H., Son, W.-Y., Levin, I., Tan, S. L., and Almog, B. (2012). Pcos patients can benefit from in vitro maturation (ivm) of oocytes. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 165(1):53–56.

Sharma, M., Nazareth, I., and Petersen, I. (2016). Trends in incidence, prevalence and prescribing in type 2 diabetes mellitus between 2000 and 2013 in primary care: a retrospective cohort study. *BMJ open*, 6(1):e010210.

Shaw, L. J., Bairey Merz, C. N., Azziz, R., Stanczyk, F. Z., Sopko, G., Braunstein, G. D., Kelsey, S. F., Kip, K. E., Cooper-DeHoff, R. M., Johnson, B. D., et al. (2008). Withdrawn: Postmenopausal women with a history of irregular menses and elevated androgen measurements at high risk for worsening cardiovascular event-free survival: Results from the national institutes of healthnational heart, lung, and blood institute sponsored womens ischemia syndrome evaluation. *The Journal of Clinical Endocrinology & Metabolism*, 93(4):1276–1284.

Sheehan, M. T. (2004). Polycystic ovarian syndrome: diagnosis and management. *Clinical Medicine & Research*, 2(1):13–27.

Shroff, R., Kerchner, A., Maifeld, M., Van Beek, E. J., Jagasia, D., and Dokras, A. (2007a). Young obese women with polycystic ovary syndrome have evidence of early coronary atherosclerosis. *The Journal of Clinical Endocrinology & Metabolism*, 92(12):4609–4614.

Shroff, R., Syrop, C. H., Davis, W., Van Voorhis, B. J., and Dokras, A. (2007b). Risk of metabolic complications in the new pcos phenotypes based on the rotterdam criteria. *Fertility and sterility*, 88(5):1389–1395.

Sir-Petermann, T., Angel, B., Maliqueo, M., Santos, J. L., Riesco, M. V., Toloza, H., and Pérez-Bravo, F. (2004). Insulin secretion in women who have polycystic ovary syndrome and carry the gly972arg variant of insulin receptor substrate-1 in response to a high-glycemic or low-glycemic carbohydrate load. *Nutrition*, 20(10):905–910.

Sirmans, S. M., Parish, R. C., Blake, S., and Wang, X. (2014a). Epidemiology and comorbidities

of polycystic ovary syndrome in an indigent population. *Journal of Investigative Medicine*, 62(6):868–874.

Sirmans, S. M., Parish, R. C., Blake, S., and Wang, X. (2014b). Epidemiology and comorbidities of polycystic ovary syndrome in an indigent population. *Journal of Investigative Medicine*, 62(6):868–874.

Sirmans, S. M. and Pate, K. A. (2014). Epidemiology, diagnosis, and management of polycystic ovary syndrome. *Clinical epidemiology*, 6:1.

Sonnenberg, F. A. and Beck, J. R. (1993). Markov models in medical decision making: a practical guide. *Medical decision making*, 13(4):322–338.

Spiegelhalter, D., Thomas, A., Best, N., and Lunn, D. (2003). Winbugs user manual.

Spiegelhalter, D. J., Abrams, K. R., and Myles, J. P. (2004). *Bayesian approaches to clinical trials and health-care evaluation*, volume 13. John Wiley & Sons.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639.

Stang, A. (2010). Critical evaluation of the newcastle-ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *European journal of epidemiology*, 25(9):603–605.

Sterne, J. A., Bradburn, M. J., and Egger, M. (2008). Meta–analysis in stata. *Systematic Reviews in Health Care: Meta-Analysis in Context, Second Edition*, pages 347–369.

Sung, Y.-A., Kim, D.-S., Yoo, S.-J., Baik, S.-H., Oh, J.-Y., and Lee, H.-J. (2010). The prevalence and phenotypes of polycystic ovary syndrome in korean women.

Sutton, A. J. and Abrams, K. R. (2001a). Bayesian methods in meta-analysis and evidence synthesis. *Statistical methods in medical research*, 10(4):277–303.

Sutton, A. J. and Abrams, K. R. (2001b). Bayesian methods in meta-analysis and evidence synthesis. *Statistical Methods in Medical Research*, 10(4):277–303.

Tan, S., Bechmann, L. P., Benson, S., Dietz, T., Eichner, S., Hahn, S., Janssen, O. E., Lahner, H., Gerken, G., Mann, K., et al. (2010). Apoptotic markers indicate nonalcoholic steatohepatitis in polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, 95(1):343–348.

Teede, H., Deeks, A., and Moran, L. (2010a). Polycystic ovary syndrome: a complex condition with psychological, reproductive and metabolic manifestations that impacts on health across the lifespan. *BMC medicine*, 8(1):41.

Teede, H., Deeks, A., and Moran, L. (2010b). Polycystic ovary syndrome: a complex condition with psychological, reproductive and metabolic manifestations that impacts on health across the lifespan. *BMC medicine*, 8(1):41.

Teede, H. J., Hutchison, S. K., and Zoungas, S. (2007). The management of insulin resistance in polycystic ovary syndrome. *Trends in Endocrinology & Metabolism*, 18(7):273–279.

Tehrani, F. R., Rashidi, H., Khomami, M. B., Tohidi, M., and Azizi, F. (2014). The prevalence of metabolic disorders in various phenotypes of polycystic ovary syndrome: a community based study in southwest of iran. *Reproductive Biology and Endocrinology*, 12(1):89.

Tehrani, F. R., Simbar, M., Tohidi, M., Hosseinpanah, F., and Azizi, F. (2011). The prevalence of polycystic ovary syndrome in a community sample of iranian population: Iranian pcos prevalence study. *Reprod Biol Endocrinol*, 9(39):39.

Thathapudi, S., Kodati, V., Erukkambattu, J., Katragadda, A., Addepally, U., Hasan, Q., et al. (2014). Anthropometric and biochemical characteristics of polycystic ovarian syndrome in south indian women using aes-2006 criteria. *International journal of endocrinology and metabolism*, 12(1).

Thurecht, L., Brown, L., and Yap, M. (2007). Economic modelling of the prevention of type 2 diabetes in australia.

Torgerson, D. J. and Raftery, J. (1999). Discounting. *Bmj*, 319(7214):914–915.

Townsend, P., Phillimore, P., and Beattie, A. (1988). *Health and deprivation: inequality and the North*. Routledge.

Tsilchorozidou, T., Overton, C., and Conway, G. S. (2004). The pathophysiology of polycystic ovary syndrome. *Clinical endocrinology*, 60(1):1–17.

Veltman-Verhulst, S. M., van Rijn, B. B., Westerveld, H. E., Franx, A., Bruinse, H. W., Fauser, B. C., and Goverde, A. J. (2010). Polycystic ovary syndrome and early-onset preeclampsia: reproductive manifestations of increased cardiovascular risk. *Menopause*, 17(5):990–996.

Viechtbauer, W. et al. (2010). Conducting meta-analyses in r with the metafor package. *Journal of Statistical Software*, 36(3):1–48.

Vrbíková, J., Bendlová, B., Hill, M., Vanková, M., Vondra, K., and Stárka, L. (2002). Insulin sensitivity and β-cell function in women with polycystic ovary syndrome. *Diabetes Care*, 25(7):1217–1222.

VrbIkova, J., Cibula, D., Dvorakova, K., Stanicka, S., Sindelka, G., Hill, M., Fanta, M., Vondra, K., and Skrha, J. (2004). Insulin sensitivity in women with polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, 89(6):2942–2945.

VrbÍková, J., Cibula, D., Dvoráková, K., Stanicka, S., Sindelka, G., Hill, M., Fanta, M., Vondra, K., and Skrha, J. (2004). Insulin sensitivity in women with polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, 89(6):2942–2945.

Vrbikova, J., Fanta, M., Cibula, D., Vondra, K., and Bendlova, B. (2009). Impaired glucose metabolism in women with polycystic ovary syndrome. *Gynecologic and obstetric investigation*, 68(3):186–190.

VrbÍková, J., Zamrazilová, H., SedláČková, B., and Šnajderová, M. (2011). Metabolic syndrome in adolescents with polycystic ovary syndrome. *Gynecological Endocrinology*, 27(10):820–822.

Vutyavanich, T., Khaniyao, V., Wongtra-ngan, S., Sreshthaputra, O., Sreshthaputra, R., and Piromlertamorn, W. (2007). Clinical, endocrine and ultrasonographic features of polycystic ovary syndrome in thai women. *Journal of Obstetrics and Gynaecology Research*, 33(5):677–680.

