The Effectiveness of Exercise in the Treatment of Polycystic Ovary Syndrome

Christopher Simon Kite

Doctor of Philosophy

ASTON UNIVERSITY

November 2020

Aston University

Christopher Simon Kite

Doctor of Philosophy - November 2020

Thesis Summary

Polycystic ovary syndrome (PCOS) is the most common endocrinopathy in reproductive-aged women. The clinical and biochemical characteristics of PCOS typically include cystic ovaries, ovulatory dysfunction, and hyperandrogenaemia. PCOS is also associated with metabolic and psychological morbidity. Typically, management of PCOS focusses upon weight loss through positive lifestyle changes, namely caloric restriction and increasing physical activity (PA). Exercise is an effective treatment for a range of populations; despite its recommendation in PCOS, little is known about its effectiveness at improving health in this population. Accordingly, three studies were completed to investigate the effect of PA in the management of PCOS.

Studies of women with PCOS that compared exercise (and diet) interventions to control conditions were meta-analysed in a systematic review. Exercise interventions improved insulin resistance, lipids, and cardiorespiratory fitness. However, the magnitude of these changes was small and the certainty of the evidence was graded as low or very low. A need for rigorously designed and sufficiently powered studies that address this question was highlighted.

In study 2, despite no differences in PA, women with PCOS were found to be more overweight, and have poorer self-esteem and quality of life (QoL) than women without PCOS. Self-esteem, BMI and a PCOS diagnosis impaired QoL, whereas PA appeared to have no effect.

Study 3 also reported less-favourable health, independent of BMI, in women with PCOS compared to controls. Cluster analysis was completed, and a larger proportion of women with PCOS were assigned to the poorer health cluster; this cluster was also less active. Furthermore, women who were more active, and spent less time sitting, had more favourable health.

In conclusion, this PhD highlights a lack of high-quality studies to investigate the role of PA in women with PCOS; this should be a research priority. However, women with PCOS who are more active, and spend less time sitting have reduced cardiovascular risk, which supports current treatment recommendations.

Key words: Polycystic ovary syndrome, physical activity, women's health, cardiovascular risk.

Acknowledgements

I would like to express my gratitude to my supervisory team, Dr James Brown, Dr Ioannis Kyrou, and Dr Ian Lahart. You have all provided me with support and guidance from the beginning of this process. Without your expertise and encouragement, it is unlikely that these projects could have been completed. Your individual inputs have been invaluable to this PhD and helped to develop my understanding of the research process and the wider academic sector. Thank you all.

The wider academic community has also supported me. Accordingly, I would like to thank Prof Harpal Randeva, Prof David Broom, and Mr Islam Afzal for their contributions to the publication of Chapter 3. Thank you to Dr Sophie Williams and all at Verity for your support in participant recruitment for the second study and also to Dr Lou Atkinson for sharing her expertise of health psychology when it was needed most. I would also like to thank my colleagues at University Centre Shrewsbury, in particular Dr Sue Taylor and Prof John Buckley, for being a constant source of advice and for engineering the time for me to work on this thesis. A special mention also goes to Kiran Shabir and Amber Lewis for their friendship and empathy; good luck to you both in your PhD journeys and careers beyond.

I would also like to thank my family. First and foremost, to my wife Mrs Tracy Kite, thank you so much for your love and support over the past 10 years. Quite simply, I could not have done this without you. I also want to thank my sons; to Ethan, Logan and Phoenix, thank you for providing the most welcome distraction of all, and for your patience throughout this process. I love you all.

I also need to thank my Father, Mr Simon Kite, and Stepmother, Mrs Marcelle Kite for providing me with the opportunity to return to my studies as a mature student. You have unconditionally given me emotional (and financial) support when needed. I would not be where I am now without your intervention and continued interest in my happiness. Thank you both.

Finally, I would like to dedicate this thesis to my late mother, Ms Jacqueline Pennington-Carlisle. I hope that I have made you proud and, in the words of Roger Waters '*how I wish you were here'.*

Contents

List of abbreviations.

(ACTH) Adrenocorticotropic hormone

(AE-PCOS) Androgen Excess and PCOS Society

(AMH) Anti-Mullerian hormone

(AMS) Application management system

(ANOVA) Analysis of variance

(AQoL) Assessment of Quality of Life

(ASRM) American Society for Reproductive Medicine

(BDI-FS) Beck Depression Inventory-Fast Screen Questionnaire

(BMI) Body mass index

(CAD) Coronary artery disease

(CIs) Confidence intervals

(CRP) C-reactive protein

(CSV) Comma separated values

(CVD) Cardiovascular disease

(DBP) Diastolic blood pressure

(DHEA) Dehydroepiandrosterone

(DHHS) Department of Health and Human Services

(DSCF) Dwass-Steel-Critchlow-Fligner

(EBBS) Exercise Benefits/Barriers Scale

(EPIC) European Prospective Investigation into Cancer and Nutrition Study

(ESHRE) European Society for Human Reproduction and Embryology

(FAI) Free androgen index

(FBG) Fasting blood glucose

(FCS) Fully conditional specification

(FI) Fasting insulin

(FIML) Full information maximum likelihood

(FSH) Follicle stimulating hormone

(FWE) Family wise error

(GRADE) Grading of Recommendations Assessment, Development and Evaluation

(HBM) Health Belief Model

(HDL-C) High-density lipoprotein cholesterol

(HOMA-IR) Homeostatic model assessment of insulin resistance index

(HPA) Hypothalamic-pituitary-adrenal

(HR) Heart rate

(HR_{max}) Maximum heart rate

(HRQoL) Health related quality of life

(*hs*CRP) High-sensitivity C-reactive protein

(ICD) International Classification of Disease

(IGF) Insulin-like growth factor

(IGT) Impaired glucose tolerance

(IPAQ-LF) International Physical Activity Questionnaire Long Last Self-administered Format

(IQR) Interquartile range

(IR) Insulin resistance

(LDL-C) Low-density lipoprotein cholesterol

(LH) Luteinizing hormone

(LTPA) Leisure-time physical activity

(MAP) Mitogen activated pathway

(MCAR) Missing completely at random

(MD) Mean difference

(MET) Metabolic equivalent of Task

(MetS) Metabolic Syndrome

(MNAR) Missing not at random

(MTA) Material transfer agreement

(NCEP ATP) National Cholesterol Education Program, Adult Treatment Programme

(NHS) National Health Service

(NIH) National Institute of Health

(OGTT) Oral glucose tolerance test

(OR) Odds ratio

(PA) Physical activity

(PCO) Polycystic ovary

(PCOS) Polycystic ovary syndrome

(PCOS-Q) Polycystic ovary syndrome-questionnaire

(PhD) Doctor of philosophy

(PMM) Predictive mean matching

(PRISMA) Preferred Reporting Items for Systematic Reviews and Meta-Analyses

(RCT) Randomised controlled trial

(RHR) Resting heart rate

(RPE) Rate of perceived exertion

(RSE) Rosenberg Self-esteem

(SB) Sedentary behaviour

(SBP) Systolic blood pressure

(SD) Standard deviation

(SEE) Self-Efficacy for Exercise

(SF-12) 12-item Short Form Health Survey

(SF-36) 36-item Health Survey

(SHBG) Sex hormone binding globulin

(SI) Système International d'Unités

(SMD) Standardised mean difference

(SPSS) Statistical package for the Social Sciences

(T2DM) Type 2 diabetes Mellitus

(UK) United Kingdom

(VO_2 max) Maximum volume of oxygen

(WC) Waist circumference

(WHO) World Health Organisation

(WHR) Waist-to-hip-ratio

(WHtR) Waist-to-height-ratio

List of Tables and Figures.

Chapter 1: Introduction

1.1. Introduction

First described by Irving F. Stein and Michael L. Leventhal (Stein and Leventhal, 1935), polycystic ovary syndrome (PCOS) is regarded as the most common endocrinopathy in reproductive-aged women (Farquhar, 2007). Depending upon the diagnostic criteria used and the population studied (Teede, Deeks and Moran, 2010; Azziz *et al.* 2016), PCOS is thought to affect between 2% (Chen *et al.,* 2008) and 21% (Lizneva *et al.* 2016) of reproductive-aged women in the general population. Despite its high prevalence, PCOS remains an under-studied endocrinopathy with a highly complex pathophysiology which has not yet been fully clarified.

PCOS clinical and biochemical characteristics typically include reproductive complications (oligomenorrhea and amenorrhea); hyperandrogenaemia [either biochemical or/and clinical with manifestations such as hirsutism, acne and androgenic alopecia (Barber *et al.*, 2006)]; and polycystic ovarian (PCO) morphology as identified via ultrasound (Adams *et al.,* 1985). In addition to the aforementioned characteristics, PCOS is also associated with metabolic comorbidity, particularly overweightness and obesity (Lim *et al.*, 2012), insulin resistance and type 2 diabetes (T2DM) (Stepto *et al.,* 2013) and other cardio-metabolic risk factors (Cussons, Stuckey and Watts, 2007). These are so commonly linked with PCOS that there have been calls for them to be included in the diagnostic criteria/characterization of PCOS, whilst clinical guidelines recommend routine screening for such metabolic comorbidities (*e.g.,* T2DM) in women with PCOS (Legro, Castracane and Kauffman, 2004). Finally, it has been reported that women with PCOS are often also burdened with psychological morbidity; particularly, greater levels of anxiety and depression (Barry *et al.,* 2011b), reduced self-esteem (Tay *et al.,* 2019) and body dissatisfaction (Himelein and Thatcher, 2006), are all associated with PCOS. It is thought that this is caused, at least in part, by the clinical manifestations which are so often associated with PCOS (Barry, 2019).

Of note, it is the widely reported success of lifestyle interventions, typically including increased exercise/physical activity (PA) and/or dietary restriction, at improving many of these health-related outcomes in a wide range of populations (Paffenbarger *et al.*, 1986; Ekelund *et al.,* 2015; Arem *et*

al., 2015), which make lifestyle interventions a suitable candidate for improving the health of women with PCOS. Accordingly, the objective of this PhD is to investigate the role of exercise/PA in the management of women with PCOS.

1.2. PCOS Diagnostic Criteria

Although originally described in the 1930s (Stein and Leventhal, 1935), PCOS was first defined/classified at a National Institute of Child Health and Human Development conference for the United States National Institute of Health (NIH) in 1990. The resulting diagnostic criteria (Zawadski and Dunaif, 1992) include the presence of clinical and/or biochemical signs of hyperandrogenism and oligo-anovulation, after the exclusion of other androgen excess or related disorders (Table 1.1); there was little consideration for the PCO morphology. Polycystic ovaries are usually identified using ultrasound, and defined by the presence of at least 12 or more follicles (cysts) measuring 2-9 mm in diameter on one ovary, or an ovarian volume >10 mL in the absence of a dominant (>10 mm) follicle (Jonard *et al.,* 2007). However, the presence of ovarian cysts has also been identified in healthy women (Franks, 1995). In fact, the PCO morphology is a highly common ultrasonic feature that is present in ~20% of reproductive-aged females, many of whom exhibit no symptoms of the syndrome whatsoever (Polson, *et al.,* 1988).

Table 1.1. Diagnostic criteria/definitions for polycystic ovary syndrome (PCOS).

PCOS Criteria	NIH (1990)*	ESHRE/ASRM (2003)*	AE-PCOS Society (2006)*
	Must have both of the marked findings.	*Must have at least two of the marked findings.*	*Must have the marked finding with either/or both of the other two.*
Hyperandrogenism	✓	✓	✓
Oligo-anovulation	✓	✓	
PCO morphology		✓	

*In addition to the above criteria, PCOS diagnosis requires the exclusion of other androgen excess or related disorders, including: hyperprolactinaemia, thyroid dysfunction, adrenal hyperplasia, androgen secreting tumours, and Cushing's syndrome. **Key:** PCO: polycystic ovarian; NIH: National Institute of Health; ESHRE: European Society for Human Reproduction and Embryology; ASRM: American Society for Reproductive Medicine; AE-PCOS: Androgen Excess and PCOS Society.

Despite omission from the NIH PCOS definition, the PCO morphology was incorporated into an internationally accepted PCOS definition following a consensus conference in 2003. The European Society for Human Reproduction and Embryology (ESHRE) along with the American Society for Reproductive Medicine (ASRM) stated that women should be diagnosed with PCOS if they have two of the three cardinal characteristics; that is, oligo-anovulation, evidence of excess androgens and the PCO morphology (Table 1.1) (Rotterdam ESHRE/ASRM, 2004). This expanded Rotterdam definition also broadened the heterogeneity of PCOS phenotypes, since women with the PCO morphology and hyperandrogenism, and those with PCO and menstrual disruption were included for the first time. Notably, the additional phenotypic subgroups created by the introduction of the Rotterdam criteria meant that the number of women with PCOS was substantially increased, and that hyperandrogenism was no longer a prerequisite for diagnosis. Recent evidence-based guidelines (Teede *et al.* 2018) continue to support the Rotterdam criteria for use in adult women.

However, there has been contention around the Rotterdam criteria; a growing body of literature implicates hyperandrogenism as the greatest determinant in the pathophysiology of PCOS (Georgopoulos *et al.*, 2014), as well as being instrumental in the metabolism-related disorders (Carmina *et al.,* 2005) which are frequently associated with this condition. This viewpoint was supported by the findings of an Androgen Excess and PCOS Society (AE-PCOS) task force (Azziz

et al., 2009) which reported that androgen excess is pivotal in PCOS and formed the AE-PCOS criteria. Thus, according to these criteria, a PCOS diagnosis must include the clinical or biochemical evidence of hyperandrogenism (Table 1.1). These findings effectively excluded the non-hyperandrogenic phenotype (PCO morphology with oligo-anovulation) that was introduced by the Rotterdam guidelines.

The heterogeneity of PCOS, alongside three internationally accepted diagnostic definitions brought about a lack of clarity in clinical practice. Therefore, a 2012 review of the guidelines (NIH, 2012) resulted in new diagnostic recommendations; a broadened version of the Rotterdam criteria was devised and this was accompanied by a detailed description of each PCOS phenotype as defined by the new criteria (Table 1.2).

Table 1.2. Polycystic ovary syndrome (PCOS) phenotypic subgroups and inclusion within diagnostic criterion.

Parameter	Phenotype A	Phenotype B	Phenotype C	Phenotype D
Hyperandrogenism	✓	✓	✓	X
Oligo-anovulation	✓	✓	X	✓
PCO morphology	✓	X	✓	✓
NIH (1990)	✓	✓	X	X
ESHRE/ASRM (2004)	✓	✓	✓	✓
AE-PCOS (2006)	✓	✓	✓	X

Key: PCO: polycystic ovarian; NIH: National Institutes of Health; ESHRE/ASRM: Human Reproduction and Embryology/American Society for Reproductive Medicine; AE-PCOS: Androgen Excess and PCOS society (Lizneva *et al.*, 2016).

The global use of multiple diagnostic criteria has reportedly raised issues of compatibility for PCOS research worldwide, leading to delayed progress in gaining an understanding of the condition and confusion within clinical practice (NIH, 2012). The NIH consensus panel (2012) proposed the use of expanded Rotterdam (2003) criteria; as with these criteria, they recommend that diagnosis should be made when at least two of the three cardinal characteristics are present, but that they should also

be accompanied with a description of the phenotypic subgroup to which the patient has been diagnosed. Accordingly, the NIH consensus panel (2012) suggested the use of four phenotypic classifications (Table 1.2), which were previously outlined by Azziz *et al.* (2006), including: Phenotype A: clinical or biochemical hyperandrogenism + oligo-anovulation + PCO morphology; Phenotype B: hyperandrogenism + oligo-anovulation; Phenotype C: hyperandrogenism + PCO morphology; and Phenotype D: oligo-anovulation + PCO morphology. This phenotypic approach allows researchers and clinicians to characterise populations based upon the presence (or absence) of PCOS's defining features (Lizneva *et al.,* 2016).

Furthermore, when making treatment recommendations in clinical practice, it is advantageous to identify the phenotype of a patient; for example, women who present with the "classic" PCOS phenotypes (A or B) often have more severe menstrual dysfunction and infertility (Kim *et al.,* 2014), and are at a higher risk of metabolic dysfunction (Azziz *et al.*, 2006). Indeed, women with these PCOS phenotypes are more hyperinsulinaemic (Welt *et al.,* 2006) and more insulin resistant (Diamanti-Kandarakis and Panidis, 2007; Moran and Teede, 2009) than women presenting with phenotypes C or D. In addition, phenotypes A and B tend to have a higher body mass index (BMI) (Welt *et al.,* 2006), increased body weight and central adiposity (Moran and Teede, 2009), as well a greater severity of dyslipidaemia (Carmina *et al.,* 2005; Kim *et al.,* 2014). The cumulative result is that those with "classic" PCOS have a greater risk for metabolic syndrome (MetS) than women with the non-classic (hyperandrogenism and PCO morphology) or non-hyperandrogenic phenotypes (Mehrabian *et al.,* 2011). Moreover, Mehrabian *et al.* (2011) report differences in the prevalence of MetS between phenotypes A and B, with higher MetS prevalence in women with phenotype B (64% prevalence) than in those with the PCO morphology (23%).

In contrast, women with hyperandrogenism and PCO (phenotype C), are reported to have intermediate levels of serum androgens and their associated clinical symptoms (*i.e.,* hirsutism and/or acne), plasma insulin, and atherogenic lipids, when compared to those with the "classic" phenotypes (Lizneva *et al.,* 2016). It is also often reported that women with non-hyperandrogenic PCOS have a lower degree of endocrine (Dewailly *et al.,* 2006; Welt *et al.,* 2006) and metabolic (Dewailly *et al.,*

2006; Zhang *et al.*, 2009) dysfunction, as well as a lower prevalence for MetS compared to hyperandrogenic PCOS (Mehrabian *et al.*, 2011).

1.3. PCOS Global Prevalence

One implication of the multiple definitions of PCOS is that it becomes difficult to precisely assess its global prevalence. Individual studies have reported PCOS prevalence rates between 2% (Chen *et al.*, 2008) and 21% (Boyle *et al.* 2012), and this variation is likely attributed to the diagnostic criteria used, but also to ethnic variation and differences in clinical practice. A recent systematic review and meta-analysis (Bozdag *et al.*, 2016) pooled data from studies that reported prevalence of PCOS according to at least one of the recognised diagnostic criteria. Based on data from 24 eligible trials, this analysis reported the PCOS prevalence and 95% confidence intervals (CIs) for each PCOS definition as follows: NIH = 6% (95% CIs: 5 to 8%; 18 trials), ESHRE/ASRM = 10% (95% CIs: 8 to 13%; 15 trials), and AE-PCOS = 10% (95% CIs: 7 to 13%; 10 trials). Data for PCOS prevalence according to the geographical location of the included studies is presented in Figure 1.1. Based on the latter, there is evidence of significant geographical heterogeneity for individual diagnostic tools which may suggest a degree of regional inconsistency when identifying PCOS symptoms or interpreting phenotypic definitions (Bozdag *et al.*, 2016).

Overall, it is clear that PCOS prevalence varies greatly depending upon the criteria used. As such, due to their expansive definitions and inclusion of additional phenotypes, the Rotterdam (2003) and AE-PCOS (2006) criteria tend to produce greater estimates of PCOS prevalence than the original NIH (1990) criteria (Sirmans *et al.*, 2014). In fact, the original use of the NIH (1990) criteria provided less variability when comparing countries/regions, but there are a few marked exceptions. For example, the study by Chen and colleagues (2008) assessed 915 women who were not presenting with a medical reason or complaint, and reported PCOS prevalence of only 2.2%. In contrast, studies in Australian Aboriginals (Boyle *et al.* 2012) and Mexican women (Goodarzi *et al.*, 2005) report significantly higher prevalence rates (15.3% and 13.0%, respectively). However, these findings may

be partly explained by limitations in sampling, since in the two latter studies where prevalence is higher, participants were identified from individuals enrolled in trials studying diabetes (Boyle *et al.* 2012) or coronary artery disease (CAD) (Goodarzi *et al.,* 2005). Both these conditions are typically associated with poor metabolic health, and progression to these conditions is more common in PCOS than in the general population.

Figure 1.1. Worldwide prevalence of PCOS according to geographic location. Data are presented as percentage of population diagnosed and 95% confidence intervals (CIs). Data unavailable for geographic regions coloured in grey (Bozdag *et al.*, 2016).

If the results of these epidemiologic studies may not be truly representative of the general population, then it is pertinent to consider studies with a more representative design. Li *et al.* (2013) randomly sampled nearly 17,000 women in a national Chinese study using the Rotterdam criteria and reported a PCOS prevalence of 5.6%. Utilising a smaller sample (n = 820), another study (Mehrabian *et al.*, 2011) randomly invited reproductive-aged women for PCOS screening and based on the Rotterdam criteria found the 15.2% of the screened women had a PCOS diagnosis. Of note, this was halved when NIH or AE-PCOS criteria were used. Another study of mainly Caucasian women assessed in adulthood (March *et al.*, 2010) found that 17.8% were diagnosed with PCOS using the Rotterdam criteria. This was 8.7% and 10.2% when the NIH and AE-PCOS criteria were applied, respectively. It is evident that, there is still considerable variation in the reported data even when greater confidence can be assumed from the applied sampling methodology and the representativeness of the sampled population.

Part of this variation may be explained by ethnic differences, whilst there are other factors to consider. One such factor is the complexity of PCOS assessments, since accurate diagnosis often requires multiple clinical and biochemical assessments, ultrasound, and likely multiple visits to specialist clinics (Lizneva *et al.*, 2016). Interestingly, Mehrabian *et al.* (2011) acknowledge such a limitation in their data. Moreover, detection rates of PCO morphology (and therefore positive diagnoses) are greater when transvaginal ultrasound is used over transabdominal methods (Farquhar *et al.*, 1994). The implication is that because of variations in routine clinical practice, or an inability to complete intensive assessments and appropriate follow-up in research projects, many studies may actually be underreporting PCOS prevalence (March *et al.*, 2010).

Notably, ethnicity may also constitute another contributing factor for geographical variations in reported PCOS prevalence. Indeed, it has been previously reported that the ethnic background of women with PCOS affects the severity of the clinical, hormonal, and metabolic characteristics associated with this syndrome (Ng *et al.*, 2007), and large phenotypic variations have been reported in different populations. For example, Polson *et al.* (1988) report that the PCO morphology is a common feature of PCOS in Western women. Polycystic ovaries were reported in 26% and 87% of

women with amenorrhea and oligomenorrhea, respectively, and in 92% of those who were hirsute. In contrast, the PCO morphology was reported in ~6% of infertile Chinese women (Ng *et al.,* 2004) and in only 1% of fertile Chinese women (Ng *et al.,* 2001). When compared to an earlier study in South Asian women (Rodin *et al.,* 1998), these numbers are much smaller, since Rodin and colleagues (1998) had previously reported a 52% prevalence of polycystic ovaries. The latter study also reported that 49% of these women had menstrual irregularity, whilst further documented that these South Asian women had comparable metabolic dysfunction (namely insulin resistance) compared to South Asian women with T2DM but no PCO morphology.

1.4. PCOS Diagnostic Components

1.4.1. Hyperandrogenism

In healthy, reproductive-aged women the ovaries and the adrenal cortex share the bulk of the steroid biosynthesis pathways, with relatively equal contributions to the circulating levels of testosterone and androstenedione. The ovaries and the adrenal cortex both secrete more androstenedione than testosterone, but ~50% of circulating testosterone is derived from the peripheral metabolism of androstenedione (Ehrmann et al., 1995). Within the ovary, the theca interna layer within the ovarian follicle produces androgens, whilst within the adrenal cortex, it is the *zona fasciculata* responsible for synthesis (Nussdorfer, Mazzocchi and Meneghelli, 1978).

Unlike protein- or peptide-producing cells, steroid-secreting cells do not store each hormone in a ready state, but synthesise them for secretion as they are required. Steroid hormone synthesising cells in the body (*i.e.* adrenal cortex, placenta and ovary) contain intracellular lipid droplets in the cytoplasm. Cholesterol (Figure 1.2), being the natural precursor to all steroid hormones, enters the cell and is stored as cholesterol esters within these lipid droplets until required (Brook and Marshall, 1996). When required, the first step of steroid synthesis is the hydrolysis of cholesterol esters; cholesterol is then transported from the lipid droplet into the mitochondria of the cell where the primary reaction is the production of pregnenolone.

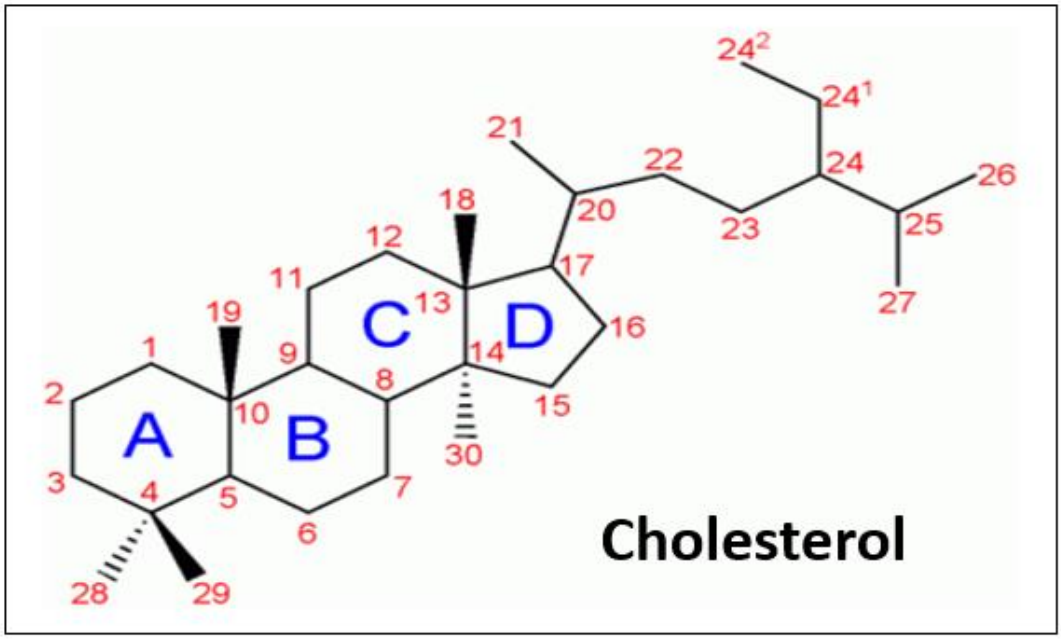


Figure 1.2. The ring-identification system of the steroid nucleus and the carbon numbering system used for steroids.

Pregnenolone is then converted to dehydroepiandrosterone (DHEA) via another two-step process along the Δ^5-pathway (Figure 1.3); in the adrenal gland, 17-hydroxyprogesterone is converted into either cortisol or sex hormones. Then 17β-hydroxysteroid dehydrogenase acts upon 17-ketosteroids to complete their conversion to testosterone, dihydrotestosterone and oestradiol (Balen *et al.*, 2005).

Figure 1.3. Schematic representation of the female steroidogenic pathway (Balen *et al.* 2005).

In healthy women, androgen secretion usually occurs in a two-fold episodic, diurnal manner that undergoes cyclic variation. A key event in the acute regulation of steroidogenesis occurs when cholesterol is delivered to the mitochondria and converted to pregnenolone (Clark *et al.*, 1995). This process is governed by trophic hormones; luteinizing hormone (LH) in the ovary and adrenocorticotropic hormone (ACTH) in the adrenal cortex (Manna *et al.*, 2009). The steroidogenic responses to LH and ACTH may also be further influenced by small peptides [*i.e.* insulin and/or insulin-like growth factors (IGFs)]. In *in vitro* models, insulin and IGF-1 and IGF-2 have been demonstrated to have a regulatory effect in ovarian theca and granulosa cells; insulin excess has been shown to increase androgen production (and promote follicular cyst development) in women with PCOS (Poretsky *et al.* 1999). Insulin action targets cytochrome P450c17α, increasing activity of 17α-hydroxylase and 17,20-lyase (Figure 1.3), resulting in increased ovarian androgen production (Nestler, 2008). Elevated insulin levels directly stimulate ovarian androgen secretion but also inhibit synthesis of sex hormone binding globulin (SHBG), thus elevating circulating free androgens and further contributing to hyperandrogenism (Diamanti-Kandarakis and Dunaif, 2012).

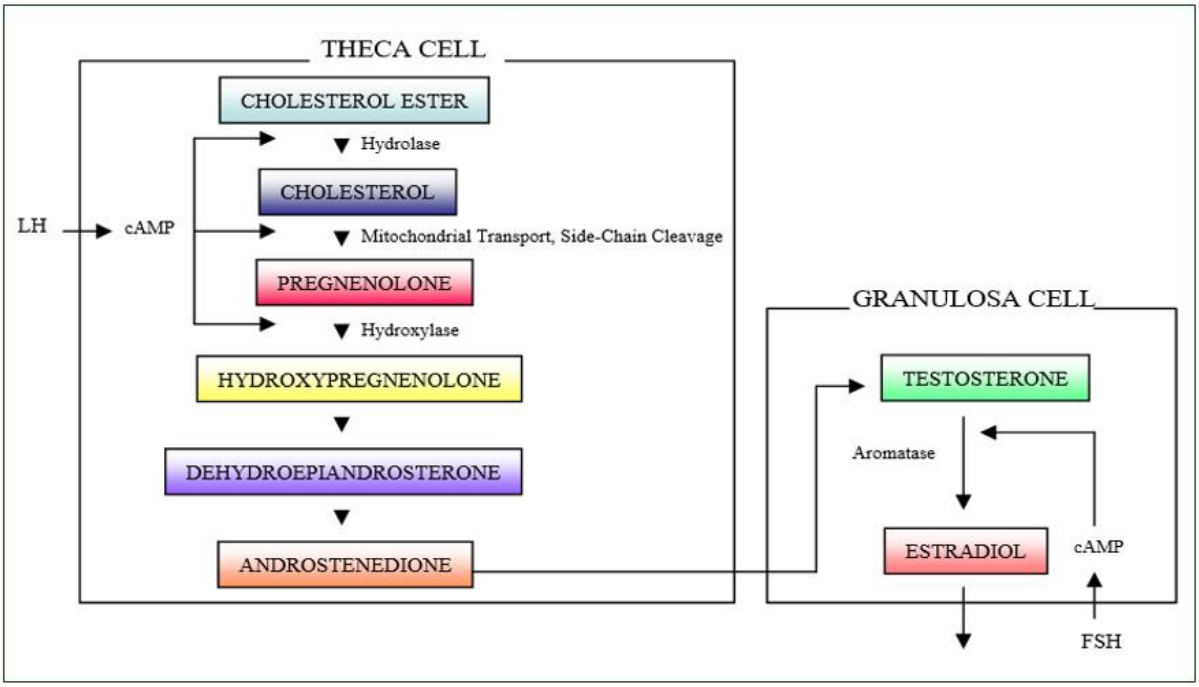


Figure 1.4. Two-cell, two-gonadotropin mechanism of follicular steroidogenesis.

Intra-ovarian androgens are an essential component of healthy follicular growth and oestradiol synthesis. However, when androgen synthesis is not fully co-ordinated (*i.e.*, excessive levels) with the needs of the growing follicle, follicular atresia occurs (Harman, Louvet and Ross, 1975). Within a healthy ovary, LH acts upon the theca cells inducing androgen synthesis, whereas it is follicle-stimulating hormone (FSH) that stimulates aromatase activity in the granulosa cells, facilitating the conversion of androstenedione and testosterone into oestrogens (Liu and Hsueh, 1986). This principle forms the basis of the two-cell, two-gonadotropin theory (Figure 1.4).

The two-cell, two-gonadotropin mechanism of oestrogen biosynthesis theory states that the theca cells secrete androgens in response to LH levels. Any synthesised androstenedione is then converted, via aromatase, to oestrogens within the granulosa cells; this process in the granulosa, is regulated by FSH (Liu and Hsueh, 1986). During a normal menstrual cycle, a dominant follicle will emerge and oestrogen content dominates over androgens that is not mediated by long-loop negative feedback effects (*i.e.*, oestrogen does not supress androgen production). These processes play an integral role in the regulation of these hormones and accordingly, overstimulation due to elevated LH results in a

downregulation of LH receptors which in turn, lowers activity of cholesterol side-chain cleavage, 17,20-lyase and 17-hydroxylase. This downregulation alters the ratio of 17-hydroxprogesterone and androgens (White *et al.,* 1995). Aromatase activity, in accordance with the development of granulosa cells, determine the rate of synthesis for androgens. A healthy follicle is able to convert androstenedione into oestrogen with efficiency but, in contrast, cystic follicles have an irregular ratio of androstenedione and oestrogen which desensitises them to the effects of FSH (via IGF-binding proteins) resulting in increased oestrogen production and decreased thecal androgen secretion (Balen *et al.,* 2005).

The increased ovarian androgen production that is often associated with PCOS is largely attributed to enhanced androgen synthesis by the follicular theca cells, which in women with PCOS often show increased expression of many genes that encode steroidogenic enzymes (McAllister *et al.,* 2015). This supports the notion that ovarian androgen excess in PCOS may be determined by genetics, at least partially. In addition to the overexpression of certain genes, Nestler *et al.* (1998) also reported that theca cells in PCOS women are more responsive (in terms of androgen secretion) to insulin and LH than healthy controls. In addition to ovarian related hyperandrogenism, 20-30% of women with PCOS also report elevated adrenal androgens, which may be the result of adrenocortical hyperfunction (Yildiz and Azziz, 2007).