Wang, E. T., Calderon-Margalit, R., Cedars, M. I., Daviglus, M. L., Merkin, S. S., Schreiner, P. J., Sternfeld, B., Wellons, M., Schwartz, S. M., Lewis, C. E., et al. (2011). Polycystic

ovary syndrome and risk for long-term diabetes and dyslipidemia. *Obstetrics and gynecology*, 117(1):6.

Wang, S. and Alvero, R. (2013). Racial and ethnic differences in physiology and clinical symptoms of polycystic ovary syndrome. In *Seminars in reproductive medicine*, volume 31, pages 365–369. Thieme Medical Publishers.

Wei, H.-J., Young, R., Kuo, I.-L., Liaw, C.-M., Chiang, H.-S., and Yeh, C.-Y. (2009). Prevalence of insulin resistance and determination of risk factors for glucose intolerance in polycystic ovary syndrome: a cross-sectional study of chinese infertility patients. *Fertility and sterility*, 91(5):1864–1868.

Welt, C., Gudmundsson, J., Arason, G., Adams, J., Palsdottir, H., Gudlaugsdottir, G., Ingadottir, G., and Crowley, W. (2006). Characterizing discrete subsets of polycystic ovary syndrome as defined by the rotterdam criteria: the impact of weight on phenotype and metabolic features. *The Journal of Clinical Endocrinology & Metabolism*, 91(12):4842–4848.

Wickham, E. P., Cheang, K. I., Clore, J. N., Baillargeon, J.-P., and Nestler, J. E. (2011). Total and high–molecular weight adiponectin in women with the polycystic ovary syndrome. *Metabolism*, 60(3):366–372.

Wijeyaratne, C. N., Seneviratne, R. d. A., Dahanayake, S., Kumarapeli, V., Palipane, E., Kuruppu, N., Yapa, C., Seneviratne, R. d. A., and Balen, A. H. (2010). Phenotype and metabolic profile of south asian women with polycystic ovary syndrome (pcos): results of a large database from a specialist endocrine clinic. *Human Reproduction*, page deq310.

Wild, S., Pierpoint, T., McKeigue, P., and Jacobs, H. (2000). Cardiovascular disease in women with polycystic ovary syndrome at long-term follow-up: a retrospective cohort study. *Clinical endocrinology*, 52(5):595–600.

Williams, L. et al. (2002). Third report of the national cholesterol education program (ncep) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel iii) final report. *Circulation*, 106(25):3143–3143.

Wiltgen, D. and Spritzer, P. M. (2010). Variation in metabolic and cardiovascular risk in women with different polycystic ovary syndrome phenotypes. *Fertility and sterility*, 94(6):2493–2496.

Winters, S. J., Talbott, E., Guzick, D. S., Zborowski, J., and McHugh, K. P. (2000). Serum testosterone levels decrease in middle age in women with the polycystic ovary syndrome. *Fertility and sterility*, 73(4):724–729.

Yaralı, H., Yıldırır, A., Aybar, F., Kabakçı, G., Bükülmez, O., Akgül, E., and Oto, A. (2001). Diastolic dysfunction and increased serum homocysteine concentrations may contribute to increased cardiovascular risk in patients with polycystic ovary syndrome. *Fertility and sterility*, 76(3):511–516.

Yildiz, B. O., Bozdag, G., Yapici, Z., Esinler, I., and Yarali, H. (2012). Prevalence, phenotype and cardiometabolic risk of polycystic ovary syndrome under different diagnostic criteria. *Human Reproduction*, 27(10):3067–3073.

Yildiz, B. O., Knochenhauer, E. S., and Azziz, R. (2008). Impact of obesity on the risk for polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, 93(1):162–168.

Yilmaz, M., Isaoglu, U., Delibas, I. B., and Kadanali, S. (2011). Anthropometric, clinical and laboratory comparison of four phenotypes of polycystic ovary syndrome based on rotterdam criteria. *Journal of Obstetrics and Gynaecology Research*, 37(8):1020–1026.

Zawadzki, J. and Dunaif, A. (1992). *Diagnostic criteria for polycystic ovary syndrome: towards a rational approach.* Boston: Blackwell Scientific Publications.

Zhang, H., Zhu, F., Xiong, J., Shi, X., and Fu, S. (2009). Characteristics of different phenotypes of polycystic ovary syndrome based on the rotterdam criteria in a large-scale chinese population. *BJOG: An International Journal of Obstetrics & Gynaecology*, 116(12):1633–1639.

Zhang, J., Fan, P., Liu, H., Bai, H., Wang, Y., and Zhang, F. (2012). Apolipoprotein ai and b levels, dyslipidemia and metabolic syndrome in south-west chinese women with pcos. *Human reproduction*, 27(8):2484–2493.

Zhao, X., Zhong, J., Mo, Y., Chen, X., Chen, Y., and Yang, D. (2010). Association of biochemical hyperandrogenism with type 2 diabetes and obesity in chinese women with polycystic ovary syndrome. *International Journal of Gynecology & Obstetrics*, 108(2):148–151.

Zhao, Y. and Qiao, J. (2013). Ethnic differences in the phenotypic expression of polycystic ovary syndrome. *Steroids*, 78(8):755–760.

Zhuang, J., Liu, Y., Xu, L., Liu, X., Zhou, L., Tang, L., Kang, D., Guo, W., He, M., Yang, F., et al. (2014). Prevalence of the polycystic ovary syndrome in female residents of chengdu, china. *Gynecologic and obstetric investigation*, 77(4):217–223.

Appendices

Appendix A

Table 1. Summary of studies evaluationg the prevalence of PCOS in women of reproductive age.

Article	Country	Sample size	Age	BMI	Prevalence of PCOS (%)			
					NIH	Rotterdam	AES	Other
Americas								
Knochenhauer et al. (1998)	US	277	18-45yrs	26.9+/-7.4	4.0	-	-	-
Azziz et al. (2004)	US	400	18-45yrs	27.8+/-7.7	6.6	-	-	-
Goodarzi et al. (2005)	US	156	34+/-8.6yrs	29.9+/-7.9	-	-	-	13(self-reported NIH)
Lo et al. (2006a)	US	644166	15-44yrs	50% of the population reported BMI 19.4%, 29.0% overweight for PCOS and control, respectively; 67.0%, 31.4% obese for PCOS and control, respectively	-	-	-	2.2(ICD-9)
Okoroh et al. (2012)	US	12171830	18-45yrs	0.6%, 1.3% obese for PCOS and control, respectively	1.11	1.59	1.20	-

Article	Country	Sample size	Age	BMI	Prevalence of PCOS (%)			
					NIH	**Rotterdam**	**AES**	**Other**
Christensen et al. (2013)	US	137502	15-19yrs	21.2% overweight, 57.3% obese for PCOS; 20.3% overweight, 16.5% obese for control	0.56	-	-	-
Sirmans et al. (2014a)	US	143413	15-45yrs	39.5%, 10.1% obese for PCOS and control, respectively	-	-	-	0.88(ICD-9)
Moran et al. (2010a)	Mexico	150	20-45yrs	27.5+/-4.6	6.0	6.6	-	-
Gabrielli and Aquino (2012)	Brazil	859	18-45yrs	Median (IQR) PCOS: 24.2(17.7-30.7) control: 24.1(18.1-30.1)	8.03	8.5	-	-
Faria et al. (2013)	Brazil	485	15-18yrs	Median(min-max) PCOS: 20.2(16.7-27.6) control: 20.4(16.3-27.1)	-	-	-	6.2(medical diagnosed PCOS)
Europe								
Michelmore et al. (1999)	UK	230	18-25yrs	Median(IQR) PCOS: 23.7(21.3-25.6) PCO only: 20.7(19.7-22.0) control: 22.3(20.4-24.7)	8.0	-	-	-