Overall, biochemical hyperandrogenaemia in the context of PCOS relies on the measurement of circulating total or free testosterone levels. In clinical practice, using the free androgen index (FAI, calculated as the ratio of the circulating total testosterone levels divided by the circulating SHBG levels and then multiplied by a constant, typically 100) is also commonplace, since FAI is thought to be a sensitive method for assessing hyperandrogenaemia (Cibula *et al.,* 2000). For clinical hyperandrogenaemia, the major clinical manifestations of hyperandrogenism include hirsutism, acne, androgenic alopecia, impaired follicular growth and ovulatory dysfunction (Barry, 2019). Hyperandrogenism can be an underlying cause of these symptoms alone, whilst in PCOS these are often further exacerbated by the concomitant presence of insulin resistance.

1.4.2. Oligo/amenorrhea

In healthy females, menstruation begins approximately three years after the onset of breast development (Swenson and Havens, 2010). A typical menstrual cycle lasts for 28 days (Figure 1.5a) and day one of the cycle is marked by the first day of menstrual bleeding (period). During the menstrual cycle 4-5 follicles will grow, but (usually) only one follicle will mature, and, whilst the dominant follicle matures (Figure 1.6a), the remaining follicles will degenerate (Elsheikh and Murphy, 2008). FSH stimulates follicular growth, but also promotes the synthesis of oestrogen in the granulosa cells surrounding the oocyte, whilst the increase in oestrogen further promotes follicular growth. Once the dominant follicle has grown to approximately 15 mm, oestrogen levels are sufficient to cause a surge in LH production by the pituitary gland, which in turn triggers ovulation (~day 14); concomitantly, oestrogen stimulates endometrial growth in preparation for pregnancy. The first phase of the menstrual cycle is defined as the follicular phase, whereas the following phase (between ovulation and menses) as the luteal phase (Wallach, McNeely and Soules., 1988). In the second phase, should the released egg not be fertilised, the endometrium is shed a further 14 days after ovulation and this cycle restarts.

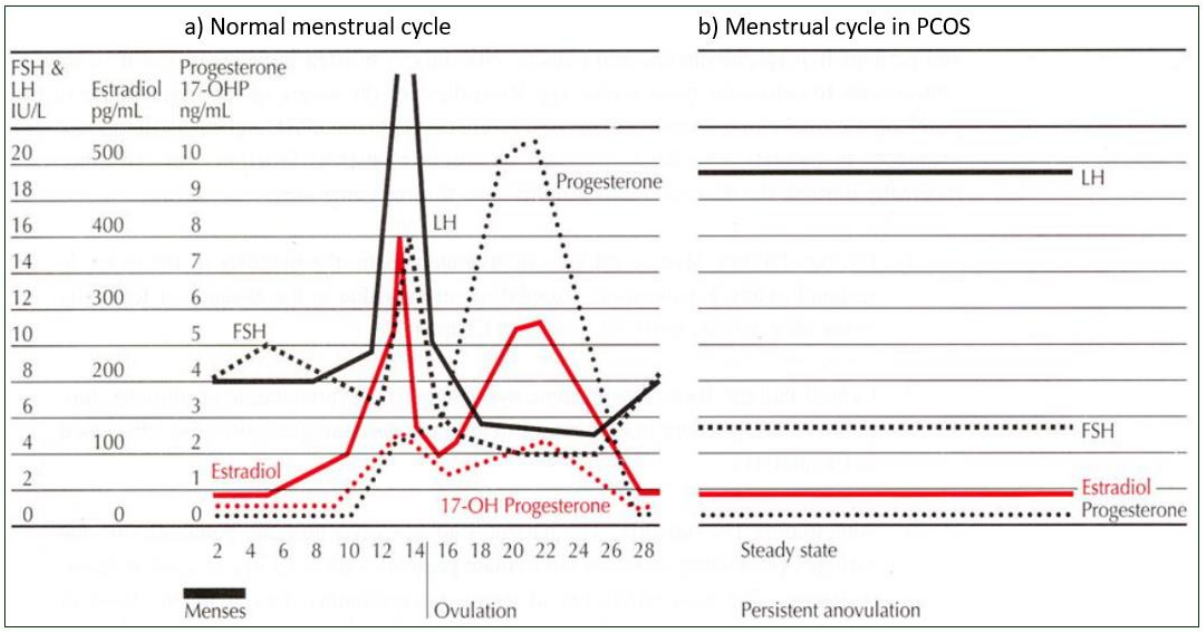


Figure 1.5. A. Hormonal responses during the menstrual cycle in healthy women; **B.** persistent ovulation in women with polycystic ovary syndrome.

Following ovulation, the empty follicle (*corpus luteum*) produces progesterone (Figure 1.5A) to cease the release of additional eggs and to prepare the womb for implantation with a fertilised egg and subsequent pregnancy. Should pregnancy not occur, the *corpus luteum* shrivels, reducing levels of oestrogen and progesterone; these falling levels allow for the shedding of the uterine lining during menstruation (Jabbour *et al.,* 2006).

The causes of menstrual disruption in PCOS are attributed to hormonal imbalances. Many women with PCOS have persistently elevated LH, and a constant level of FSH (Figure 1.5B). Persistently elevated LH stimulates the theca cells to produce an increased volume of androgens (namely androstenedione and testosterone). Alongside this, unwavering levels of FSH result in paused follicular development (Figure 1.6b), failure of follicular maturation and ultimately, no ovulation (Elsheikh and Murphy, 2008). Failure to reach the luteal phase, means that insufficient progesterone is produced to induce menstruation and periods become irregular (oligomenorrhea) or absent (amenorrhea). The immature follicles then remain within the ovary and form the 'cysts', as observed on ultrasound, that are synonymous with the condition (Couse *et al.,* 2004).

Figure 1.6. Follicular maturation, ovulation and luteal phase in a healthy ovary (a) and, paused follicular maturation and development of cysts within a polycystic ovary.

Disturbance of the menstrual cycle is a cardinal characteristic of PCOS. As such, oligomenorrhea (menstruation occurs at intervals of between 35 days and 6 months) or amenorrhea (no menstrual bleeding for > 6 months) and anovulation (the ovaries do not release an oocyte during the menstrual cycle) are frequently associated with the condition (Balen et al., 2005). Indeed, in a sample of 1741 women, Balen *et al.* (1995) reported that 47% experienced oligomenorrhea, ~19% amenorrhea, and ~30% reported regular menses. Similarly, Robinson *et al.* (1993) reported that 73.6% of PCOS women in their study had either oligomenorrhea or amenorrhea and that the PCO morphology is detected in the majority of women with oligomenorrhea.

The presence, and severity of menstrual disruption differs greatly between women with PCOS. These variations are often associated with comorbidities, such as obesity, insulin resistance, elevated androgens and concentrations of LH. When Balen *et al.* (1995) stratified PCOS women by BMI, a greater proportion (32%) of women had regular menstrual cycles when their BMI was < 30 kg/m^2. In contrast, only 22% of those with a BMI > 30 kg/m^2 experienced regular cycles. In addition, oligomenorrheic women with PCOS also had decreased insulin sensitivity when compared to healthy controls and, regularly menstruating PCOS women. This indicates that PCO morphology in combination with insulin resistance is associated with anovulation and menstrual irregularities (Robinson *et al*., 1993).

1.4.3. PCO Morphology

There have been many descriptions and definitions of the PCO morphology, and, as imaging technology has progressed, these have been further refined. Based upon histological samples taken at wedge resection of the ovaries, Stein and Leventhal (1935) first characterised the ovarian features from seven hirsute, amenorrheic women which included a prominent theca, fibrotic thickening of the tunica albuginea and multiple cystic follicles. Goldzieher and Green (1962) have also reported an 'excessive' number of antral follicles (2-6 mm in diameter) in a similar population. The introduction of transabdominal and transvaginal ultrasound scanning has altered the method in which the PCO morphology is identified.

The seminal work of Adams *et al.* (1985) used transabdominal ultrasound to define a PCO as containing at least 10 follicles, between 2-8 mm, in one plane, and either arranged peripherally around a dense core of ovarian stroma, or interspersed throughout an increased amount of stroma. This definition is commonly cited in subsequent studies that used ultrasound scanning to detect the PCO morphology. Transvaginal ultrasound scanning has largely superseded previous methods, as it offers a higher resolution and is more comfortable for the patient. Thus, this method may provide a more accurate view of the internal structures of the ovaries, avoiding apparently homogenous ovaries as described with transabdominal scans, which may be particularly beneficial in women with obesity (Balen *et al.,* 2003). It is likely that three-dimensional ultrasound may further enhance the detection of the PCO morphology, and may be utilised more frequently in the coming years.

The most comprehensive definition of the PCO morphology which is currently applied was introduced by a consensus meeting on PCOS held by the ESHRE/ASRM (2003). Based on this, PCO morphology should be characterised by either, 12 or more follicles measuring 2-9 mm in diameter, or ovarian volume >10 cm^3. In the context of this consensus, it was further stated that the subjective appearance of the ovaries, follicular distribution and stromal echogenicity should not be utilised in defining the PCO morphology and that women who have the PCO morphology, but in the absence of ovulatory dysfunction or hyperandrogenism should not be diagnosed with PCOS.

1.5. PCOS and Metabolic complications

1.5.1. Insulin Resistance

Hyperinsulinaemic insulin resistance is commonly associated with PCOS, and women with PCOS, either lean or obese, are more insulin resistant than BMI matched controls (Dunaif, *et al.*, 1989; Norman, *et al.,* 2001; DeUgarte, Bartolucci and Azziz, 2005). Iuorno (2007) defines insulin resistance as the reduced ability of insulin to stimulate glucose uptake. Insulin has a number of subcellular actions that are unrelated to glucose uptake, but there are two major pathways of insulin subcellular signalling: (1) the glucose transporter type 4 (GLUT-4), primarily located in adipose tissues and striated muscle, is stimulated by insulin to increase glucose uptake intracellularly (Sivitz

et al., 1989); and (2) the mitogen activated protein (MAP) kinase activation pathway, which is primarily involved in the regulation of mitogenesis. Whilst glucose uptake is downregulated in insulin resistance, the MAP kinase pathway may be upregulated, increasing lipogenesis in T2DM (Koistinen et al., 2003). Generally, the resistant pathway refers to the former; that is decreased insulin mediated intracellular uptake of glucose.

Of note, insulin resistance is one of the two major contributing factors to the development of T2DM (Kahn, Hull and Utzschneider, 2006), which is extremely prevalent in women with PCOS (Teede *et al.,* 2006). Higher rates/degrees of insulin resistance and the increased presence of concomitant obesity, β-cell dysfunction and dyslipidaemia increase the risk of T2DM in PCOS (Ovalle and Azziz, 2002). Legro *et al.* (1999) suggested that up to 40% of women with PCOS may be insulin resistant, whilst Peppard *et al.* (2001) report that 25% of premenopausal women with T2DM may also have PCOS.

Although not included in the PCOS diagnostic criteria, insulin resistance is clearly associated with the pathophysiology of PCOS, but the exact underlying mechanisms of these links remain not fully clarified. In contrast to eumennorhic women with insulin resistance, who have decreased cellular insulin receptors that are inversely proportional to the degree of hyperinsulinaemia (Olefsky and Kolterman, 1981), women with PCOS reportedly have normal insulin receptor numbers which have a normal affinity for insulin (Conway *et al.*, 1994; Tarkun *et al.*, 2005). Abnormalities in subcellular signalling pathways are considered as a likely mechanism for insulin resistance in PCOS. Indeed, in the initial steps of the insulin signalling pathway, downregulated insulin mediated glucose transport in adipocytes has been reported (Dunaif *et al.,* 2001), as well as decreased insulin stimulated lipolysis (Rosenbaum, Haber and Dunaif, 1993). In addition, Dunaif (1995) reported that the fibroblasts of women with PCOS exhibited decreased serine autophosphorylation of the insulin receptor.

Overall, the presence of insulin resistance in a substantial proportion of women with PCOS means that screening for metabolic irregularities in PCOS (*e.g.,* for MetS and T2DM) should be commonplace, so that relevant prevention and treatment interventions should be promptly applied in the management of these patients. For a large proportion of women with PCOS, this aspect of the

management of PCOS requires equal attention to that of addressing menstruation and fertility problems.

1.5.2. Overweightness/Obesity in PCOS

PCOS is also commonly associated with overweightness and obesity (Legro, 2012), whilst excess body weight may also worsen hyperandrogenism and menstrual disturbances in women with PCOS (Kiddy *et al.,* 1990; Balen *et al.,* 1995; Liou *et al.,* 2009). Moreover, a number of studies have suggested that PCOS may also be more prevalent in women with overweightness or obesity; for example, a Spanish study (Alvarez-Blasco *et al.,* 2006) reported that 28% of women with overweightness/obesity met PCOS diagnostic criteria. However, women with obesity and PCOS are more likely to develop worse reproductive clinical symptoms (Gambineri *et al.,* 2002), and therefore contact a medical professional for support and so this increases the opportunity for a PCOS diagnosis to be made, which may partially exaggerate the association between PCOS and obesity (Lim *et al.,* 2012).

In addition to increased risk of overweightness/obesity, it has also been reported that women with PCOS are more likely to have increased visceral adiposity when compared with BMI-matched controls without PCOS (Karabulut *et al.,* 2012; Carmina *et al.,* 2009). Indeed, in the general population, it is waist circumference as opposed to BMI, which better explains the obesity-related health-risks (Janssen, Katzmarzyk and Ross, 2004; Huxley *et al.,* 2009). Similar effects are observed in women with PCOS, with women with PCOS and higher levels of central obesity exhibiting increased prevalence of insulin resistance (Lord *et al.,* 2006; Karabulut *et al.,* 2012), as well as higher severity of metabolic dysregulation (Pasquali *et al.,* 1994; Lord *et al.,* 2006), hyperandrogenism (Svendsen *et al.,* 2008), and reproductive disruption (*e.g.,* anovulation) (Carmina *et al.,* 2009).

Despite the apparent association between PCOS and obesity, determining the exact obesity prevalence in this population remains problematic. Within the general population, obesity prevalence is known to vary depending upon factors such as age, ethnicity, education status and geographical

location (Santos and Barros, 2003). In addition to these demographic characteristics, determining the precise prevalence of obesity in women with PCOS is further complicated by the use of multiple diagnostic criteria (Moran and Teede, 2009), and a potential lack of data that is representative of the general population.

Lim and colleagues (2012) completed a systematic review and meta-analysis in order to provide an evidence-based evaluation of the prevalence rates of overweightness, obesity and central obesity in women with PCOS. Individual studies included in this analysis reported prevalence of overweight and obesity rates ranging from 6% (Ansarin *et al.,* 2007) to 100% (Peppard *et al.,* 2001; Glueck *et al.,* 2003; Villaseca *et al.,* 2004), whilst the performed meta-analysis identified a polled effect estimate of 61% (95% CI: 54-58%, 21 studies, 3132 participants). When Lim *et al.* (2012) separated data for obesity only, prevalence rates ranged from 12.5% (de Vries *et al.,* 2007) to 100% (Peppard *et al.,* 2001), and the pooled effect estimate was 49% (95% CI: 42-55%, 18 studies, 4160 participants). In addition, the prevalence range for central obesity was 20% (Gul *et al.,* 2008) to 86% (Glueck *et al.,* 2003), with a pooled prevalence of 54% (95% CI: 43-62%, six studies, 1191 participants). For all three analyses, women with PCOS demonstrated statistically higher values than control women without PCOS; all meta-analyses had statistically significant levels of heterogeneity (P <.001; $I^2 \geq 84\%$) reducing confidence in effect estimates. It should also be noted that the majority of studies included in these analyses are drawn from medical settings that may introduce considerable selection bias. Indeed, studies utilising unselected populations have reported that BMI was markedly lower in unselected groups (*i.e.,* not medically identified), and that prevalence of obesity was comparable to data from the general population (Ezeh, Yildiz and Azziz, 2013; Luque-Ramírez *et al.,* 2016). This implies that at least a proportion of the observed higher obesity prevalence in PCOS may be due to self-referral to healthcare settings in response to exacerbated symptoms.

Contrasting evidence also exists regarding whether increased BMI drives PCOS development; again, PCOS incidence was greater in women that had obesity when these women had been selected based upon referral for bariatric surgery (Gosman *et al.,* 2010), dietary consultation (Alvarez-Blasco *et al.,* 2006) or in those self-reporting symptoms (Laitinen *et al.,* 2003). In contrast, Yildiz *el al.* (2008)

found that increasing BMI had little influence upon PCOS prevalence in an unselected population. Whilst this methodology may help to reduce selection bias, a similarly designed study (Yildiz *et al.*, 2012) presented conflicting results, since it noted that as mean BMI increased, so too did instances of PCOS. This contrasting evidence makes it difficult to establish whether (and to what extend) obesity drives PCOS. Thus, based on the available evidence, it seems more likely that PCOS is present, but possibly undetected, and that increased adiposity exacerbates symptoms leading to clinical presentation (Lizneva *et al.*, 2016).

Irrespective of the order of exacerbated clinical manifestations, it is widely reported that women with PCOS and overweightness/obesity have poorer health than their lean counterparts, with the relevant impact of PCOS and obesity being both independent and additive (Dunaif *et al.*, 1989). In fact, Dunaif *et al.* (1992) suggest that insulin action in women with obesity without PCOS is comparable to that of lean women with PCOS. Risk of impaired glucose tolerance (IGT) is also commonly associated with obesity; normal-weight women with PCOS seldom present with IGT, yet their counterparts with overweight/obesity are at a much higher risk (Ehrmann *et al.*, 1999). Similarly, dyslipidaemia (Legro, Kunselman and Dunaif, 2010) and MetS (Ehrmann *et al.*, 2006) are also associated with the degree of increased BMI in these women.

When the severity of hormonal and reproductive PCOS symptoms are considered, their relationship with obesity is less certain. Previous evidence has not identified any clear relationship between obesity and menstrual regularity (Solomon *et al.*, 2002), hyperandrogenaemia, hirsutism (Carmina *et al.*, 1992) or the identification of PCO morphology using ultrasound (Legro *et al.*, 2005). However, obesity is reportedly a predictor of how well women with PCOS respond to treatment, particularly treatments aimed at improving fertility outcomes (*e.g.*, ovulation and likelihood of pregnancy) (Legro, 2012). Indeed, clomiphene resistance (Imani *et al.*, 1999) and a failure to conceive via *in vitro* fertilisation (Rausch *et al.*, 2009) are all negatively affected proportionate to the degree of obesity. In addition, a reduction in BMI can promote favourable effects upon other circulating factors, such as free androgen and IGF-1 levels, which are shown to predict ovulatory function (Imani *et al.*, 2000).

1.6. PCOS and Psychological wellbeing

1.6.1. PCOS and Depression - Anxiety

Given the aforementioned disease burden, it is somewhat unsurprising that women with PCOS exhibit more often mental health complications (*e.g.*, depression and anxiety) than their counterparts without PCOS (Farrell and Antoni, 2010). A previous systematic review used standardised mean difference (SMD, calculated using Hedges' g with values of 0.2, 0.5 and >0.8 indicating, respectively, a small, moderate and large difference between groups) meta-analyses to assess the difference in depression and anxiety between women with PCOS and controls, showing that the former had higher levels of both depression (Hedges' g: 0.82, 95% CI: 0.73 to 0.92; 13 studies, 2257 participants) and anxiety (Hedges' g: 0.54, 95% CI: 0.33 to 0.75; 6 studies, 375 participants) than the latter (Barry, Kuczmierczyk and Hardiman, 2011). Of note, Barry and colleagues (2011) state that there is some uncertainty around the clinical importance of these reported differences, since, although effect sizes were large, and women with PCOS had scored significantly higher, the real magnitude of the difference and average scores was representative of no depression in the controls and mild depression in the women with PCOS. Similarly, anxiety scores were only representative of modest clinical differences, with the average control group scores being slightly lower than the norm, and the scores for the women with PCOS being 'mildly elevated'. Whilst this may seem as not significant, it is postulated that even a modest increase in a hyperandrogenic condition could further compound symptoms. Given that women produce ~25% of their testosterone from their adrenal glands (Burger, 2002), prolonged anxiety/stress may lead to elevated testosterone levels and a greater severity of symptoms by chronic activation of the adrenal glands via the hypothalamic-pituitary-adrenal (HPA) axis (Reiche, Nunes and Morimoto, 2004). This effect may be further exaggerated due to an increased HPA response in women with PCOS (Benson *et al.*, 2009).

Moreover, the findings of a more recent systematic review and meta-analysis (Cooney *et al.*, 2017) tended to agree with the findings of Barry, Kuczmierczyk and Hardiman (2011), since the prevalence of women with PCOS that had depression was 36.6% compared to only 14.2% in control women. Meta-analysis revealed that women with PCOS had increased odds of any depressive condition (odds

ratio, OR: 3.78, 95% CI: 3.03 to 4.72; 18 studies) and further increased odds of moderate/severe depressive symptoms (OR: 4.18, 95% CI: 2.68 to 6.52, 11 studies) compared to women without PCOS. When anxiety was analysed, the median prevalence was 41.9% in women with PCOS compared to 8.5% in a polled control. This also represents an increased likelihood of any anxiety (OR: 5.62, 95% CI: 3.22 to 9.80, 9 studies) or moderate/severe anxiety symptoms (OR: 6.55, 95% CI: 2.87 to14.93; 5 studies) in women with PCOS compared to controls. Of note, when women were matched for BMI, fewer studies were included in the analysis, but these odds were largely unaffected.

As to the underlying cause of this increased psychological comorbidity, Barry (2019) states that there are differences between the causes of depression and anxiety in women with PCOS. Indeed, there is likely a host of factors which either directly or indirectly contribute to depression in PCOS, including increased body weight, acne, and hirsutism (Barry, Qu and Hardiman, 2018). Furthermore, it is likely that individuals who have been diagnosed with a chronic condition without definite treatment, such as PCOS, will experience an acute and chronic impact upon their quality of life, with subsequent mood disturbances as a result (Barry, 2019). In this context, it is postulated that, as opposed to elevated androgen levels directly causing depression, it is their influence upon the distressing symptoms of PCOS which promote a depressive state (Pastore *et al.*, 2011; Borghi *et al.*, 2018; Batool *et al.*, 2016).

Interestingly, anxiety in PCOS is also hypothesised to be associated with reactive hypoglycaemia (a state of lowered glucose which typically occurs after meals) which has previously been reported to be four times more frequent in women with PCOS than controls (Atluntas *et al.*, 2005). A condition called ‘tense-tiredness’ often occurs in relation to hypoglycaemia, which results in the individual feeling depressed, anxious and fatigued (Thayer, 1989). Although increased fatigue is not widely reported in PCOS (Hollinrake *et al.*, 2007), it has been suggested that women with PCOS demonstrate more of the psychological manifestations associated with tense-tiredness than their healthy counterparts (Barry *et al.*, 2011a). Of note reactive hypoglycaemia could be controlled through dietary interventions (*e.g.*, low-glycaemic index meals).

1.7. Treatment options for PCOS

1.7.1. International Guidelines for PCOS

Due to the heterogeneous nature of PCOS, multiple diagnostic criteria and varying phenotypes, clinical practice for the management of PCOS is largely inconsistent (Teede *et al.,* 2018). Although a series of international guidelines have previously been developed (Balen *et al.,* 2016; Goodman *et al.,* 2015; Conway *et al.,* 2014; Legro *et al.,* 2013; Teede *et al.,* 2011), many of these are now outdated. Furthermore, it is thought that there are gaps in the guidance provided due to a narrow scope, potential limitations of the methods used to develop them, and limited involvement from women with PCOS (Teede *et al.,* 2018). Thus, clinical guidelines and recommendations for the management of women with PCOS remain relatively unclear. In 2018, Teede *et al.* developed comprehensive evidence-based guidelines that are intended to '*serve as a single source of international evidence-based recommendations to guide clinical practice with the opportunity for adaptation in relevant health systems'*. The following sections will briefly summarise treatment recommendations for the aforementioned symptoms of PCOS.

1.7.2. Hormonal contraceptives for PCOS

For the management of irregular menstrual cycles and hyperandrogenism symptoms, treatment with combined oral contraceptives is recommended (Teede *et al.,* 2018). The progestin contained within hormonal contraceptives has a supressing effect upon LH, which in turn reduces ovarian androgen production. Furthermore, the contained oestrogen increases circulating SHBG, thus reducing the amount of bioavailable androgens, which are responsible for hyperandrogenism-related PCOS symptoms (Legro *et al.,* 2013). The effectiveness of hormonal contraceptives upon carbohydrate metabolism, lipid profiles and overweight/obesity are less clear and may be variable depending upon oestrogen dose and the progestin type that are prescribed (van der Vange, Kloosterboer, and Haspels, 1987). Moreover, oral contraceptives may often be prescribed in combination with other pharmacological agents, such as insulin sensitizers and antiandrogens.

1.7.3. Other pharmacological interventions for PCOS

With regard to anovulation and improvements at conception and live birth rates, letrozole should be considered as first line pharmacological treatment when available, otherwise other ovulation induction agents, such as clomiphene citrate or gonadotropins, can be used (Teede *et al.,* 2018). Women who are clomiphene citrate resistant may also be offered laparoscopic ovarian surgery as a second line treatment for anovulatory infertility (Seow *et al.*, 2008); laparoscopic ovarian drilling has been shown to induce ovulation in 83% of clomiphene citrate resistant women with PCOS (Campo *et al.,* 1993). Although not as effective as oral contraceptives (Morin-Papunen *et al.,* 2003), the use of metformin has also been shown to improve menstrual cycle regularity (Nestler *et al.*, 1998) and ovulation rates (Tang *et al.,* 2012) in women with PCOS. Furthermore, metformin, alongside lifestyle changes are often recommended for women with PCOS and obesity, particularly when metabolic complications are present since this treatment appears to offer the greatest health benefit in those with T2DM risk factors or IGT (Teede *et al.*, 2018). A previous systematic review in reproductive-aged women with PCOS and overweight/obesity has demonstrated that metformin was successful at promoting weight loss in women with PCOS (~2.9% decrease in body weight) when compared to a placebo (Nieuwenhuis-Ruifrok *et al.,* 2009) which was equivalent to orlistat treatment for weigh loss (Padwal, Li and Lau, 2003). However, when women were adhering to a lifestyle weight loss intervention with increased PA and dietary control, there was additional benefit of adding metformin (Ladson *et al.*, 2011; Nieuwenhuis-Ruifrok *et al.,* 2009).

1.7.4. Lifestyle interventions for PCOS

Overall, diet and exercise should form the principle treatment recommendation for women with PCOS and obesity (Legro *et al.,* 2013). As such, in order to achieve/maintain a healthy weight, optimise hormone levels, metabolic profiles, and improve general health and quality of life, women with PCOS are advised to follow a healthy diet and engage in more and regular PA (Teede *et al.*, 2018). In women with PCOS and obesity, weight loss of 5-10% has been shown to induce clinically important changes across a range of outcomes and accordingly, lifestyle changes should be

recommended to all women with PCOS and increased BMI (Kiddy *et al.*, 1992). In addition to diet and exercise, lifestyle interventions typically include behaviour change components that are designed to promote long-term adherence to an improved lifestyle.

Given that PA has been proven to be effective at improving physical and psychological health in multiple populations and that it forms part of the primary treatment recommendation (lifestyle) in PCOS, further investigation as to its effectiveness is warranted. Indeed, there are no published large randomised controlled trials assessing the role of exercise in the management of PCOS, and the corresponding confidence – using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach/system (Schünemann *et al.*, 2013) – in international recommendations is low or very low (Teede *et al.*, 2018).

1.8. Physical Activity and Exercise for PCOS

1.8.1. Defining physical activity (PA)

PA has previously been defined as '*any bodily movement produced by skeletal muscle that results in energy expenditure*' (Caspersen *et al*., 1985). On this basis, every human performs PA every minute of every day just to sustain life. However, the volume of PA completed varies from person to person, and for each individual over time, based upon their personal choices and circumstances. Caspersen *et al.* (1985) further state that exercise differs from PA, with the former being a subset of PA that is structured, planned and uses repetitive bodily movements with the aim of improvement or maintenance of physical fitness levels.

This definition, along with others (Knuttgen, 1978; Rogers and Cavanagh, 1984; Faulkner, 2003) have proved to be slightly contentious. Winter and Fowler (2009) state that any definition of exercise (and terms associated with it) must adhere to universally adopted principles of science and satisfy the Système International d'Unités (SI), suggesting that these previous definitions of exercise fail to meet SI requirements. Winter and Fowler's (2009) concern surrounds the wording of '*repetitive bodily movement*', arguing that isometric muscular contraction (*i.e.* when there is no change in the

length of the contracting muscle and no corresponding movement) has been entirely overlooked. Accordingly, they suggest a potentially more suitable definition based on which exercise is a potential disruption to homeostasis by muscle activity that is either exclusively, or in combination, concentric, eccentric or isometric. Furthermore, they also cite that distinctions between PA and exercise must depend upon the context and circumstances. In the context of this thesis, the two terms (exercise and physical activity) are applied interchangeably.

Following hypertension (responsible for 13% of global mortality), tobacco usage (9%) and elevated blood glucose levels (6%), physical inactivity represents the fourth leading risk factor for global mortality, accounting for 6% of global deaths (WHO, 2009). Additionally, physical inactivity is thought be pivotal in the development of non-communicable diseases and, particularly as a key determinant of energy expenditure, is considered a determinant of overweightness and obesity (the fifth leading risk factor for mortality; 5%). Accordingly, the World Health Organisation (WHO) implemented global PA guidelines, recommending that adults (aged 18-64) complete at least 150 minutes of moderate-intensity aerobic exercise or 75 minutes of vigorous-intensity exercise per week (WHO, 2011). Alternatively, an equivalent amount using a combination of both intensities (moderate and vigorous) can be completed. Furthermore, aerobic activities should be completed in bouts lasting at least 10 minutes, whilst muscle strengthening exercises, involving the major muscle groups, should also be completed at least 2 days per week (WHO, 2011). Recommendations for adolescents (<18 years) and the elderly (≥ 65 years) vary slightly to better suit each demographic. Within the UK, similar guidelines are outlined by the Department of Health (2011).

When quantifying the intensity of PA, multiples of metabolic equivalents of task (MET) is the method that is commonly used. One MET is defined as the energy that an average individual expends when sitting quietly, and is based upon the relative rate of oxygen consumption (one MET is 3.5 ml/kg/min of oxygen) (Ainsworth *et al.*, 1993). This allows the intensity of PA to be categorised in absolute terms, as light (<3 METs), moderate (3-6 METs) or vigorous (>6 METs) (Saxton, 2011). Walking at ~5 km/h is an example of an activity that is roughly three METs, whereas brisk walking (~7 km/h) increases the METs to approximately five (Ainsworth, 2000). An activity that is five METs

would therefore expend five times more energy (~17.5ml/kg/min of oxygen) than when the body is at rest. In epidemiological studies, PA guidelines are often expressed as MET-minutes or MET hours per week (MET-h/wk), although this is often not appropriate for use by the general public. To complete 150 MET-minutes (or 2.5 MET-h) of PA, a five MET activity would need to be completed for 30 minutes. However, the same volume can be achieved by completing a 10-MET activity for 15 minutes. Individuals following WHO PA guidelines should therefore achieve 8.3-16.7 MET-h/wk, which epidemiological studies suggest may be enough to elicit favourable changes (Liu *et al.*, 2015; Li *et al.*, 2015; Kelly *et al.*, 2014).

1.8.2. Epidemiological evidence for physical activity

Early studies by Morris and colleagues (1953) were arguably the first to investigate the effects of exercise and PA. Indeed, when comparing the relatively active double-decker bus conductors to the more sedentary drivers, they reported lower incidence and later onset of CAD, as well as lower CAD-related mortality. Another landmark study with a 16-year follow up of >3000 dockworkers observed that those who were the most active had lower CAD mortality than their less active co-workers (Paffenbarger *et al.*, 1970), independently of smoking status, BMI or blood pressure. A later study by the same authors, which analysed ~17,000 Harvard Alumni in a 16-year follow-up, reported an inverse dose-response association between PA levels and all-cause mortality rates (Paffenbarger *et al.*, 1986).

More recently, the European Prospective Investigation into Cancer and Nutrition Study (EPIC) completed a 12-year follow up of more than 300,000 individuals to assess whether associations between PA levels and all-cause mortality are influenced by body fat (Ekelund *et al.*, 2015). In this study, PA was associated with lower all-cause mortality irrespective of BMI or waist circumference, and it was concluded that if all the inactive study participants were at least moderately active, mortality rates would be substantially reduced. In addition, a number of systematic reviews, including meta-analyses, have collated and evaluated the evidence of PA at reducing all-cause mortality and morbidity. One such review pooled data from 661,137 participants, and found that

those who completed some moderate to vigorous leisure time PA, but less than national guidelines, had a 20% reduction in all-cause mortality. By contrast, those completing 3-5 times the recommended levels, saw a 39% reduction in all-cause premature mortality risk (Arem *et al.,* 2015). Similar findings have been reported elsewhere, with Moore *et al.* (2012) identifying increases in life expectancy when leisure time PA was increased, and Hupin *et al.* (2015) reporting all-cause mortality risk reductions of up to 35% in older adults, whilst Samitz *et al.* (2011) found relative risk reductions of 9% and 4% for each 1-hour increment per week of vigorous- or moderate-intensity PA, respectively.