Article	Country	Sample size	Age	BMI	Prevalence of PCOS (%)			
					NIH	Rotterdam	AES	Other
Diamanti-Kandarakis et al. (1999)	Greece	192	17-45yrs	27.2+/-0.7, 28.7+/-2.7 for 1 feature group 28.9+/-1.6 for PCOS 25.9+/-0.6 for control	6.8	-	-	-
Asunción et al. (2000)	Spain	154	18-45yrs	23.8+/-3.2	6.5	-	-	-
Sanchón et al. (2012)	Spain	592	≥ 18yrs Median:17-33*	25+/-5, 24+/-5 for women from Spain and Italy, respectively	5.4	-	-	-
Lindholm et al. (2008)	Sweden	147	25-40yrs	29.5+/-5.7 for PCOS; 25.1+/-6.9, 28.0+/-4.4 for 1 feature group; 24.8+/-4.4 for control	-	-	-	4.8(self-reported NIH)
Lauritsen et al. (2014)	Denmark	447	20-40yrs	23.1+/-3.5	-	16.6	-	-
Asia								
Chen et al. (2008)	China	915	19-45yrs	20.9+/-3.0	-	2.4	2.2	-
Ma et al. (2010)	China	2111	19-45yrs	-	-	6.11	-	-
Li et al. (2013a)	China	15924	19-45yrs	22.2+/-4.2, 22.1+/-3.4 for PCOS and control, respectively	-	5.6	-	-

Article	Country	Sample size	Age	BMI	Prevalence of PCOS (%)			
					NIH	Rotterdam	AES	Other
Jiao et al. (2013)	China	1600	19-45yrs	25.8% obese for PCOS control not known	-	8.25	-	-
Zhuang et al. (2014)	China	1645	12-44yrs	-	7.1	11.2	7.4	-
Sung et al. (2010)	Korea	-	16-39yrs	-	4.4	6.3	5.1	-
Nidhi et al. (2011)	India	460	15-18yrs	14.3% overweight, 7.14% obese for PCOS; control not known	2.61	9.13		-
Gill et al. (2012)	India	1520	18-25yrs	21.7+/-5.5 for PCOS, 19.5+/-3.0 for control	3.7	-	-	-
Joshi et al. (2014)	India	600	15-24yrs	19.4+/-3.7	-	22.5	10.7	-
Kumarapeli et al. (2008)	Sri Lanka	2915	15-39yrs	Median(IQR) PCOS: 24.2(16.1-32.3) control not known	-	6.3	-	-
Vutyavanich et al. (2007)	Thailand	1095	18-40yrs	23.8+/-5.1 for PCOS, control not known	5.7	-	-	-
Middle East								
Musmar et al. (2013)	Iran	137	18-24yrs	22.2+/-3.1	7.3	-	-	-
Hashemipour et al. (2004)	Iran	1000	14-18yrs	9.1% obese	-	-	-	3(clinical PCOS)
Mehrabian et al. (2011)	Iran	820	17-34yrs	19% overweight 9% obese	7.0	15.2	7.92	-
Asgharnia et al. (2011)	Iran	1850	17-18yrs	21.1+/-3.6 for PCOS, control not known	11.34	-	-	-

Article	Country	Sample size	Age	BMI	Prevalence of PCOS (%)			
					NIH	Rotterdam	AES	Other
Tehrani et al. (2011)	Iran	929	18-45yrs	Median (IQR) PCOS: 27.0(25.8-37.3), 25.0 (21.6-29.2), 25.6 (23.0-31.2) control: 26.4(23.1-29.4)	7.1	14.6	11.7	-
Esmaeilzadeh et al. (2014)	Iran	1549	16-20yrs	21.9+/-4.1	-	-	-	8.3 (criteria not stated)
Rashidi et al. (2014)	Iran	602	18-45yrs	36.9% overweight, 22.1% obese	4.8	14.1	12.0	-
Yildiz et al. (2012)	Turkey	392	18-45yrs	Obese: 10.2% for total 25% for NIH 15.4% for Rotterdam 15% for AES	6.1	19.9	15.3	-
Al Khaduri et al. (2014)	Oman	3644	12-45yrs	25.5+/-4.7	-	7.0	-	-
Attlee et al. (2014)	UAE	50	17-23yrs	22.9+/-3.5	-	-	-	20(criteria not stated)
Oceania								
Lowe et al. (2005)	Australia	100	-	23.4+/-3.8 for women reported (11%)	-	12	-	-
March et al. (2010)	Australia	728	27-34yrs	Median 25.7 (22.5-30.9)	8.7	11.9	10.2	-
Boyle et al. (2012)	Australia	248	15-45yrs	Median 27.0 (22.5-32.6)	15.3	-	-	-
Joham et al. (2013)	Australia	8612	28-33yrs	28.0+/-7.2 25.1+/-5.6 for PCOS, control, respectively	-	-	-	5.8(self-reported PCOS)

Appendix B

Table 2. Summary of studies assessing impaired glucose tolerance (IGT) and type 2 diabetes (DM2) in women with and without PCOS.

Study	Total case/control	Age(yrs)	BMI(kg/m^2)	PCOS definition	IGT/DM2 definition
Europe					
Rajkhowa et al. (1996)	72/39	Mean: cases: 26 controls: 30	Mean: cases: 31.6 controls: 25.9	Rotterdam	WHO
Ciampelli et al. (1998)	35/11	cases: 21-36 controls: 28-30	Not stated	NIH	NDG
Cibula et al. (2000)	28/752	45-59	Mean(SD): cases: 28.0(4.2) controls: 28.2(5.2)	NIH	Fasting glucose $\geq$ 7mmol/L current medical treatment
Yaralı et al. (2001)	30/30	Mean(SD): cases: 27.9(6.1) controls: 31.4(6.5)	Mean(SD): cases: 27.3(6.0) controls: 25.0(3.3)	NIH	WHO
Faloia et al. (2004)	50/20	Mean(SD): cases: 23(5) controls: 12 lean, 8 overweight	cases: 23 lean, 27 overweight controls: 12 lean, 8 overweight	NIH	WHO

Study	Total case/control	Age(yrs)	BMI(kg/m^2)	PCOS definition	IGT/DM2 definition
Diamanti-Kandarakis et al. (2005)	29/22	Mean(SD): cases: 25.8(5.0) controls: 28.1(4.0)	Mean(SD): cases: 27.2(7.7) controls: 23.3(5.7)	NIH	OGTT, not stated
Álvarez-Blasco et al. (2006)	32/72	Mean(SD): cases: 26(7) controls: 32(8)	Mean(SD): cases: 34.8(6.6) controls: 35.2(6.2)	NIH	Not stated
Attaoua et al. (2008)	207/100	Mean(SD): cases: lean: 23.1(0.5) obese: 26.3(0.6) controls: 31.4(1.1)	Mean(SD): cases: 23.0(0.3) controls: 22.2(0.4)	Rotterdam	ADA(2006)
Alvarez-Blasco et al. (2009)	141/102	Not stated	Mean(SD): cases: 30(8) controls: 31(8)	NIH	ADA
Vrbikova et al. (2009)	228/643	cases: 18-39 controls: 20-39	cases: 121 obese or overweight, 107 lean controls: not stated	Rotterdam	WHO
Karaer et al. (2010)	31/31	18-45	Mean: 28.9	Rotterdam	OGTT
Luque-Ramírez et al. (2010)	112/86	Mean(SD): cases: 27(6) controls: 28(6)	Mean(SD): cases: 31(8) controls: 30(8)	NIH	ADA

Study	Total case/control	Age(yrs)	BMI(kg/m^2)	PCOS definition	IGT/DM2 definition
Pall et al. (2010)	75/23	14-45	Mean(SD): 54 obese, 52 lean controls: 22(2)	NIH	OGTT, glucose level 100-126mg/dL and/or 2-hr glucose level of 140-199mg/dL for IGT, glucose level $\geq$ 126mg/dL and/or 2-hr glucose level $\geq$ 200mg/dL
Amato et al. (2011)	125/144 220/144 177/144	13-43	Mean(SD): cases: 30.6(7.7) controls: 31.3(7.7)	NIH Rotterdam AES	OGTT, ADA
Hudecova et al. (2011a)	84/87	Age at follow-up: > 35yrs mean follow-up: 13.9yrs	Mean(SD) at follow-up: cases: 29.2(5.7), 26.2(5.2) controls: 25.6(4.2)	Rotterdam	WHO
Schmidt et al. (2011)	25/95	40-59 at baseline	Mean(SD) at follow-up: cases: 27.1(5.0) controls: 26.4(4.8)	Rotterdam	Medical or dietary treatment for diabetes and/or diagnosis of diabetes

Study	Total case/control	Age(yrs)	BMI(kg/m^2)	PCOS definition	IGT/DM2 definition
Celik et al. (2013)	252/117	Mean(SD): cases: 24.8(5.5) controls: 25.9(5.7)	Mean(SD): cases: 26.1(5.7) controls: 24.9(4.3)	Rotterdam	OGTT, ADA
Dalamaga et al. (2013)	316/102	18-35	Mean(SD): cases: 28.7(8.0), 24.9(6.3) controls: 24.6(4.9)	Rotterdam	WHO
Celik et al. (2014) Celik	84/45	Mean(SD) at baseline: cases: 24.7(6.5) controls: 27.4(5.6)	Mean(SD) at follow-up: cases: 27.5(6.9) controls: 24.6(5.0)	Rotterdam	ADA
North America					
Dunaif et al. (2001)	14/12	19-41	Mean(SD): cases: 40.5(1.9) controls: 40.5(1.6)	NIH	WHO
Phy et al. (2004)	7/18	Mean(SD): cases: 30.9(4.5) controls: 31.1(2.6)	Mean(SD): cases: 30.9(10.2) controls: 25.0(5.8)	NIH	ADA
Apridonidze et al. (2005)	106/-	cases: 29.1-31 controls: not stated	cases: 33.7-39.2 controls: not stated	NIH	ADA 1997
Legro et al. (2005)	71/23	Mean: cases: 27, 29.6 controls: 36.2	Mean(SD): cases: 35.7(7.5), 38.7(8.3) controls: 29.3(6.5)	NIH	ADA, WHO
Boudreaux et al. (2006)	97/95	Mean: cases: 38.0 controls: 40.0	Mean: cases: 31.6 controls: 26.2	NIH	Fasting glucose $\geq$ 7.0mmol/L or doctor diagnosis

Study	Total case/control	Age(yrs)	BMI(kg/m^2)	PCOS definition	IGT/DM2 definition
Leibel et al. (2006)	36/21	12-19	Mean(SD): cases: 30.3(7.6), 37.9(8.2) controls: 24.9(4.5)	Rotterdam	ADA
Lo et al. (2006b)	11035/55175	Mean: cases: 30.7 controls: 30.8	cases: 67% obese controls: 31.4% obese	NIH	Hospital discharge or $\geq$ 2 outpatient diagnosis
Shaw et al. (2008)	104/286	Mean(SD): cases: 62.5(10) controls: 65.8(9)	Mean(SD): cases: 31.1(7) controls: 28.4(6)	NIH	Fasting glucose $\geq$ 7.8mmol/L or medication
Kawai et al. (2009)	185/120	Mean(SD): cases: 30.8(6.3), 28.4(6.1) controls: 41.6(10.6), 32.3(10.9)	Mean(SD): cases: 38.2(7.3), 36.3(8.0) controls: 31.6(8.4), 26.8(6.8)	NIH	OGTT, ADA
Nur et al. (2009)	101/40	10-21	Mean(SD): cases: 33.2(5.9) controls: 32.4(5.3)	Rotterdam	ADA
Wang et al. (2011)	53/1074	20-32 at baseline 34-46 at follow-up	At baseline: cases: 22.6% obese controls: 19.7% obese	NIH	Fasting plasma glucose $\geq$ 126mg/dL or use of diabetic medications
Wickham et al. (2011)	13/13	18-40	Mean: cases: 30.7 controls: 30.0	NIH	OGTT, ADA
Roe et al. (2013)	148/57	cases: 13-20 controls: 14-20	cases: 38% obese controls: 9%	AES (majority also met NIH)	Glucose $\geq$ 100mg/dL
Sirmans et al. (2014b)	1689/5067	15-45	cases: 39.5% obese controls: 10.1% obese	NIH (ICD-9)	ICD-9

Study	Total case/control	Age(yrs)	BMI(kg/m^2)	PCOS definition	IGT/DM2 definition
South America					
Dos Reis et al. (1995)	29/19	18-37	cases: 14 lean, 15 obese controls: 10 lean, 9 obese	NIH	WHO
Sir-Petermann et al. (2004)	146/97	cases: 15-35 controls: 15-35	Median: cases: 29.0 controls: 24.8	NIH	WHO
Echiburú et al. (2008)	159/93	15-36	Mean(SD): cases: 28.7(6.1) controls: 25.5(4.3)	NIH	WHO
Márquez et al. (2008)	50/70	cases: 16-43 controls: 20-45	Mean(SD): cases: 33.3(8.1) controls: 23.4(2.7)	NIH	Fasting glucose $\geq$ 7.8mmol/L or history of DM2
Wiltgen and Spritzer (2010)	195/25 240/25	14-35	Mean(SD): cases: 31.1(8.0), 27.0(6.4) controls: 27.0(3.6)	NIH Rotterdam	OGTT
East Asia					
Sawathiparnich et al. (2005)	6/6	Mean: cases: 14.1 controls: 14.5	Mean: cases: 37.4 controls: 34.2	NIH	ADA
Lee et al. (2009)	194/162	cases: 16-41 controls: 20-39	Mean(SD): cases: 26.8(4.5), 22.3(3.8) controls: 22.2(2.7), 23.7	Rotterdam	WHO (1999)
Wei et al. (2009)	356/974	19-44	Mean(SD): cases: 22.4(4.2) controls: 20.6(2.4)	NIH	OGTT, WHO

Study	Total case/control	Age(yrs)	BMI(kg/m^2)	PCOS definition	IGT/DM2 definition
Huang et al. (2010)	90/40 128/40	Median: cases: 18 controls: 19	Median: NIH cases: 20.2 Rotterdam cases: 20.0 controls: 19.6	NIH Rotterdam	ADA
Zhao et al. (2010)	647/717 818/717 678/717	cases: 18-41 controls: 20-45	NIH cases: 21.8% obese Rotterdam cases: 19.7% obese AES cases: 21.1% obese controls: 8.8% obese	NIH Rotterdam AES	ADA
Li et al. (2012)	56/26	15-19	Mean(SD): cases: 22.0(4.9) controls: 19.7(1.6)	Rotterdam	OGTT 2-h PG $\geq$ 140mg/dL but $\leq$ 200mg/dL
LIANG et al. (2012)	220/70	Mean(SD): cases: 26.9(5.8) controls: 28.3(4.4)	Mean(SD): cases: 25.9(6.1) controls: 28.3(4.4)	Rotterdam	OGTT
HU et al. (2014)	234/39	Mean(SD): cases: 29.2(3.3), 30.6(3.4) 31.5(4.0) controls: 31.8(4.3)	Mean(SD): cases: 26.6(10.9), 26.4(4.6) 25.5(9.1) controls: 21.3(3.2)	Rotterdam	OGTT
South Asia					
Bhattacharya and Jha (2011)	264/116	cases: 16-39 controls: 16-35	Mean(SD): cases: 26.5(4.7), 28.6(4.2) controls: not stated	NIH	WHO
Guleria et al. (2014)	50/50	18-35	Mean(SD): cases: 24.6(4.9) controls: 23.9(3.8)	NIH	OGTT

Study	Total case/control	Age(yrs)	BMI(kg/m^2)	PCOS definition	IGT/DM2 definition
Middle East					
Moini and Eslami (2009)	273/276	Mean(SD): cases: 27.9(4.2) controls: 31.1(5.8)	Mean(SD): cases: 27.9(2.8) controls: 25.6(4.4)	Rotterdam	self-reported DM2
Australia					
Chan et al. (2013)	109/133	18-60	Mean(SD): cases: 29.5(1.4), 29.8(1.1) 32.1(1.5), 35(1.7) controls: 22.2(0.6), 24.5(1.0) 27.7(0.9), 27.8(0.8)	Rotterdam	Self-reported DM2 history

Appendix C

Table 3. Summary of studies assessing metabolic syndrome (MetS) in women with and without PCOS.