Other reviews have also assessed the association between PA levels and the risk of chronic disease or disease-specific mortality. Kyu *et al.* (2016) reported that higher levels of PA were associated with a reduced risk of developing breast cancer, colon cancer, diabetes, CAD or ischemic stroke, and that the greatest risk reductions (14-28%) came from those who were in the highest PA group (5-6 times the current WHO recommendations). Similarly, Wahid *et al.* (2016) found that moving from being inactive to meeting recommended PA levels resulted in a 26% lower risk of T2DM, 20% lower risk of CAD and 18% lower risk of ischemic stroke.

Overall, the available evidence strongly supports the notion that participation in moderate to vigorous PA is sufficient to reduce the risk of developing a host of conditions, and may prolong life. Despite this, there is an apparent lack of large scale randomised controlled trials (RCTs) that investigate exercise therapy in women with PCOS (Harrison *et al.,* 2011). Despite this limited availability of existing evidence, Legro *et al.* (2013) still argue that the perceived benefits of exercise are strong enough to recommend it as a first line treatment in the management of PCOS.

1.8.3. Exercise and PCOS

Given that PA has been shown to be effective in other insulin resistant populations, it is reasonable to assume that it would also be an effective treatment strategy for women with PCOS. Indeed, as aforementioned increased regular exercise is a first-line recommendation for the management of women with PCOS, along with wider lifestyle changes (Teede *et al.,* 2018). Despite this current

recommendation, there is little consensus about the optimal frequency, intensity, duration, or type of PA that would be most beneficial within this female patient population. Therefore, women with PCOS and clinicians typically adopt a generic approach to exercise prescription (*i.e.* following PA guidelines), often alongside pharmacological intervention, with uncertainty of the effectiveness of PA (Stepto *et al.*, 2019).

Identifying this marked gap in the existing studies/evidence in this field was the primary motivation for the present PhD thesis. As such, the chapters hereafter will synthesise the existing evidence for exercise/PA in women with PCOS and present studies which aimed to offer better insight on its role in improving health and wellbeing in this understudied population.

Chapter 2: Methods

2.1. Methods for a systematic review and meta-analysis: Exercise, or exercise and diet for the management of polycystic ovary syndrome

2.1.1. Systematic review protocol and registration

This systematic review was prospectively registered on the Prospero International Prospective Register of Systematic Reviews (CRD42017062576) and is reported based on the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Moher *et al.*, 2009).

2.1.2. Search Methods for Identification of Studies

The inclusion and exclusion criteria for eligible studies in this systematic review are presented in Table 2.1. In brief, 1) trials had to include women of a reproductive age who had received a PCOS diagnosis; 2) trials must have employed a randomised, or quasi-randomised experimental (intervention) design that 3) measured the chronic (*i.e.* long-term) effects of exercise, or exercise and diet combined, in women with PCOS. Exercise was defined according to the definition set out by Winter and Fowler (2009): "*a potential disruption to homeostasis by muscle activity that is either exclusively, or in combination, concentric, eccentric or isometric*". Accordingly, we included all methods of exercise training, including continuous aerobic exercise (*e.g.*, walking, jogging, running, swimming or cycling); high-intensity interval training; resistance training (*e.g.*, weight lifting); flexibility training; and yoga, Tai Chi, and Pilates. Trials were eligible if they compared a minimum of two conditions (*e.g.*, exercise compared to control, or exercise compared to diet) and used either a within-subject crossover design, or a between-subject comparison to a control, or alternative treatment group, as long the effect of exercise could be isolated. Studies that included follow-up testing, at least one month after completion of the intervention, were also eligible for inclusion.

The following databases were searched in June 2017: CENTRAL (in the Cochrane Library), PubMed, CINAHL, SCOPUS, EMBASE (via Web of Science), SportDiscus (via EBSCOhost), and PsycINFO (via OvidSP). There was no time limit specified for trial inclusion and only fully published, peer-reviewed papers were included; grey literature was not eligible for inclusion. In

addition, no language restrictions were placed on the search, meaning that foreign language papers could be included subject to translation. A search algorithm was developed for PubMed (Table 2.2), which was then modified for each additional database search.

One reviewer completed (CK) the initial database searches and the titles and abstracts were exported into EndNote (EndNote X8.2, build 11343; Clarivate Analytics, Philadelphia, USA) and duplicate records were removed. The resulting database was transferred into Covidence (Covidence v1344; Melbourne, Australia) and an additional screen for duplicates was completed. Two reviewers (CK and IL) then independently screened each title and abstract excluding papers that were clearly ineligible. At the next stage, PDFs were retrieved, and the same two reviewers independently screened each full-text, ensuring that those which were eligible were retained in the Covidence software. At each stage of screening, all disagreements on paper eligibility were resolved by discussion between the two reviewers, whilst any unresolved disagreements were decided by arbitration from a third reviewer (DB).

Table 2.1. Eligibility criteria for including studies in this systematic review.

Inclusion Criteria:

1. **Study Design:** randomised controlled trials and quasi-randomised controlled trials.
2. **Types of Participants:** reproductive-aged women with a diagnosis of polycystic ovary syndrome (PCOS) based on the National Institute of Health (NIH) diagnostic criteria (1990), or the Rotterdam ESHRE/ASRM (2003) diagnostic criteria, or the AE-PCOS Criteria (2006). Trials were also included where the PCOS diagnosis had been verified by a general practitioner or specialist clinician.
3. **Comparators:** exercise vs usual care/control, exercise combined with diet vs usual care/control, exercise combined with diet vs diet only. Exercise combined with diet vs exercise only, exercise vs diet, exercise combined with pharmaceutical vs pharmaceutical.
4. **All outcomes:** expected outcomes included: primary outcomes, such as blood pressure, fasting blood glucose, insulin and lipid concentrations; and secondary outcomes, such as body mass index, cardiorespiratory fitness, testosterone, free androgen index, and health-related quality of life measures.

Exclusion Criteria:

1. **Study Design:** case studies, cross sectional and non-randomised controlled trials.
2. **Types of Participants:** males, adolescent females, post-menopausal women, women without PCOS.
3. **Comparators:** women with PCOS *vs.* healthy controls, pharmaceutical *vs.* exercise, pharmaceutical *vs.* diet, diet *vs.* diet, surgical *vs.* any other condition.

Table 2.2. Search algorithm devised for advanced search in PubMed database. Algorithm adapted and implemented across additional databases.

Search	Query
#1	Polycystic ovary syndrome [MeSH Terms]
#2	Polycystic ovar* [Title/Abstract]
#3	PCOS [Title/Abstract]
#4	PCOD [Title/Abstract]
#5	Stein levent* [Title/Abstract]
#6	PCO [Title/Abstract]
#7	#1 OR #2 OR #3 OR #4 OR #5 OR #6
#8	Exercise [MeSH Terms]
#9	Exercise movement techniques [MeSH Terms]
#10	Exercise Therapy [MeSH Terms]
#11	Exercise [Title/Abstract]
#12	Physical education and training [MeSH Terms]
#13	Physical fitness [MeSH Terms]
#14	Physical fitness [Title/Abstract]
#15	Physical exertion [MeSH Terms]
#16	Sports [MeSH Terms]
#17	Physical Activity [MeSH Terms]
#18	Sport* [Title/Abstract]
#19	Physical activity [Title/Abstract]
#20	Physical activities [Title/Abstract]
#21	Walking [MeSH Terms]
#22	Walk* [Title/Abstract]
#23	Resistance Training [MeSH Terms]
#24	Muscle training [Title/Abstract]
#25	Strength training [Title/Abstract]
#26	Endurance training [Title/Abstract]
#27	Interval training [Title/Abstract]
#28	Intermittent training [Title/Abstract]
#29	Fitness [Title/Abstract]
#30	Swimming [MeSH Terms]
#31	Swim* [Title/Abstract]
#32	Bicycling [MeSH Terms]
#33	Bicycl* [Title/Abstract]
#34	Cycling [Title/Abstract]
#35	Cycle [Title/Abstract]
#36	Strengthening [Title/Abstract]
#37	#8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36
#38	#7 AND #37
#39	Randomized controlled trial [Publication Type]
#40	Controlled clinical trial [Publication Type]
#41	Randomized [Title/Abstract]
#42	Placebo [Title/Abstract]
#43	Clinical trial as topic [MeSH Terms]
#44	Randomly [Title/Abstract]
#45	Trial [Title]
#46	#39 OR #40 OR #41 OR #42 OR #43 OR #44 OR #45
#47	#38 AND #46
#48	Animals [MeSH Major Topic] NOT Humans [MeSH Major Topic]
#49	#47 NOT #48

In some instances, there were multiple publications for the same trial; these papers were linked together, and the earliest paper of the trial was used as the primary reference. However, the earliest paper was only used as the reference and all data were extracted from these papers with the most comprehensive available data (*i.e.,* largest available sample size) included for each outcome. Data were extracted from eligible studies and compiled in an Excel spreadsheet (Microsoft Excel v16.04849.1000; Microsoft Corporation, Washington, USA). Trial data for each individual outcome were then extracted and combined in meta-analyses using Review Manager (RevMan 5.3.5, Copenhagen, Denmark).

All trial outcomes were eligible for inclusion following the search, but the primary outcomes were those linked to cardiovascular disease (CVD) risk (*e.g.*, blood pressure, lipids, and glucose). Secondary outcomes were cardiorespiratory fitness, anthropometric measures, androgen levels, pro-inflammatory markers, and psychosocial outcomes.

2.1.3. Assessment of risk of bias in included studies

The Cochrane Collaboration's tool for assessing risk of bias was used; this tool allows assessment of seven specific domains (sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective outcome reporting, and other bias). These domains are associated with five potential sources of bias (selection bias, performance bias, detection bias, attrition bias, and reporting bias). The risk of bias tool also allows reviewers to assess other sources of bias by adding fields they deem to be appropriate to the review. Accordingly, three additional domains were added: group similarity at baseline, adherence to intervention and contamination (*i.e.,* control participants partaking in activities similar to the intervention). Two reviewers (CK and IA) assessed risk of bias, and a third reviewer (IL) arbitrated conflicts not due to assessor error. Recommendations from the *Cochrane Handbook* (Higgins and Altman, 2011) were followed, and each bias parameter was graded as either high, low, or unclear risk.

Studies with >20% of data missing were deemed to have a high risk of attrition bias. We also considered studies with between-group baseline differences that may have affected the outcome (*e.g.*, age, BMI, or activity status), less than 75% adherence to the prescribed intervention, and those with contamination in the control group, as high risk of 'other sources of bias' (Furmaniak *et al*., 2016). It should also be noted that, in exercise trials it is difficult to blind participants and researchers (particularly when supervised) to the interventions. This resulted in a judgement of a high risk of performance bias being made; this does not infer that the methodological quality of the trial is poor, but rather that the inevitable bias related to lack of blinding has been acknowledged by the reviewers. A risk of bias table is presented in the appendices (Appendix 7.1) and risk of bias summarised in the results (Figure 3.2 and Figure 3.3).

2.1.4. Strategy for data synthesis

Where data from ≥2 trials were available, pooled intervention effect estimates and their 95% confidence intervals (CIs) are presented. A random effects model was adopted as it allows for differences in the treatment effect between studies (Riley, Higgins and Deeks, 2010). Meta-analytical methods for involving continuous outcomes assume that data are normally distributed; therefore, we excluded data from the meta-analysis if they were clearly skewed, or if study results were reported with median and range values, and/or non-parametric tests had been used in the subsequent analysis.

All outcomes from each trial were presented as continuous data and, based on the *Cochrane Handbook's* recommendations (Deeks *et al.*, 2011), the random-effects method for meta-analysis was utilised to combine data (DerSimonian and Laird, 1986). Mean ± standard deviation (SD) data for both change from baseline to immediately post-intervention and immediately post-intervention values only were pooled in separate meta-analyses. Where data were missing, the RevMan calculator was used to convert standard errors, CIs or t-values to SD. *A priori*, the analysis was based on the change from baseline data as it removes a large component of between-person variability (Deeks *et al*., 2011). However, analysis of immediately post-intervention data was also included in order to

nullify the effect of selective reporting, and to better indicate whether there was a treatment effect regardless of baseline values.

Where individual trials had reported the same outcome using the same scale, mean difference (MD) was used. Where outcomes were reported using various scales, the units of measurement were converted to the most common measure [*e.g.*, fasting insulin (FI) converted from pmol/L to μIU/mL]. If this was not possible (*e.g.*, relative and absolute values reported for cardiorespiratory fitness), then standardised mean difference (SMD) was used. Where a trial contained multiple intervention arms that were eligible for inclusion (*e.g.*, Almenning *et al.*, 2015; Thomson *et al.*, 2008), the outcome data from both groups were combined using methods recommended in the *Cochrane Handbook* (Deeks *et al.*, 2011). If a trial incorporated a crossover design (Roessler *et al.*, 2013) then only data up to the point of crossover were included in the meta-analysis.

To assess the quality of the evidence for the primary outcomes, the Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) approach was utilised (Schünemann *et al.*, 2013). GRADE is a systematic framework for evaluating the quality and certainty of evidence for making recommendations in clinical practice (Guyatt *et al.*, 2008). If a systematic review is able to provide a best estimate, from all the available evidence, of the effect size for each outcome, then GRADE allows a quality rating across a body of evidence (*i.e.* for each outcome). This is usually completed by taking the lowest quality of evidence from all of the outcomes that are most relevant to the research question and therefore, critical to the decision making process (Langer *et al.*, 2012; Guyatt *et al.*, 2013). GRADE utilises four levels of evidence: high, moderate, low and very low; evidence from RCTs starts out as high quality and is then downgraded based upon predetermined metrics. They are a risk of bias, imprecision around the effect estimate [*i.e.* wide 95% CIs around the effect estimate (Guyatt *et al.*, 2011)], inconsistency, indirectness and publication bias. Accordingly, the findings for the most prevalent outcomes in the current review were graded against these criteria. Those primary outcomes in the review were: systolic blood pressure (SBP), diastolic blood pressure (DBP), blood glucose, FI, homeostatic model assessment of insulin resistance index (HOMA-IR), total cholesterol (TC), low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein

cholesterol (HDL-C), and triglycerides. To develop the 'Summary of Findings' tables, the GRADEpro GDT software was used; two review authors independently graded the quality of evidence for each primary outcome. Quality was not downgraded based upon lack of blinding alone; as stated in Section 2.1.3, this was due to the common difficulties associated with blinding participants (performance bias) and exercise supervising personnel. However, should the blinding of outcome assessors not be adequately reported, risk of detection bias was classified as high- or unclear-risk. However, the quality of evidence was downgraded based on risk of bias only if a lack of blinding was accompanied by high risks of bias in other domains (*e.g.,* selection bias and incomplete outcome reporting).

2.1.5. Investigation of heterogeneity

To evaluate the heterogeneity of results for each outcome, the chi-square (χ^2) test was used to quantify whether between study differences were due to chance alone. Where statistical heterogeneity was identified, the I^2 statistic was used to quantify its magnitude. Although the I^2 is not a measure of absolute heterogeneity, it describes the percentage of variability in the point estimates that is due to heterogeneity rather than sampling error (Deeks *et al.*, 2011). As recommended by Deeks and colleagues (2011), the degree of between study heterogeneity was interpreted as: 0-40% '*might not be important*', 30-60% '*may represent moderate heterogeneity*', 50-90% '*may represent substantial heterogeneity*', and 75-90% '*considerable heterogeneity*'. The importance of the observed I^2 value depends on the magnitude and direction of effects, as well as the strength of evidence for heterogeneity. Accordingly, a visual inspection of each forest plot was completed, and there was an assumption of statistical heterogeneity if little or no overlap of the 95% CIs for the results of individual studies was observed. If there was evidence of at least substantial heterogeneity (≥ 50%), the potential cause was investigated by study population groups. In this instance, the largest outlier was identified, that trial was removed from the analysis and the I^2 was re-assessed. If heterogeneity was not reduced to at least a moderate level (≤ 60%), it was also assessed in subgroup analyses and reported in the results section.

2.1.6. Assessment of reporting biases

To investigate publication bias, if there were ten or more trials included in an individual analysis a funnel plot was used to explore the possibility of small study effects; this is because there is often a tendency for smaller studies to report larger beneficial effects (Sterne *et al.*, 2000). Furthermore, when there is a small number of studies (< 10), the power of tests is too low to distinguish chance from real asymmetry (Deeks *et al.*, 2011). In the current systematic review, only BMI had ≥ 10 studies included within its analysis, so this was the only outcome where a funnel plot was completed.