Study	Total case/control	Age(yrs)	BMI(kg/m^2)	PCOS definition	MetS definition
Europe					
Faloia et al. (2004)	50/20	Mean(SD): cases: 23(5), 21(5) controls: age-matched	Mean: cases: 22 for lean, 32 for overweight controls: 20 for lean, 37 for overweight	NIH	ATP III
VrbÍková et al. (2004)	69/73	Mean(SD): cases: 25.2(4.7) controls: age-matched	Median: cases: 23.0 controls: 21.9	Rotterdam	ATP III
Álvarez-Blasco et al. (2006)	32/72	Mean(SD): cases: 26(7) controls: 32(8)	Mean(SD): cases: 34.8(6.6) controls: 35.2(6.2)	NIH	ATP III
Carmina et al. (2006)	282/85	cases: 18-40 controls: 18-35	Mean(SD): cases: 27.2(0.3) controls: 23.3(0.6)	Rotterdam	ATP III, WHO

Study	Total case/control	Age(yrs)	BMI(kg/m^2)	PCOS definition	MetS definition
Attaoua et al. (2008)	107/100	Mean(SD): cases: lean: 23.1(0.5) obese: 26.3(0.6) controls: 31.4(1.1)	Mean(SD): cases: lean: 23.0(0.3) obese: 34.9(0.4) controls: 22.2(0.4)	Rotterdam	ATP III
Fulghesu et al. (2010)	71/94	cases: 13-18 age-matched controls	Mean(SD): cases: 24.0(0.7) controls: 22.6(0.5)	Rotterdam	Adolescent-specific 3 out of 5
Gambineri et al. (2009)	200/200	14-49	Mean(SD): cases: 31.7(7.4) controls: 31.0(7.6)	Rotterdam	ATP III, IFD, AHA
Tan et al. (2010)	186/73	Mean(SD): cases: 28.4(6.7) controls: 28.4(8.4)	Mean(SD): cases: 31.5(8.3) controls: 23.0(3.1)	AES	Not stated
Veltman-Verhulst et al. (2010)	456/240	Mean(SD): case: 30.6(4.4) controls: 29.0(4.9)	Mean(SD): cases: 26.6(7.0) controls: 26.1(5.5)	Rotterdam	ATP III
Hudecova et al. (2011a)	84/87	Age at follow-up: > 35 mean follow-up: 13.9yrs	Mean at follow-up: cases: 29.2(5.7), 26.2(5.2) controls: 25.6(4.2)	Rotterdam	WHO
Amato et al. (2011)	125/144 220/144 177/144	13-43	Mean(SD): cases: 30.6(7.7) controls: 31.3(7.7)	NIH Rotterdam AES	ATP III
Rizzo et al. (2011)	350/185	18-40	Mean(SD): cases: 27(7) controls: 23(5)	AES	AHA
Vrbíková et al. (2011)	43/48	Mean(SD): cases: 16.8(2.0) controls: 17.5(1.9)	Mean(SD): cases: 23.6(6.2) controls: 23.4(5.3)	Rotterdam	IDF adolescent

Study	Total case/control	Age(yrs)	BMI(kg/m^2)	PCOS definition	MetS definition
Panidis et al. (2013)	905/277 1223/277	≤ 39	Mean(SD): NIH cases: 28.1(7.0) Rotterdam cases: 27.5(6.9) controls: 26.6(6.7)	NIH Rotterdam	ATP III, AHA, IDF, Joint definition
Albu et al. (2015)	398/126	Mean(SD): cases: 24(7) controls: 28(10)	Mean(SD): cases: 26.1(10.9) controls: 25.6(12)	Rotterdam	ATP III
North America					
Glueck et al. (2003)	138/1887	Mean(SD): cases: 31(7) controls: not stated	Not stated	NIH	ATP III
Apridonidze et al. (2005)	106/-	Mean: cases: 29.1-31 controls: not stated	Mean: cases: 33.7-39.2 controls: not stated	NIH	ATP III: BMI 31 as a surrogate for WC 88cm
Dokras et al. (2005)	127/1887	cases: 18-49 controls: not stated	Not stated	NIH	WHO
Leibel et al. (2006)	36/21	12-19	Mean(SD): cases: 30.3(7.6), 37.9(8.2) controls: 24.9(4.5)	Rotterdam	Modified ATP III, modified AHA
Coviello et al. (2006)	49/165	cases: 14-19 controls: 12-19	Mean(SD): cases: 32(9) controls: 23(5)	NIH	Modified ATP III
Welt et al. (2006)	418/64	cases: 18-45 controls: 18-37	Mean(SD): cases: 32.0(8.6), 27.0(6.8), 24.7(5.4) controls: 27.3(6.8)	Rotterdam	Not stated

Study	Total case/control	Age(yrs)	BMI(kg/m^2)	PCOS definition	MetS definition
Shroff et al. (2007a)	24/24	21-50	Mean(SD): cases: 36(5.4) controls: 35(3.3)	NIH	AHA
Shroff et al. (2007b)	258/110	18-45	Mean(SD): cases: 35(9), 36.6(10.9) 35.5(9.1), 29.4(9.3) controls: 29(6)	Rotterdam	3 out of 5: BMI>30, TG≥1.7mmol/L, HDL-C<1.3mmol/L BP≥130/85, FPG≥6.1mmol/L or DM2 presence
Shaw et al. (2008)	104/286	Mean(SD): cases: 62.5(10) controls: 65.8(9)	Mean(SD): cases: 31.1(7) controls: 28.4(6)	NIH	ATP III
Gulcelik et al. (2008)	60/60	Mean: cases: 24.6 controls: 26.1	Mean: cases: 28 controls: 26.7	Rotterdam	ATP III
Nandalike et al. (2012)	28/28	13-18	Mean(SD): cases: 44.8(8.8) controls: 40.2(4.7)	Rotterdam	Weiss criteria
Roe et al. (2013)	148/57	cases: 13-20 controls: 14-20	cases: 38% obese controls: 9% obese	AES (majority also met NIH)	Modified cook criteria
South America					
Wiltgen and Spritzer (2010)	195/25 240/25	14-35	Mean(SD): cases: 31.8(8.0), 27.0(6.4) controls: 27.0(3.6)	NIH Rotterdam	Not stated
Melo et al. (2011)	150/146 226/146 175/146	Mean(SD): cases: 26.6(5.1), 26.2(5.7), 25.9(5.3), 27(4.5) controls: 28.9(0.5)	Mean(SD): cases: 31.3(8.7), 28.5(8.8), 34.1(8.7), 29.5(9.9) controls: 24.4(4.9)	NIH Rotterdam AES	Modified ATP III

Study	Total case/control	Age(yrs)	BMI(kg/m^2)	PCOS definition	MetS definition
Ramos and Spritzer (2015)	199/99	Mean(SD): 22.7(7.1)	Mean(SD): cases: 29.6(6.4) controls: 27.0(6.0)	Rotterdam	Joint Interim criteria
East Asia					
Cheung et al. (2008)	288/98	Mean(SD): cases: 30.2(6.4) controls: 33.4(5.9)	Mean(SD): cases: 25.8(5.9) controls: 21.3(2.6)	Rotterdam	ATP III, AHA
Zhang et al. (2009)	248/85 719/85 344/85	Mean(SD): cases: 26(4.9), 25(5.1), 27(3.7), 26(4.5) controls: (27(5.3))	Mean(SD): cases: 36.5(8.6), 35.8(9.3) 30.9(8.3), 28.6(6.5) controls: 27.3(5.2)	NIH Rotterdam AES	IDF
Huang et al. (2010)	90/40 128/40	Median: cases: 18 controls: 19	Median: NIH cases: 20.2 Rotterdam cases: 20.0 controls: 19.6	NIH Rotterdam	IDF
Zhang et al. (2009)	254/342 406/342 279/342	17-40	Mean(SD): cases: 22.9(4.2) controls: 20.8(2.6)	NIH Rotterdam AES	ATP III
LIANG et al. (2012)	220/70	Mean(SD): cases: 26.9(5.8) controls: 28.3(4.4)	Mean(SD): cases: 25.9(6.1) controls: 23.4(5.2)	Rotterdam	ATP III
Li et al. (2013b)	397/2732 833/2732 708/2732	19-45	Mean(SD): cases: 22.2(4.2) controls: 22.1(3.4)	NIH Rotterdam AES	ATP III
Chen et al. (2015)	224/198	Mean(SD): 27.5(6.4)	Mean(SD): 24.6(5.8)	AES	ATP III
HU et al. (2014)	234/39	Mean(SD): cases: 29.2(3.3), 30.6(3.4), 31.5(4.0) controls: 31.8(4.3)	Mean(SD): cases: 26.6(10.9), 26.4(4.6), 25.5(9.1) controls: 21.3(3.2)	Rotterdam	China Diabetes Association criteria