2.1.7. Subgroup analyses

A subgroup analysis was conducted for each outcome where there were data from > 2 studies. The study characteristics that formed the subgroups for analysis were: BMI upon study entry (≤ 24.9 kg/m^2, 25.0-29.9 kg/m^2 or ≥ 30.0 kg/m^2), intervention type (aerobic exercise, resistance training, or interventions that combined both modalities), intervention duration (≤ 12 weeks or > 12 weeks), and the delivery format of the intervention (supervised, unsupervised, or mixed delivery). Outcome data were separated by subgroup and subtotal summary statistics were presented. The available data were insufficient to complete three of the sub-analyses from the original PROSPERO protocol; they are: exercise intensity, combined treatments, and behaviour change components. However, these findings have been reported qualitatively where available.

2.1.8. Sensitivity analysis

For all outcomes where a statistical effect was observed, sensitivity analyses were completed. This allowed an assessment of the effect of small sample size studies (*n* <30 total participants), and the effect of removing studies with high overall bias risk. Due to the nature of the interventions, performance bias was removed from the reviewer's judgement. Because all studies exhibited at least one domain where risk of bias was unclear, only studies with at least one domain (excluding performance bias) where risk of bias was deemed to be high, were removed.

2.2. Methods for a case-control study: A comparison of self-reported energy expenditure, health-related quality of life, and attitudes towards physical activity in women with and without polycystic ovary syndrome.

2.2.1. Study Design

This protocol has been written following guidelines set out in the 22-item checklist (Appendix 7.4), Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement (von Elm *et al.*, 2008). Favourable ethical opinion for this observational case-control trial was awarded from the School of Life and Health Sciences Ethics Committee at Aston University in January 2019 (Ethics reference 1442).

Recruitment of reproductive-aged women (aged 18-45) with and without PCOS took place between January 9th and May 9th 2019 via advertisements on social media, through PCOS support groups and in multiple online forums hosted by Verity, the UK-based PCOS charity. Inclusion/exclusion criteria are presented in Table 2.3. A range of questionnaires were used to collect the study data; questionnaires were completed online using the survey software, Qualtrics© XM (Qualtrics XM, Provo, Utah, USA); a URL link was either emailed directly to participants who expressed an interest in participation or participants could access the questionnaires directly via the same URL posted on social media channels.

Table 2.3. Inclusion/exclusion criteria for participants in the current study

Inclusion criteria:
1. Reproductive aged women aged 18-45 years
2. A self-reported diagnosis of PCOS, or a self-report of being healthy (*i.e.*, free from any other chronic condition)
Exclusion criteria:
1. Males
2. Females <18 years or >45 years
3. Any chronic health condition other than PCOS
4. Pregnancy
5. Involvement in any exercise intervention in the last 12-months

2.2.2. Sample Size

No single outcome was used to determine the required sample size. However, there was still a need to ensure that the sample size was adequate to detect meaningful between group differences, and to ensure that false positive/negative results were not reported. Therefore, a generic sample size calculation based on a minimally interesting effect size (δ) of 0.5, a power (1-β error probability) of 0.80 and α error probability of 0.05 was used. Data was input into jpower (The Jamovi Project, v1.1.9.0) assuming a two-tailed hypothesis, indicating a sample size of 128 participants (64 per group) was required. The likelihood of reliably detecting various effect sizes is reported in Table 2.4 and a power contour for the current design in Figure 2.1.

Table 2.4. Statistical power by effect size based upon recruitment targets

True effect size	Power to detect	Description
$0 < d = 0.349$	≤50%	Likely miss
$0.349 < d = 0.499$	50-80%	Good chance of missing
$0.499 < d = 0.642$	80-95%	Probably detect
$d = 0.642$	≥95%	Almost surely detect

Key: *d*: effect size.

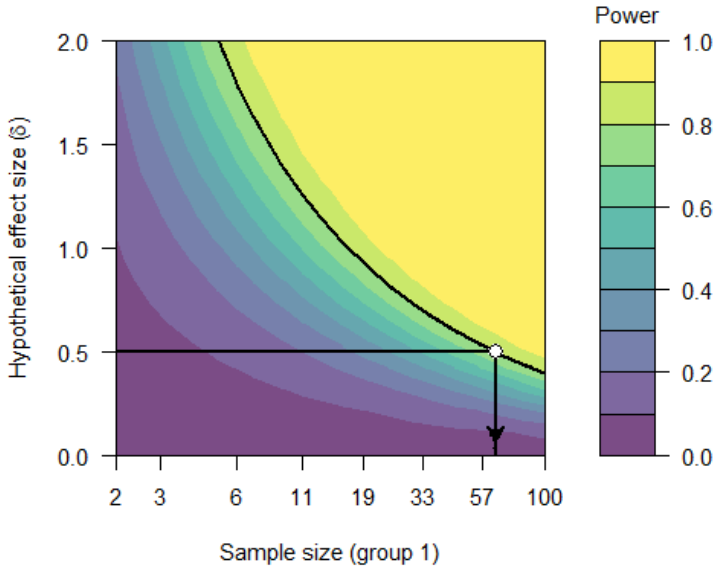


Figure 2.1. Power contour for current study design. The point represents the specified study sample size and effect size.

2.2.3. Variables (Devising the theoretical model)

To explore the relationships between participant's circumstances, the role of PA, and their influence upon health-related quality of life (HRQoL), the constituent parts of Rosenstock and colleagues' (1988) extended Health Belief Model (HBM) were utilised. The HBM is a widely used theoretical model associated with preventative health behaviours; it stipulates that the likelihood of an individual's engagement in health protective behaviours (*e.g.,* PA) depends on their perception of the severity of illness (*e.g.,* PCOS, metabolic syndrome, T2DM) and also the perceived balance of cost versus benefit of taking action (Becker and Maiman, 1975). Those who believe that the severity of illness is great, and that there are more positives than negatives for engaging in PA, are more likely to adopt it as a behaviour. Whilst the HBM was originally designed to focus on disease, and not prediction of PA behaviours (Berger *et al.,* 2015), it was felt that it represented an appropriate framework to guide the selection of outcomes for the path model that is included in this study (Figure 2.2).

The HBM incorporates three major components that may affect an individual's readiness to engage in a specific health behaviour. Whilst it is suggested that the individual's threat perception (*i.e.,* how likely they are to get the disease, and how bad it would be if they did) and response effectiveness (*i.e.,* what are the potential benefits/barriers of engaging?) may be the main determinants of behaviour (Abraham *et al.,* 2016), there are also external variables which are also likely to contribute. Whilst the HBM has always included external variables (*e.g.,* age, gender, socio-economic status, etc.), it wasn't until 1988 that self-efficacy was introduced, which in certain studies, has proved to increase the predictive capabilities of this model (Hay *et al.,* 2003). In order to capture variables that would constitute the external components, a sociodemographic questionnaire was devised, and a measure of self-efficacy for exercise and self-esteem was also implemented. To obtain the outcomes related to response effectiveness, a measure of participant perceptions of benefits and barriers to PA was used, and for threat perception, a disease specific (PCOS) questionnaire was incorporated. The primary outcome of interest is HRQoL; in both cases and controls, physical and mental health was

evaluated using the 12-item Short Form (SF-12) Health Survey and the differences between these populations reported.

As stated above, it was anticipated that PA, as the target health behaviour would mediate the relationship between the constituent components of the HBM and HRQoL. PA-related outcomes are self-reported energy expenditure (MET-mins/wk), achieved through multiple domains of PA, sitting time (mins/wk) and the differences between women with PCOS and their healthy counterparts.

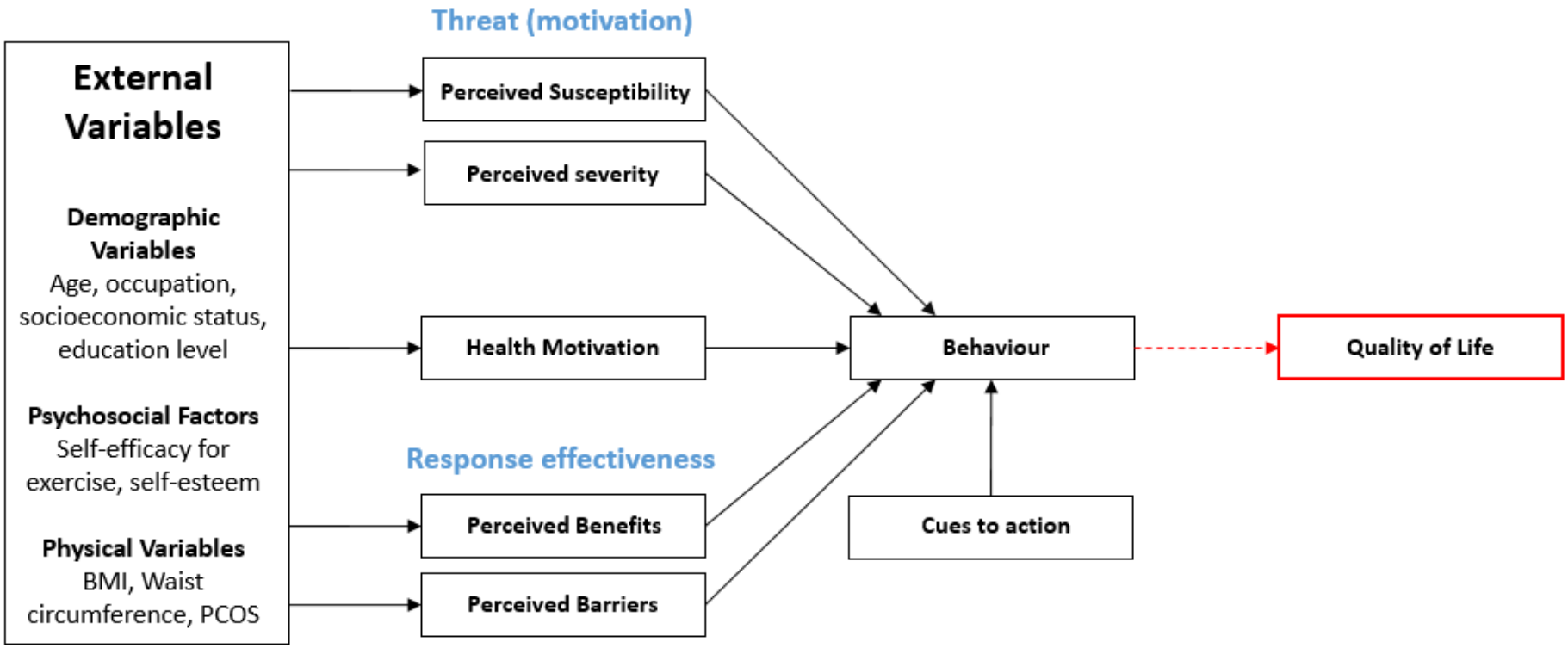


Figure 2.2. Adapted Health Belief Model. The behaviour of interest is physical activity. Red box denotes hypothesised outcome (HRQoL) affected by changes in physical activity and/or mediated by additional components.

2.2.4. Outcome Measurements

Health Related Quality of Life

The 12-item Short Form (SF-12) Health Survey (Appendix 7.5) (Ware *et al.,* 1996) was used to measure HRQoL from the patient's perspective. The SF-12 is a shortened version of a longer generic survey; it presents a subset of 12 items from the SF-36 Health Survey (Ware and Sherbourne, 1992) and provides physical and mental component summary scores. Its shortened length was designed to reduce the burden upon respondents whilst achieving the required standards of precision needed to compare multiple health dimensions between groups. The SF-12 has been previously validated against the SF-36; product moment correlations between the physical (r = 0.94-0.96) and mental (r = 0.94-0.97) summary measures from each questionnaire are reportedly very high, suggesting that the SF-12 is a practical alternative to the SF-36 for use in a number of countries, including the UK (Gandek *et al.*, 1998).

For the path analysis, the individual mental and physical health scores were combined to give an overall composite score for quality of life (Ware, Kosinski, and Keller 1998). Whilst this practice is not commonplace, it was felt a reasonable approach due to a high degree of correlation between each domain score and the composite score and for simplification of the path analysis model.

Participants in the PCOS patient group were also given a PCOS specific questionnaire called the PCOS-Q (Appendix 7.6). The PCOS-Q is a 26-item questionnaire that was developed to assess the impact of PCOS symptoms, and their associated treatments, across five domains, each related to a common symptom of PCOS (Cronin *et al.,* 1998): emotions (eight items), body hair (five items), body weight (five items), infertility (four items) and menstrual problems (four items). Participants respond to each of the 26 items on the PCOS-Q by selecting an answer on a 1-7 scale; seven is representative of optimal function and one the poorest function. Each item is weighted equally when scored meaning that each domain is presented as a score out of seven regardless of the number of items.

In lieu of another PCOS-specific questionnaire, Jones *et al.* (2004) found that the PCOS-Q demonstrated construct validity; statistically significant correlations were reported between the emotions domain from the PCOS-Q and the most relevant domains from the SF-36; that is the role emotional (r = 0.49; P <.01) and mental health (r = 0.62; P <.01) domains. Whilst both the PCOS-Q and SF-36 have previously been used in exercise intervention studies with women with PCOS (Stener-Victorin *et al.* 2013; Vizza *et al.* 2016), the PCOS-Q may allow a greater insight into the key dimensions of PCOS and their effect upon HRQoL (Teede *et al.*, 2018).

Physical Activity and Sedentary Behaviour

Self-reported PA was measured using the International Physical Activity Questionnaire Long Last 7 Days Self-administered Format (IPAQ-LF; Appendix 7.7). The IPAQ-LF is widely used in clinical settings and PA research and asks participants to recall their last seven days of PA, equating this into four comprehensive domains: leisure-time physical activity (LTPA), domestic and gardening activities, work-related physical activity, and transport-related physical activity. The structure then allows separate scoring for walking and either vigorous, or moderate-intensity activity within the other domains. Summation of self-reported PA duration, multiplied by weekly frequency and normative metabolic equivalent of task (MET) data (Ainsworth, *et al.*, 2011) provides continuous data, reported as MET-minutes per week (MET-mins/wk). As sitting is an activity equivalent to 1-MET, sitting time is reported as minutes per week (mins/wk).

When data were analysed, data cleaning instructions from the Guidelines for Data Processing and Analysis of the IPAQ (IPAQ Research Committee, 2005) were followed. The main issues arising were: where minutes had been entered in the hour's column they were manually corrected; where the sum total of all Walking, Moderate and Vigorous time variables exceeded 960 mins/day they were excluded from the analysis; responses of <10 minutes were recoded to zero; for each summed behaviour (*i.e.*, total Walking, total Moderate- and total Vigorous-intensity) that was >180 minutes, the total value was truncated to exactly 180 minutes. These data processing rules mean that outliers are removed and also ensures that highly active people remain classified as highly active, whilst reducing the chance that those who are less active are incorrectly classified.

Criterion validity of the IPAQ-LF has previously shown fair to moderate agreement when compared with a Computer Science and Application's Inc. accelerometer (CSA model 7164, Florida, USA); the pooled Spearman's coefficient (p) = 0.33, and 95% CI were 0.26 to 0.39, although these values increased when categorical estimates of PA (*i.e.* sufficient or insufficient levels) were compared (Craig *et al.*, 2003).

Anthropometric and Sociodemographic information

Participants completed a study-specific questionnaire (Appendix 7.8) to ascertain anthropometric and sociodemographic data for all participants. Questions about the participant's physical characteristics were asked; these included self-reported age, height and weight (from which BMI was calculated as weight in kg divided by the height squared in metres) and waist circumference (centimetres). Furthermore, participants were asked if they had ever been diagnosed with PCOS; if they responded affirmatively, they were asked to specify the time (in years and months) since PCOS diagnosis and to identify the specific phenotype associated with their diagnosis. The four options available were: polycystic ovaries (PCO) menstrual disruption and excess androgens; PCO and menstrual disruption; PCO and excess androgens or; menstrual disruption and excess androgens. Participants could also answer as 'do not know'. Questions about participant ethnicity, marital status, occupational status and education level, whether they have children and their approximate household income were also asked.

Self-esteem

Orth and Robins (2014) define the construct of self-esteem as an individual's subjective evaluation of their own worth as a person; historically, self-esteem has also been viewed as an indication of mental and social life adjustment (Harter, 1989), as well as a mediator of behaviour (Marsh, 1993). In the general population, exercise and PA have often been related to an individual's perceived self-esteem levels (Fox, 1997); whilst a common belief is that improvements in self-esteem may stem from the positive bodily changes (*e.g.*, reduced weight, improved muscle tone, smaller waist circumference) that result from participation in PA (Weinberg and Gould, 2019), there is a counter

argument that self-esteem may be increased from perceptions of improved performance, or additional biological/physical factors that may not be typically associated with physical fitness (Sonstroem, 1997). Although self-esteem (Strauss, 2000; Strong *et al.*, 2005) and PA/exercise (Fox, 1999) are acknowledged as important determinants of psychological wellbeing, less is known about the inter-relationship between the three; increasing PA reportedly improves self-esteem, but it is unclear whether those with higher self-esteem are more likely to be physically active.