Study	Total case/control	Age(yrs)	BMI(kg/m^2)	PCOS definition	MetS definition
South Asia					
Bhattacharya and Jha (2011)	51/45	Mean(SD): cases: 17.1(1.6) controls: 16.7(1.7)	Mean(SD): cases: 26.4(4.6) controls: 25.4(4.6)	AES	Joint Interim criteria
Wijeyaratne et al. (2010)	395/205	Median: cases: 25 controls: 29	Median: cases: 25 controls: not stated	Rotterdam	ATP III
Karoli et al. (2013)	54/55	Mean(SD): cases: 28.5(6.2) controls: 27.8(7.5)	Mean(SD): cases: 27.2(5.4) controls: 26.8(6.7)	Rotterdam	ATP III
Guleria et al. (2014)	50/50	18-35	Mean(SD): cases: 24.6(4.9) controls: 23.9(3.8)	NIH	IDF
Thathapudi et al. (2014)	204/204	17-35	Mean(SD): cases: 29.3(4.2) for obese, 22.0(1.7) for lean controls: 23.4(3.2)	AES	ATP III, IDF
Middle East					
Çalışkan et al. (2007)	182/182	Mean(SD): cases: 23.2(4.5) controls: 23.6(4.6)	Mean(SD): cases: 25.0(5.1) controls: 23.5(2.9)	Rotterdam	APT III, WHO, AHA, IDF, Rotterdam
Yilmaz et al. (2011)	85/44 127/44 103/44	Mean(SD): cases: 25.5(6.4), 25.3(5.2), 25.5(6.5), 25.3(5.8)	Mean(SD): cases: 25.4(5.6), 23.4(4.7), 21.7(3.2), 25.3(5.2) controls: 22.0(2.5)	NIH Rotterdam AES	Study-specific 3 out of 5 criteria
Lankarani et al. (2009)	55/59	15-40	Mean(SD): cases: 24.9(5.3) controls: 21.6(2.6)	NIH	ATP III

Study	Total case/control	Age(yrs)	BMI(kg/m^2)	PCOS definition	MetS definition
Rahmanpour et al. (2012)	30/71	Mean(SD): cases: 17.7(1.0) controls: 17.7(1.3)	Mean(SD): cases: 23.4(3.1) controls: 21.2(2.9)	Rotterdam	IDF for adolescent
Hosseinpanah et al. (2014)	134/414	18-45	Mean(SD): cases: 26.8(5.8) controls: 26.8(4.9)	Rotterdam	Joint interim statement
Tehrani et al. (2014)	30/517 85/517 72/517	18-45	Mean(SD): cases: 25.4(5.0), 26.4(4.8), 27.2(4.4), 24.1(5.5) controls: 26.6(5.0)	NIH Rotterdam AES	Clinical diagnostic criteria applied for Iranian adult (3 out of 5)
Australia					
Cussons et al. (2008)	168/883	cases: 25-54 controls: 25-53	Mean(SD): cases: 32.3(8.1) controls: 25.8(5.8)	NIH	ATP III, IDF
Hart et al. (2011)	34/169 61/143	14	Mean(SD): 22.8(3.8)	NIH Rotterdam	ATP III, IDF, WHO, EGIR

Appendix D

Table 4. Summary of critical appraisal of studies reporting PCOS-related metabolic disorders using the Newcastle-Ottawa Scale for case-control studies.

Study	Selection (max 4 stars)	Comparability (max 2 stars)	Exposure (max 3 stars)
Dos Reis et al. (1995)	★		★★
Rajkhowa et al. (1996)	★★	★★	★★
Ciampelli et al. (1998)	★		★★
Cibula et al. (2000)	★	★★	★
Dunaif et al. (2001)	★★★	★★	★★
Yaralı et al. (2001)	★★	★	★★
Glueck et al. (2003)	★★	★	★★
Faloia et al. (2004)	★	★★	★★
Phy et al. (2004)	★	★★	★★
Sir-Petermann et al. (2004)	★★★		★★
Apridonidze et al. (2005)	★★		★
Legro et al. (2005)	★★★★		★★
Diamanti-Kandarakis et al. (2005)	★★	★★	★★
Dokras et al. (2005)	★★★	★★	★★
Sawathiparnich et al. (2005)	★	★★	★
VrbÍková et al. (2004)	★	★	★★
Álvarez-Blasco et al. (2006)	★★★	★	★★
Boudreaux et al. (2006)	★★★	★★	★★★
Carmina et al. (2006)	★★	★	★★
Coviello et al. (2006)	★★	★★	★★
Leibel et al. (2006)	★	★★	★★
Lo et al. (2006b)	★★	★	★★
Welt et al. (2006)	★	★★	★★
Çalışkan et al. (2007)	★★	★★	★★
Shroff et al. (2007a)	★★	★★	★★
Shroff et al. (2007b)	★★★	★★	★★
Cheung et al. (2008)	★★	★★	★★
Cussons et al. (2008)	★★	★★	★
Echiburú et al. (2008)	★★★	★	★★
Gulcelik et al. (2008)	★★	★★	★★
Márquez et al. (2008)	★★	★	★★

Study	Selection (max 4 stars)	Comparability (max 2 stars)	Exposure (max 3 stars)
Attaoua et al. (2008)	★★★	★	★★
Shaw et al. (2008)	★		★
Alvarez-Blasco et al. (2009)	★★★★	★	★★
Bhattacharya and Jha (2011)	★★		★★
Moini and Eslami (2009)	★★	★	★★
Fulghesu et al. (2010)	★★	★★	★★
Gambineri et al. (2009)	★★	★★	★★
Kawai et al. (2009)	★★		★★
Lankarani et al. (2009)	★★★	★	★★
Lee et al. (2009)	★★	★	★
Nur et al. (2009)	★★	★★	★★
Vrbikova et al. (2009)	★		★
Wei et al. (2009)	★★		★★
Zhang et al. (2009)	★★	★	★★
Huang et al. (2010)	★★★	★★	★★
Karaer et al. (2010)	★	★★	★★
Luque-Ramírez et al. (2010)	★★★	★★	★★
Pall et al. (2010)	★★★	★	★★
Tan et al. (2010)	★★★	★	★★
Veltman-Verhulst et al. (2010)	★★		★★
Wiltgen and Spritzer (2010)	★		★★
Zhao et al. (2010)	★★		★
Amato et al. (2011)	★★★	★★	★★
Hudecova et al. (2011a)	★★	★	★★
Hudecova et al. (2011b)	★★	★	★★
Melo et al. (2011)	★★★		★★
Rizzo et al. (2011)	★★	★★	★★
Schmidt et al. (2011)	★★	★	★★
VrbÍková et al. (2011)	★	★★	★★
Wang et al. (2011)	★★★	★★	★★★
Wickham et al. (2011)	★★	★★	★★
Wijeyaratne et al. (2010)	★★★	★★	★★
Yilmaz et al. (2011)	★★		★★
Karoli et al. (2013)	★	★★	★★
Li et al. (2012)	★★★	★	★★
LIANG et al. (2012)	★	★	★★
Nandalike et al. (2012)	★★★	★	★★
Rahmanpour et al. (2012)	★★		★★
Zhang et al. (2012)	★★★		★★
Dalamaga et al. (2013)	★★★		★★
Celik et al. (2013)	★★★		★★
Chan et al. (2013)	★	★	★★
Li et al. (2013b)	★★★	★	★★

Study	Selection (max 4 stars)	Comparability (max 2 stars)	Exposure (max 3 stars)
Panidis et al. (2013)	★★	★	★★
Roe et al. (2013)	★★★	★	★★
Albu et al. (2015)	★★★	★	★★
Celik et al. (2014)	★★		★★★
Chen et al. (2015)	★★★	★	★★
Guleria et al. (2014)	★★★	★★	★★
Hosseinpanah et al. (2014)	★★★		★★
HU et al. (2014)	★★	★	★★
Tehrani et al. (2014)	★★★	★	★★
Thathapudi et al. (2014)	★★★	★	★★
Sirmans et al. (2014b)	★★	★	★★
Ramos and Spritzer (2015)	★★★	★	★★
Hart et al. (2011)	★★★		★

Appendix E

Table 5. Read codes to identify PCOS cases.