Therefore, the Rosenberg Self-esteem (RSE) Scale (Appendix 7.9) was used to measure participant's self-esteem. The scale utilises a four-point Likert scale allowing participants to respond to ten statements about themselves; a higher score indicates a greater level of self-esteem. The RSE Scale has been previously used in studies of women with PCOS (Açmaz *et al.*, 2013), and in some instances (Bazarganipour *et al*., 2014; Bazarganipour *et al.*, 2013) has been used to demonstrate that self-esteem plays an important role, as a mediating factor, in the HRQoL of women with PCOS. The RSE Scale has received extensive psychometric analysis and empirical validation causing Gray-Little *et al.* (1997) to conclude that the RSE Scale provides a highly reliable, and internally consistent measure of self-esteem, thereby justifying its extensive use in psychosocial research.

Self-efficacy for Exercise

The Self-Efficacy for Exercise (SEE) Scale (Appendix 7.10) was administered to assess perceived motivational barriers to completion of PA. The SEE Scale is a revision of the unpublished work of McAuley (1990), a 13-item questionnaire that was intended to measure self-efficacy barriers to exercise. The revision of McAuley's instrument was the result of an interdisciplinary study attempting to understand factors influencing adherence to a walking programme in older adults (Resnick and Spellbring, 2000). Based upon participant responses and feedback, the SEE Scale was devised; respondents are tasked with scoring from 0-10 (zero being not confident and ten being very confident) how confident they are that they could exercise for 20 minutes, three times per week given a variety of situations. The SEE Scale is a 9-item instrument and the total score is calculated by

summing the responses to each question; the scale has a possible scoring range of 0-90 and a higher score is indicative of higher self-efficacy for exercise.

Resnick and Jenkins (2000) used structured equation modelling to test internal reliability and found that the SEE Scale demonstrated evidence of internal consistency (α-coefficient = 0.92). Furthermore, they assessed the construct validity of the scale using the mental and physical domains from the SF-12; when age and gender were controlled, the mental (F = 38.9, P <.05) and physical (F = 24.3, P <.05) domains from the SF-12 significantly predicted participant SEE. In addition, criterion validity was demonstrated as SEE Scale scores were able to significantly predict exercise activity (F = 78.8; P <.05), accounting for 30% of the variance in exercise activity.

Perceived Benefits/Barriers

The Exercise Benefits/Barriers Scale (EBBS) (Appendix 7.11) was used to measure participant's perceived benefits and barriers to participation in exercise (Sechrist *et al.*, 1987*a*). The EBBS requires respondents to rate their agreement with 43 statements (benefit items = 29; barrier items = 14) using a four point Likert scale. Answers are scored from 1 to 4 (*e.g.* strongly agree = 4; strongly disagree =1), with the 14 barrier items being reverse scored; total scores can range between 43-172, with a lower score indicative of fewer perceived benefits and greater perceived barriers. During development (Sechrist *et al.*, 1987*b*), the EBBS demonstrated extremely high internal consistency (Cronbach's α = 0.953) and test-retest reliability (r = 0.89) for benefit barrier (Cronbach's α = 0.886; r = 0.77) items, across a 2-week period.

2.2.5. Risk of Bias

This study was designed with a consideration of reducing potential risk of bias. In this context, the first consideration for this study design concerned the selection of participants. Thus, the included cases (*i.e.*, women with PCOS) and controls in this study are clearly defined, and explicit inclusion/exclusion criteria were defined. However, this study relies upon self-reporting of either

having PCOS or being healthy, and whilst it would have been preferential to have the included cases and controls confirmed by a medical practitioner, this was not feasible in the context of this PhD work. Nevertheless, all study participants were recruited through the same methods (via social media channels) and during the same period, until recruitment targets had been met. The recruitment target was determined by a sample size calculation (Section 2.2.2), and recruitment was ceased once the required number of participants had completed a sufficient amount of the questionnaires. A case would be included in the study only if they had completed at least one questionnaire in addition the demographics information.

Once the study recruitment had been completed, data analysis was completed using a Microsoft Excel spreadsheet. Each participant was assigned a participant identifier code and this was recorded against their membership to either the case or control group on a separate Excel workbook; this allowed the group membership identifier to be removed from the working file and to process each questionnaire response blinded to group allocation. The exception to this were the PCOS-Q responses, as they were only completed by the women with PCOS making blinding impossible.

2.2.6. Statistical Methods

Data analysis: All data analyses were performed on an encrypted laptop. Questionnaires were scored according to their individual criteria and data were collated in a password protected Excel spreadsheet (Microsoft Excel v16.04849.1000; Microsoft Corporation, Washington, USA). Statistical analysis was completed in IBM SPSS (SPSS, v.23.0.0, IBM Corporation, NY, USA) and in IBM SPSS Amos (IBM SPSS Amos, v.25.0.0, Amos Development Corporation, PA, USA).

Due to the sample size (≥ 20), the Shapiro-Wilk test of normality was completed on each variable (Shapiro, Wilk and Chen, 1968); this was performed on the total data set and also split data according to the diagnosis of PCOS response. Non-parametric tests were used for data that are non-normally distributed. Whilst non-parametric tests may be less efficient at detecting genuine between group

differences than their parametric counterparts, they may be more robust as they are less influenced by extreme observations (Kirkwood and Stern, 2003). Non-parametric tests make no assumption about the shape of the distribution, but use the information about the rank of each data point. Furthermore, it is the median values that are compared when testing for differences and the probability of ranked order is tested (Field, 2014). Therefore, where data are non-normally distributed median and interquartile range were reported and independent samples Mann-Whitney U tests were completed to highlight between group differences. The Mann-Whitney U test compares the ranks of observations in two observed groups; data from the two groups are pooled and the ranks of each observation calculated. The rank of each group are separately summed and termed R_1 and R_2 and then entered into the following calculations to give the tests statistics U_1 and U_2 (Ennos, 2007):

$$U_1 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2$$

$$U_2 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

In these formulae, the sample size of each group (*i.e.,* Women with PCOS or Control) is represented by n_1 and n_2, respectively. The smaller *U* value is then compared to the critical values table for the relevant group sizes and the null hypothesis (*i.e.,* there is no difference in the median between groups) is rejected if *U* is lower, or equal to a critical value. Between group median difference, effect size (Cohen's *d*) and significance values (*P*) were reported for all non-parametric outcomes.

Where data are normally distributed, mean ± standard deviation (SD) were reported, and if there was no violation of the assumption of equal variances, as tested by Levene's test, a Student's t-test was used to test for differences between women with PCOS and their healthy counterparts. Should the assumption of equality of variance be violated, then Welch's t-test was used. Mean difference (MD), 95% confidence intervals (95% CIs) of the MD, Cohen's *d* and *P* were reported for all parametric variables. In these analyses, pairwise exclusion was used to deal with missing values.

A separate analysis was completed on the domain scores from the PCOS-Q. Because these data were non-parametric, a Durbin-Conover pairwise comparison was used to identify statistical differences between domains. Data for each comparison (median difference and *P*) are reported in Table 4.4.

Correlations: Due to the prevalence of non-parametric variables, Kendall's rank correlation (τ_b) was chosen to measure the strength of association between two variables. Kendall (1970) developed a rank correlation coefficient that offers an alternative to its counterpart, the Spearman's rank correlation coefficient. Kendall's τ_b is calculated in the following way: for each pair of subjects it is recorded whether the subjects are rank ordered in the same way for the two variables of interest, that is a concordant pair (P), ordered in opposite ways, a discordant pair (Q), or equal for one of the variables, a tied pair (X_0 or Y_0). Therefore, τ_b is the proportion of concordant pairs minus the proportion of discordant pairs; τ_b would be +1.00 if the rankings are identical, and -1.00 if they are exactly opposite (Bland, 2015):

$$\tau_b = \frac{P - Q}{\sqrt{(P + Q + X_0)(P + Q + Y_0)}}$$

Where variables were highly correlated and deemed to be reporting similar effects (*e.g.,* body mass, BMI and waist circumference), the variable with the largest sample size was retained. Where these variables were domains from a questionnaire (*e.g.,* mental and physical domains of the SF-12), the total score was used as the variable in the regression.

Path Analysis: The SPSS data file was transferred to SPSS Amos. In order to generate a complete data set, full information maximum likelihood (FIML) regression imputation was used to account for missing data. Although pairwise or listwise deletion are commonly used by many applied researchers, it was felt that the use of FIML was favourable in order to preserve the sample size. Whilst other methods reportedly have no solid theoretical basis for their use, Enders and Bandalos (2001) found that when dealing with missing data, FIML provided data estimates that were unbiased and more efficient than other methods. FIML also yielded the lowest proportion of convergence failures and provided near optimal Type I error rates.

HRQoL, as measured by the SF-12, was used as the endogenous (dependent) variable; the remaining variables were arranged into a path model (Figure 4.3) to indicate causal relationships between the exogenous (diagnosis of PCOS, and BMI), mediating (self-efficacy for exercise, self-esteem, perceived benefits/barriers of exercise and MET-mins/wk) and endogenous variables (SF-12 total scores). A diagnosis of PCOS, and BMI were selected as exogenous variables because it is assumed that their variance is caused entirely by factors outside of the causal model (Wuensch, 2016).

Because the causal model has two exogenous variables, the standardised path coefficients (β weights) are in fact partial regression coefficients which measure the direct effects that one variable has upon another, whilst controlling for a set of correlated exogenous variables (Sarwono, 2018). A diagnosis of PCOS serves as one of the exogenous variables in the path model; this is a binary variable (*i.e.*, respondents could only answer either yes or no) which functions only as a predictor variable in the regression equations, and therefore does not affect the methods of the mediation analysis (Breisch and Rajagopal, 2010). Standardised path coefficients are calculated from the unstandardised regression coefficient using the following formula:

$$Standardised\ regression\ coefficient\ = Unstandardised\ \beta\ weight \times \frac{\sigma_{independent\ variable}}{\sigma_{dependent\ variable}}$$

In addition to the direct effects of variables upon the endogenous variable, indirect effects from both the exogenous and mediating variables partially account for some of the variability in the model. Indirect effects are calculated by multiplying the standardised β weights along each causal pathway; the indirect pathways, via all mediating variables are then summed to give the total indirect effect of the independent variable upon the dependent. To calculate the total effect, the direct effects are added to the indirect effects. Standardised regression coefficients, and their 95% CI's for direct, indirect and total effects are presented in Table 4.7. In addition, a squared multiple correlation (R^2) is calculated for each endogenous variable in the model; the R^2 is indicative of the proportion of

variance in the dependent variable (*i.e.*, HRQoL) explained by the other variables in the path model. This is calculated using the following formula:

$$R^2 = 1 - \frac{\sigma^2\ residual}{\sigma^2\ endogenous}$$

2.3. Methods for a cluster analysis: Clustering of cardiometabolic risk factors and their association with physical activity and sedentary time in women with and without polycystic ovary syndrome.

2.3.1. Study design

The UK Biobank is a large open access national health resource with the aim of improving preventative measures, diagnostic processes and treatment of a wide range of serious diseases. Between 2006 and 2010, approximately 500,000 UK adults were enrolled following an invitation through population-based National Health Service (NHS) patient registers (Sudlow *et al.*, 2015). Enrolled participants were aged 40-69 years and had to live within 25 miles of one of the 22 Biobank assessment centres located in England, Wales or Scotland (Littlejohns *et al.*, 2019). All participants provided written informed consent before providing baseline measures; they answered questions (either electronically or verbally) on sociodemographic, lifestyle (including physical activity), and health-related characteristics. They also provided blood, urine, and saliva samples allowing for data on a range of biomarkers, many of which are associated with CVD development [*e.g.*, blood levels of lipids, glucose, and C-reactive protein (CRP)] and/or manifestations of PCOS [*e.g.*, testosterone, oestradiol, and SHBG] to be collected.

The UK Biobank were granted ethical approval by the North West Multicentre Research Ethics Committee, and the National Information Governance Board for Health and Social Care in England and Wales, and the Community Health Index Advisory Group in Scotland. As the aims of the current study fall within the remit of the UK Biobank project objectives, no further ethical approval was required from Aston University, or any other local institution or organisation. However, to obtain the required data, an application was made to the UK Biobank (Application Reference Number 52771), in which the objectives and methods were reviewed and approved; furthermore, the host institution (Aston University) had to agree to a material transfer agreement (MTA) and pay an access charge.

As per the application process, the outcomes that would be investigated had to be specified prior to the data release. The included outcomes were selected according to diagnostic features of metabolic

syndrome (waist circumference, blood pressure, blood glucose, and lipid profile). The physical (weight, BMI, cardiorespiratory fitness, inflammatory and androgenic biomarkers) and psychosocial (anxiety, depression, mood, self-esteem, and quality of life) characteristics that are often associated with PCOS were also included. It was also requested that outcomes related to self-reported PA, exercise and sedentary behaviour, in addition to objectively measured PA were included.

2.3.2. Data Handling

Once the UK Biobank had approved the study application, data were made available through their Application Management System (AMS). The main dataset was supplied in an encrypted format and required several steps in order to be accessed. The first step involved the downloading of a series of 'helper' programmes in order to decrypt and convert to the main dataset. They are: *ukbmd5* for ensuring the encrypted main dataset has downloaded correctly, *ukbunpack* for decrypting the main dataset, and *ukbconv* for converting the decrypted dataset into a suitable format [*i.e.* a Comma Separated Values (csv) file). Once these were installed, the main dataset was downloaded through the AMS; a 32-character authentication code (MD5 checksum) was supplied when access to the data was granted and it was used here to initiate the download. Once downloaded, the main data file needed to be validated; a Windows command prompt was opened and *ukbmd5* was run to complete the validation. Next, the *ukbunpack* was run to decrypt and decompress the downloaded file into a custom UK Biobank format. Finally, the dataset was converted using *ukbconv* into a csv file suitable for use in Microsoft Excel, SPSS, and Jamovi. The converted file was immediately password protected and saved onto an encrypted, Aston University owned laptop.

The final dataset included all female participants from the UK Biobank (~250,000 participants). From the full dataset, women with PCOS were identified; the inclusion for a 'case' participant was a diagnosis of PCOS as recorded by the UK Biobank. When selecting case-control participants from datasets such the UK Biobank, the case group should comprise all (or a high proportion) participants diagnosed with the outcome disease (*i.e.* PCOS), whereas the referent or control group(s) should be a random sample that are free from the disease at the specified sampling point (Läärä, 2011). It is

however important that within each group (*i.e.*, case and each control) that no participant is sampled in more than one group. Although the comparator groups in the present study were age and/or BMI-matched, they were randomly selected from the matched sample.

In the present study, all participants with PCOS were identified in Microsoft Excel (v.16.0.4849.1000); this was done using the International Classification of Disease (ICD-10) codes that define participant characteristics in the UK Biobank. The ICD-10 for PCOS is E28.2, which appeared in the dataset as E282. Using the cut and paste function, each row of data for women with PCOS was transferred onto a separate tab on the same Excel workbook and the Sort function was used to organise them according to their age at assessment (*i.e.* from the youngest to the oldest). The next phase involved the selection of two comparator groups: 1) a non-PCOS age-matched cohort and 2) a non-PCOS age- and BMI-matched cohort. These participants were selected concurrently as follows:

Age-matched cohort: all eligible controls (*i.e.*, all participants without PCOS) were also ordered in the dataset based upon their age at assessment (youngest to oldest). Participants were then matched on a two-to-one ratio for each age category (*i.e.*, 40 years, 41 years, 42 years, *etc.*) of women with PCOS. For example, if there were 15 women with PCOS that were 40 years old, then the first 30 women that were aged 40, were extracted from the control dataset and pasted into a separate tab of the same Excel workbook.

Age- and BMI-matched cohort: similarly, the sort command was used to identify these participants. The participants with PCOS were sorted based upon their age (youngest to oldest) and then each age category was sorted based upon their BMI values. For example, if there were 15 women with PCOS that were 40 years of age, they were all selected as one age category. The sort command was used to arrange them with the lowest BMI value first, ascending to the highest. The same procedure was repeated in the dataset of non-PCOS women and once again, the non-PCOS group were allocated on a two-to-one ratio. To identify those participants who would be included, each BMI value from the cohort with PCOS was recorded and then within the control dataset, the two BMI values that were

closest, and within the corresponding age category determined which participants would be included. Once identified, these data were cut from the non-PCOS data and pasted onto a separate tab on the same Excel workbook. To allow statistical analyses, each cohort was labelled (either PCOS, Age, or BMI + Age) and the data were merged into one dataset.

In addition to the above data, the ICD-10 was also used to capture non-PCOS morbidity incidence across all three study groups.