medcode	description
C164.12	Stein - Leventhal syndrome
C165.00	Polycystic ovarian syndrome
7E25300	Endoscopic drilling of ovary
K591100	Oligomenorrhoea
K591200	Primary oligomenorrhoea
K591300	Secondary oligomenorrhoea
K594.00	Irregular menstrual cycle
K594z00	Irregular menstrual cycle NOS
22D8.00	O/E - hirsutism
C161000	Hypersecretion of ovarian androgen
M240.00	Alopecia
M240000	Alopecia unspecified
M240200	Male pattern alopecia
M240300	Frontal alopecia of women
M240400	Premature alopecia
M240z00	Alopecia NOS
M241.00	Hirsutism - hypertrichosis
M260.00	Acne varioliformis
M260000	Acne frontalis
M260z00	Acne varioliformis NOS
M261.00	Other acne
M261000	Acne vulgaris
M261100	Acne conglobata
M261600	Cystic acne
M261A00	Pustular acne
M261E00	Acne excoriee des jeunes filles
M261F00	Acne fulminans
M261G00	Acne agminata
M261J00	Acne necrotica
M261K00	Acne keloidalis
M261X00	Acne, unspecified

medcode	description
M261z00	Other acne NOS
Myu6300	[X]Other androgenic alopecia
Myu6800	[X]Other acne
Myu6F00	[X]Acne, unspecified
4473	Serum testosterone
4473100	Serum testosterone level abnormal
4474	Free androgen index
4474100	Free androgenic index abnormal
447G.00	Plasma testosterone level
447H.00	Androgen level
4Q26.00	Dihydrotestosterone level
4Q2E.00	Free testosterone level
4Q2F.00	Calculated free testosterone
ZRBs.00	Ferriman and Galwey score
K53..11	Ovarian cysts
K532.00	Other ovarian cysts
K532z00	Ovarian cyst NOS
Kyu9500	[X]Other and unspecified ovarian cysts
B540.00	Malignant neoplasm of adrenal gland
B540000	Malignant neoplasm of adrenal cortex
B540100	Malignant neoplasm of adrenal medulla
B540z00	Malignant neoplasm of adrenal gland NOS
B587.00	Secondary malignant neoplasm of adrenal gland
B7H0.00	Benign neoplasm of adrenal gland
B8yy100	Carcinoma in situ of adrenal gland
B922.00	Neoplasm of uncertain behaviour of adrenal gland
BB5h.00	[M]Adrenal cortical tumours
BB5h000	[M]Adrenal cortical adenoma NOS
BB5h100	[M]Adrenal cortical carcinoma
BB5h300	[M]Adrenal cortical adenoma, heavily pigmented variant
BB5h400	[M]Adrenal cortical adenoma, clear cell type
BB5h500	[M]Adrenal cortical adenoma, glomerulosa cell type
BB5h600	[M]Adrenal cortical adenoma, mixed cell type
BB5hz00	[M]Adrenal cortical tumours NOS
BBCF.00	[M]Adrenal rest tumour
BBD7.00	[M]Extra-adrenal paraganglioma, NOS
C15..00	Disorders of adrenal glands
C153.00	Other corticoadrenal overactivity
C155000	Adrenal medullary insufficiency
C15y.00	Other specified adrenal disorders
C15yz00	Other specified adrenal disorder NOS
C15z.00	Adrenal gland disorder NOS
Cyu4A00	[X]Other specified disorders of adrenal gland

medcode	description
PK1..00	Anomalies of adrenal gland
PK10.00	Aberrant adrenal gland
PK12.00	Accessory adrenal gland
PK13.00	Hypoplasia of adrenal gland
PK14.00	Ectopic adrenal gland
PK1y.00	Other specified anomalies of adrenal gland
PK1y000	Congenital cyst of adrenal gland
PK1yz00	Other congenital anomaly of adrenal gland NOS
PK1z.00	Anomalies of adrenal gland NOS
2226.11	O/E - cushingoid facies
C150.00	Cushing's syndrome
C150000	Idiopathic Cushing's syndrome
C150100	Iatrogenic Cushing's syndrome
C150111	Drug-induced Cushings syndrome
C150200	Pituitary dependent Cushing's syndrome
C150300	Ectopic ACTH secretion causing Cushing's syndrome
C150500	Alcohol-induced pseudo-Cushing's syndrome
C150z00	Cushing's syndrome NOS
Cyu4500	[X]Other Cushing's syndrome
F395100	Myopathy due to Cushing's syndrome
C150400	Nelson's syndrome
BB5y400	[M]Prolactinoma
C131000	Hyperprolactinaemia
C134011	Hypoprolactinaemia
B542.00	Malignant neoplasm pituitary gland and craniopharyngeal duct
B542000	Malignant neoplasm of pituitary gland
B542z00	Malig neop pituitary gland or craniopharyngeal duct NOS
B7H2.00	Benign neoplasm of pituitary gland and craniopharyngeal duct
B7H2.11	Pituitary adenoma
B7H2000	Benign neoplasm of pituitary gland
B7H2z00	Benign neoplasm of pituitary and craniopharyngeal duct NOS
B8yy300	Carcinoma in situ of pituitary gland
B920.00	Neop uncertain behaviour pituitary and craniopharyngeal duct
B920000	Neoplasm of uncertain behaviour of pituitary gland
B920z00	Neop uncertain behaviour pituitary and craniopharyngeal NOS
BB5V.00	[M]Pituitary adenomas and carcinomas
BB5Vz00	[M]Pituitary adenoma or carcinoma NOS
C13..00	Disorders of pituitary gland and its hypothalamic control
C131.00	Other anterior pituitary hyperfunction
C134.00	Other anterior pituitary disorder
C134z00	Other anterior pituitary disorder NOS
C134z11	Anterior pituitary hormone deficiency NEC

C137.00	Iatrogenic pituitary disorders
C137z00	Iatrogenic pituitary disorder NOS
C13z.00	Pituitary disorders NOS
Cyu4400	[X]Other disorders of pituitary gland
Cyu4M00	[X]Hyperfunction of pituitary gland, unspecified
K5B1.00	Female infertility of pituitary - hypothalamic origin
K5B1000	Primary pituitary - hypothalamic infertility
K5B1z00	Female infertility of pituitary - hypothalamic cause NOS
PK24.00	Anomalies of pituitary gland
PK24000	Aberrant pituitary gland
PK24z00	Anomaly of pituitary gland NOS

Appendix F

Sample JAGS codes of Bayesian models for evidence synthesis

```
### Prevalence estimation for PCOS population with different ethnic backgrounds
   model{
    for( i in 1:I){
     x[i]~dbin(betau[i],m[i])
     logit(betau[i])<- gamma[i]
     gamma[i]~dnorm(mu.gamma,tau.gamma)
     }

    # prior distribution
    mu.gamma~dnorm(-2.75,6.25)

    # uniform distribution for tau.gamma
    sigma.gamma~dunif(0,2)
    tau.gamma<-pow(sigma.gamma,-2)

    # halfcauchy (as replacement prior for tau.gamma for sensitivity analysis)
    # sigma.gamma <- abs(z.gamma)/pow(epsilon.gamma,0.5)
    # z.gamma ~ dnorm(0,tau.z.gamma)
    # epsilon.gamma ~ dgamma(0.5,0.5)
    # tau.z.gamma <- pow(B.gamma,-2)
    # B.gamma ~ dunif(0,0.5)
    # tau.gamma<-pow(sigma.gamma,-2)
    # tau.gamma<-pow(sigma.gamma,-2)

    p<-exp(mu.gamma)/(1+exp(mu.gamma))

  }

### Meta-regression model for the evaluation of relative risk of metabolic disorders
    among women with PCOS and controls
   model{
    for (s in 1:S){
     pcos0evt[s]~dbin(pi0[s],pcos0tl[s])
     pcos1evt[s]~dbin(pi1[s],pcos1tl[s])
     logit(pi0[s])<-alpha+X1[s]*beta0+X2[s]*beta1+X3[s]*beta2+X4[s]*beta3
     logit(pi1[s])<-alpha+X1[s]*beta0+X2[s]*beta1+X3[s]*beta2+X4[s]*beta3+delta[s]

     delta[s]~dnorm(mu.delta,tau.delta)
     beta0~dnorm(1.702,0.01)
     beta1~dnorm(1.296,0.01)
     beta2~dnorm(1.058,0.01)
     beta3~dnorm(0.967,0.01)
     alpha~dnorm(-3.585,0.01)
     gamma[s]<-exp(delta[s])
     }

    # Prior distribution
    mu.delta~dnorm(0,2)
    sigma.delta~dunif(0,1)
```

```
tau.delta<-pow(sigma.delta,-2)

# student-t distribution (as replacement prior for mu.delta for sensitivity analysis)
# mu.delta~dt(0,0.5,v)
# v~dunif(0,8)

rho<-exp(mu.delta)
p.harm<-1-step(mu.delta)

}
```

Appendix G

Diagnostic plots for the assessment of model convergence with Gelman-Rubin statistics provided for all model parameters.
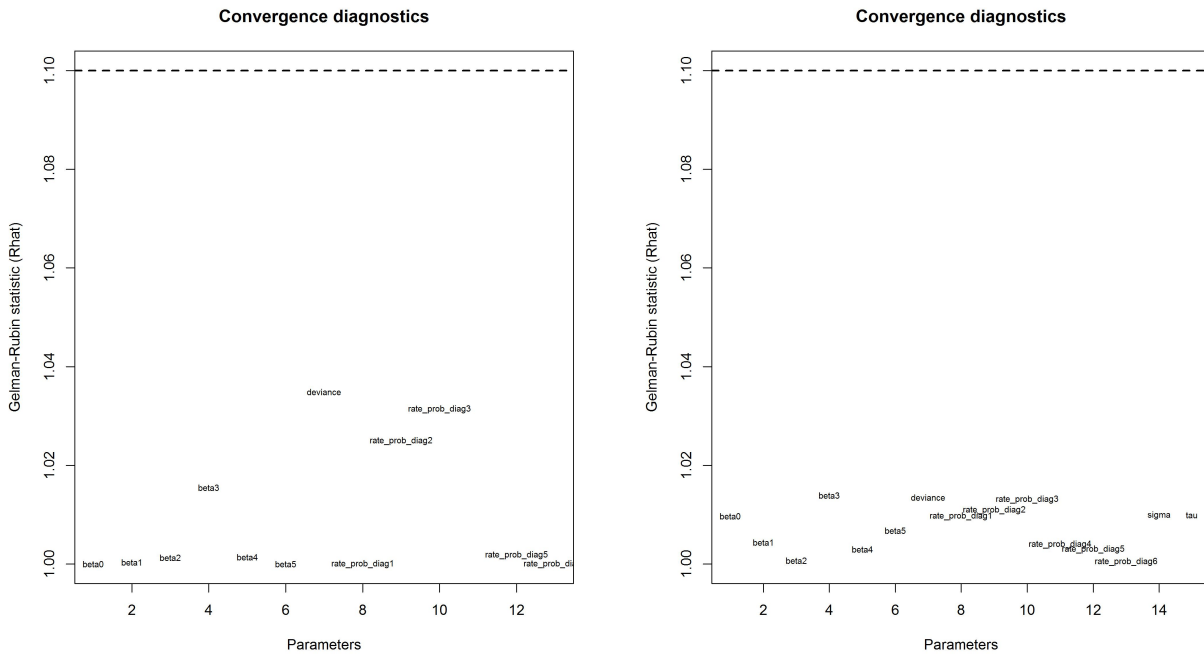

Figure G.1: Convergence assessment for model estimating the incidence rates from probable to diagnosed PCOS for standard model (left) and hierarchical model (right).
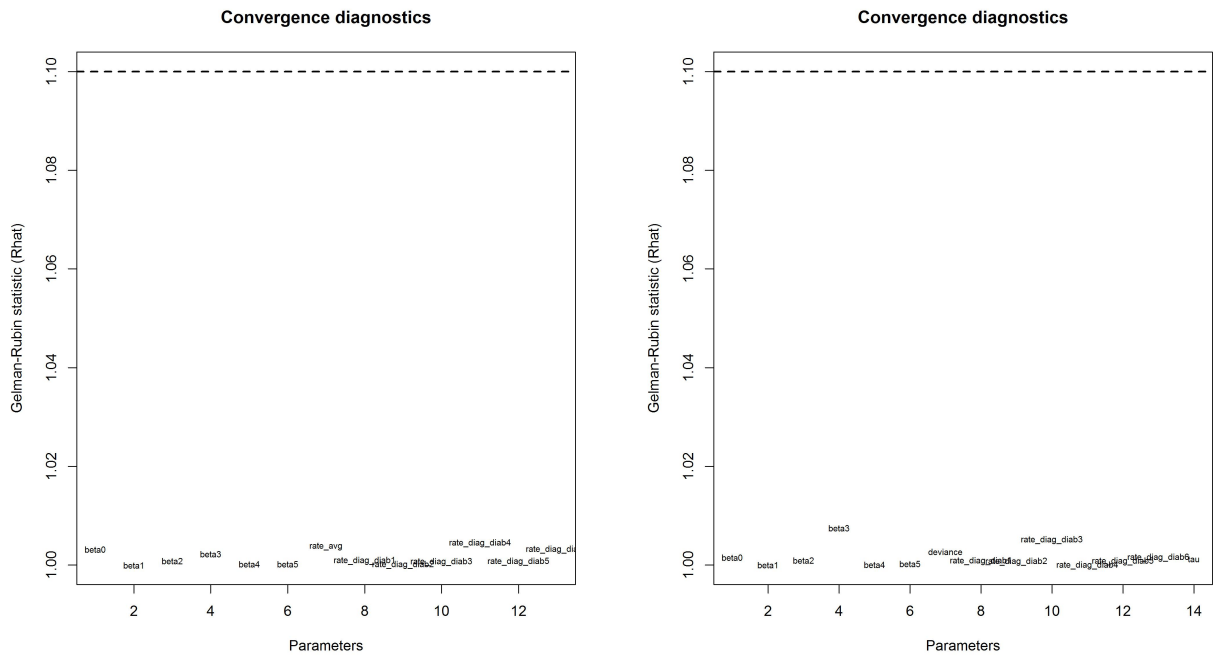
Figure G.2: Convergence assessment for model estimating the incidence rates from diagnosed PCOS to diabetes for standard model (left) and hierarchical model (right).
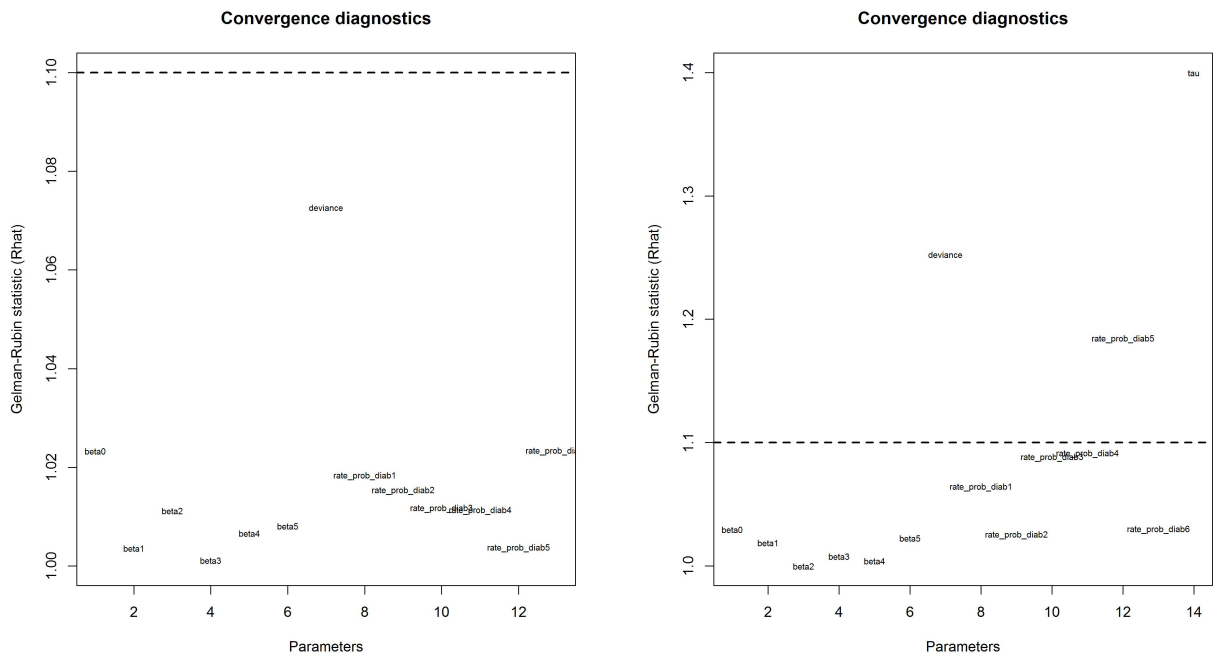

Figure G.3: Convergence assessment for model estimating the incidence rates from probable PCOS to diabetes for standard model (left) and hierarchical model (right).