2.3.3. Data Analysis

Statistical analysis was completed in Jamovi (The Jamovi project: v1.2.16). Descriptive statistics were presented for each outcome according to individual cohort. Mean and SD were reported alongside median and interquartile range (IQR); the number of missing values (and corresponding cohort percentage) were also reported. Assumptions of normality were checked in Jamovi by completing Shapiro-Wilk tests of normality and analysis of skewness and kurtosis values. None of the included outcomes met the assumption of normality across all three groups; therefore, the following non-parametric analysis was adopted: a Kruskal-Wallis [one-way analysis of variance (ANOVA) on ranks] to compare outcomes in the three independent groups (PCOS vs Age vs BMI + Age). A statistically significant Kruskal-Wallis test is indicative that the outcome's median value in at least one group is different to the median of at least one other group. In the current study, levels of significance were set at $P < .05$, and in order to gauge the magnitude of influence, an effect size was also reported for each comparison. When completing a Kruskal-Wallis test, there is no definitive effect size statistic that should be calculated to report magnitude of influence. Many statisticians use eta-squared (η^2) or partial eta-squared (η_p^2) but these have been cited as biased effect size estimators (Albers and Lakens, 2017) and their use as an effect size estimator in ANOVA is discouraged (Skidmore and Thompson, 2013). A less biased alternative to η^2 is the epsilon squared (ε^2) effect size (Lakens, 2015); whilst a ε^2 of zero would indicate no effect, there is no formal way to determine the magnitude of an effect greater than zero. In this instance, recommendations from Rea and Parker (1992) have been used to report the strength of the observed effect. That is: <0.01 as negligible; 0.01

< 0.04 as weak; 0.04 < 0.16 as moderate; 0.16 < 0.36 as relatively strong; 0.36 < 0.64 as strong; and > 0.64 a very strong.

What the Kruskal-Wallis test does not report is, if differences are present between groups, where do the differences occur, and for how many pairs of groups there is a difference. For the analysis of specific group pairings, a post-hoc test is required. The Dwass-Steel-Critchlow-Fligner (DSCF) post-hoc test is often used when the null-hypothesis is rejected following a Kruskal-Wallis test. The DCSF test is a two-sided, non-parametric procedure that provides family-wise error (FWE) protection; FWE refers to the probability of making at least one type I error when completing a series of statistical tests (*i.e.*, at least one false conclusion when multiple hypotheses are tested) (Hochberg and Tamhane, 1987). Following the one-way Kruskal-Wallis test, the DCSF test procedure is used to calculate a Wilcoxon (W) rank sum test statistic on each pair of groups with the "family" (Hollander, Wolfe, and Chicken, 2014). If data are tied, the average ranks are used in the individual rank sum statistics and statistical significance levels (*p*-value) are calculated using the distribution of the range of independent, standard-normal variables. It is this mechanism that differs from typical Mann-Whitney U calculation and that provides the FWE protection. Therefore, the DCSF post-hoc test was completed on any Kruskal-Wallis analyses that were statistically significant; between pair significance are identified in bold text in Table 5.5.

2.3.4. Categorical data

Where data were categorical (*e.g.*, IPAQ activity group, meeting PA guidelines, health ratings and satisfaction scores), Chi-square (χ^2) and tests of frequencies were completed to determine whether there were any statistical difference (P <.05) between expected and observed frequencies based upon number of respondents.

2.3.5. Cluster analysis for body composition and cardiometabolic outcomes

Cluster analysis is a set of data reduction techniques that are used to identify natural groupings (or clusters) within a dataset. The objective of cluster analysis is to group cases (*i.e.* participants) together so that all observations within the same cluster are as similar to each other (homogenous) as possible, and conversely, cases in each cluster differ as much as possible to those in other clusters (Kaufman and Rousseeuw, 1990). Where cluster analysis varies from other data reduction techniques (*e.g.*, factor analysis or principal component analysis) is that whereas they tend to group data by similarities across variables of a dataset (or columns), cluster analysis groups cases based upon similarities across rows (Columbia Public Health, 2020)

A typical measure of similarity between variables is the correlation coefficient, which provides information about the degree of corresponding changes between variables. However, it is not the most suitable parameter to make a comparison of cases across variables. Although a simple correlation coefficient may provide information about whether the pattern of responses or measures are similar between cases, it does not provide any information about the distance between each case (Field, 2000). An alternative measure for comparing cases, particularly for continuous variables, is the Euclidean distance (Everitt *et al.*, 2011). The Euclidean distance is the geometric distance between two objects (cases in this instance). If cases are labelled *a* and *b* then their Euclidean distance can be expressed in terms of the following equation:

$$d = \sqrt{(a_1 - b_{1)} + (a_2 - b_2) \cdots (a_n - b_n)} = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2}$$

In essence, this formula reveals the difference between cases (a – b) by taking their scores on each variable (i) and calculating the difference. Because differences can be negative or positive, differences need to be squared before they are added together. Once the summed squared differences have been totalled, the square root is taken to give the measure of variance. The smaller the Euclidean distance, the greater the degree of similarity (Field, 2000). Before cluster analysis can be completed, several pre-analysis checks are needed to ensure the data is suitable for these analyses. These checks

are: 1) rows must be observations (participants) and columns should be variables, 2) Any missing data should be removed or imputed, and 3) all data must be standardised to make variables comparable (Kassambara, 2017).

2.3.6. Outlier Identification

All inferential statistics are generally sensitive to outliers and cluster analyses are no different. Indeed, values that are far from the mean can alter results considerably. Robust statistical analysis typically deals with the location of the large heterogeneous proportion of data alongside an unknown number of outliers (Hennig, 2003). When outliers are identified, a decision must be made about how to treat them. Typically, a researcher may return to the data collection instrumentation to determine whether the outlier is due to an equipment malfunction or a data entry error (Parke, 2013). However, because the current study utilises data from the UK Biobank, it is not possible to check the origin of the outlier, but in order to preserve as many cases as possible, it is also undesirable to simply delete the data.

Descriptive statistics for each group (*e.g.*, PCOS, Age-matched and Age + BMI-matched) were reported for the outcomes selected for the cluster analysis. In addition to median and IQR, the value of the 5^{th} and 95^{th} centiles were calculated (Appendix 7.13). The IQR indicates how spread out the middle values of the data are for each group, and can also identify outliers. If a data point was below 1.5* IQR the 25^{th} centile, or 1.5* IQR above the 75^{th} centile it was marked as an outlier. Box plots were created for each outcome and data labels on each box plot identified those cases that lay beyond the pre-specified limits (*i.e.*, Q1 - 1.5*IQR or Q3 + 1.5*IQR) for each outcome; these outliers were then transformed to a specified percentile of the data. In this instance, a 90^{th} percentile Winsorization was performed that transformed all data below the 5^{th} percentile to the 5^{th} percentile and all data above the 95^{th} percentile to match the 95^{th} percentile (Aguinis, Gottfredson and Joo, 2013). Winsorization is not without drawbacks; systematic bias is introduced into results of analyses, although the degree of bias is less than if truncation (*i.e.* deletion of data point or case) had been performed. The alternative approach would have been to leave the outlier in place, but this may have

heavily influenced the manner in which clusters were formed (Garcia-Escudero and Gordaliza, 1999).

2.3.7. Cluster variable selection

Variables used to determine clusters should be a thorough representation of the underlying construct of interest (Everitt *et al.* 2011). In this instance, the objective was to investigate whether outcomes associated with metabolic health cluster differently in each study population. Whilst variable choice is highly important, the consensus of many statisticians is that clustering should utilise all possible variables as long as they fit the model of interest (Steinbach, Kumar and Tan, 2005). Furthermore, variables that do not describe a great deal of the variance in Euclidean distances between observations should have a lower weighting when assigning cases to a cluster (Field, 2000).

Whilst cluster analysis should include as many variables as are relevant, there is also a need to ensure that variables are not reporting similar phenomena, or that they are made up as a product of other variables. Therefore, Pearson's correlation coefficient analysis was completed for all complete variable pairs that were deemed relevant to the study objective; that is outcomes associated with metabolic health, or those deemed to be important for PCOS. The analysis revealed that many outcomes were statistically significant so, where it was felt that variables were reporting similar effects (*e.g.*, BMI, waist circumference and body fat percentage), the magnitude of correlation was considered, along with the completeness of data before outcomes were removed from subsequent analysis.

The variables selected for subsequent analyses were: waist circumference (cm), Diastolic BP (mmHg), Systolic BP (mmHg), C-reactive protein (mg/L), HbA1c (mmol/mol), HDL-C (mmol/L), IGF-1 (nmol/L), LDL-C (mmol/L), SHBG (nmol/L), testosterone (nmol/L), and triglycerides (mmol/L).

2.3.8. Missing value analysis

With regard to the completeness of data, an analysis of missing values was completed in SPSS (SPSS, v.23.0.0, IBM Corporation, NY, USA). Missing data can be defined in two ways: user missing data or system missing data. User missing data may be coded by the user due to incompatibilities between case and variable (*e.g.,* length of pregnancy in males), whereas system missing data are missing data that is not present in the dataset (Heymans and Eekhout, 2019). Further investigation of the complexity of missing data patterns (*i.e.* where values are systematically missing together) was investigated using the Missing Value Analysis and Multiple Imputation menu items in SPSS.

Initially, missing data for each outcome was summarised, along with the number of participants with incomplete data. This confirmed that there is at least one missing value for each outcome to be included in the cluster analysis. Analysis of incomplete data also revealed the percentage of cases that have incomplete data; using listwise analysis, a substantial proportion of the sample would be absent from the statistical analysis. Initially, missing data were summarised (Appendix 7.14 and 7.15) and missing value patterns were identified from the output (Appendix 7.16). The ten most common missing value patterns were investigated and where recurring patterns were evident, these were summarised in the results. The volume of recurring patterns, (*i.e.,* where values were often missing together) was indicative that data was missing not at random (MNAR) and this assumption was tested using Little's (1988) MCAR test. Collectively, these recurring patterns are strong indicators that these data are not missing completely at random (non-MCAR); this notion is further supported by the result of Little's MCAR test (χ^2 = 478.272, df = 354; P <.001) meaning that missing value multiple imputation can be completed.

Jakobsen *et al.* (2017) outline criteria that would justify the use of multiple imputation. They state that complete case analysis can be completed if the proportion of missing data is below five percent; if the potential impact of missing data is deemed negligible, it can simply be ignored in the final analysis (Jakobsen *et al.*, 2014). In contrast, if large proportions (> 40%) of data are missing from key variables then complete case analysis could also be used with the caveat that limitations of findings are clearly discussed. If imputation methods are used to replace missing data in these

circumstances, then study results should only be used to generate hypotheses (Clark and Altman, 2003). Finally, if the missing completely at random and missing at random assumptions are not met based on the characteristics of missing data, results have an increased risk of incomplete outcome bias (Higgins and Green, 2011). Whilst no statistical methods are guaranteed to negate this bias (Sterne *et al.*, 2009), multiple imputation can be used to estimate these missing values.

2.3.9. Data Imputation

Data imputation was completed in SPSS. All of the outcomes selected for inclusion in the model (based upon correlations and missing data) were selected for imputation; five iterations saved within a new dataset were requested. For the imputation methods, a custom approach was adopted and fully conditional specification (FCS) selected. FCS is the Bayesian approach to imputation and accounts for imputation uncertainty by both adding error variance to the predicted values and by accounting for the uncertainty in estimating the regression coefficients of the imputation model (Heymans and Eekhout, 2019). Bayesian theory states that there is not one "true" regression coefficient but that the coefficients themselves follow a distribution (Gelman *et al.*, 2014). Following this, predictive mean matching (PMM) was selected as the model type; PMM is the default procedure when multivariate imputation by chained equations is used and is the method that predicts (and selects) data to replace missing values (Rubin, 1987).

Pooled statistics were reported for each of the included outcomes and the dataset was combined (and original data removed) to allow clustering to be completed on the data. Prior to the cluster analysis, standardisation of the data was completed. Standardisation of variables is a necessary step to take because dissimilarity measures, such as Euclidean distance, are extremely sensitive to variability in the magnitude of scales (Milligan and Cooper, 1988). Standardisation of values not only equalises the magnitude and the variability of the variables within the model but also their relative weighting (Anderberg, 2014; Romesburg, 1984). Standardised values are derived from Z-scores; Z-scores are simply values that have been given a common standard, which is a mean of zero and a standard

deviation (SD) of one. Z-scores are calculated by first subtracting the mean score from each individual score and then dividing the remainder by the SD of the overall mean:

$$Z_x = \frac{X_i - \bar{X}}{SD_x}$$

Z-score standardisation is a type of linear transformation; transformed scores follow the same distribution as that of the original data and, if plotted against each other, data points would be linear.

2.3.10. Cluster Analysis

Once outcome scores had been standardised, a cluster analysis was completed using the non-hierarchical K-means clustering in SPSS. K-means clustering divides observations into discrete groups based upon the Euclidean distance between the mean of included outcomes. The primary objective of the K-means algorithm is to minimise the sum of distances between the points and the cluster centroid to which they belong (Everitt *et al.* 2011). The K-means algorithm utilises iterative refinement in order to find the best solution; that is assigning points to a cluster based upon the size of the squared Euclidean distance, recalculation of the cluster centroid and then repeating until convergence is achieved. The algorithm has reached convergence when cases are no longer reassigned when the centroids are recalculated (Hartigan and Wong, 1979).

K-means analysis should be completed when variables are quantitative and at either the interval or ratio level (as in the current study); hierarchical clustering would be used if data were binary or counts. To complete the analysis in SPSS, all standardised variables to be included in the model were selected and the number of clusters (*k*) were selected as two. The maximum number of iterations were set to 15 with a convergence criterion of zero, meaning that no new iterations would be generated when a complete iteration does not move any of the cluster centroids. Cluster membership was recorded for each case, which created a new categorical outcome within the dataset. It was also specified that initial cluster centres would be recorded and an ANOVA test between clusters for each outcome was generated. To assess the strongest solution (*i.e.* number of clusters) this process was

repeated using different numbers of clusters (up to five) and comparing results. The number of iterations before convergence was achieved, alongside the between cluster ANOVA results were used as confirmations that the strongest solution had been found. The optimal number of clusters identified was two and this was achieved after 12 iterations.

Once cluster membership for each case had been determined, the categorical variable was used to generate descriptive statistics for each cluster. Furthermore, the number of cases (and relative percentages) from each study group and their cluster membership was reported; a graphical representation of each outcome, split by study group and cluster membership was also reported. Finally, the descriptive statistics for outcomes relating to PA, sedentary time, and co-morbidities were reported for each cluster and the appropriate independent samples t-test (Mann-Whitney) was completed for each outcome.

2.3.11. Classification using physical activity and sedentary time

For this analysis, participants were categorised according to their weekly physical activity level (MET-mins/week from IPAQ responses) and their time spent (h/day) in sedentary behaviour (SB) (summed screen time and driving time). The four categories are High PA + Low SB, High PA + High SB, Low PA + Low SB and Low PA + High SB. The groups were defined using quartiles from the respective outcome. Low PA included data from quartiles 1 and 2 of the MET-mins/wk outcome and High PA from quartiles 3 and 4. Similarly, Low SB used data from quartiles 1 and 2 of the summed sitting time variable, with high SB coming from quartiles 3 and 4 (Engelen *et al.*, 2017). The visual binning function in SPSS was used to identify these quartiles and allocate participants into a new categorical variable.

Once categories had been determined, a frequencies analysis was run to determine how participants had been allocated into groups based upon their cohort classification; these data were reported graphically and participants with missing data (*i.e.* missing MET-mins/week or missing sitting time) were removed from subsequent analyses. Shapiro-Wilk tests of normality were conducted, and a

non-parametric approach was then used. Descriptive statistics (median and IQR) were reported and a Kruskal-Wallis test completed to assess variability between groups for each outcome. Where statistically significant effects were reported, pairwise comparisons (Mann-Whitney U) were completed and results reported either graphically or in a tabular format.

Finally, to calculate the probability of disease incidence between groups (*i.e.* low-risk *vs* high-risk behaviours), logistic regression analyses were used in SPSS. The disease of interest was set-up as a dichotomous variable (*e.g.*, diagnosed or not diagnosed) and then input as the dependent variable in the model. The PA risk categories, High PA + Low SB and Low PA + High SB were selected as a categorical covariate in the model and the analysis run. The exponential of the coefficient [exp(β)] is called the odds ratio that describes, in this instance, the odds of disease incidence between the two groups. However, the odds ratio was transformed into a probability score using the formula: P = odds / (1 + odds). This allows reporting of the increase in probability that a participant allocated in the high-risk group may have additional morbidities.

Chapter 3: Results

Study 1.

Exercise, or exercise and diet for the management of polycystic ovary syndrome: a systematic review and meta-analysis

3.1. Introduction

Polycystic ovary syndrome (PCOS) is the most common endocrinopathy in reproductive-aged women, that depending upon the studied population and the applied diagnostic criteria (Table 1.2), affects 6-21% of this female population worldwide (Boyle *et al.*, 2012; Lizneva *et al.*, 2016; Ma *et al.*, 2010). PCOS is characterised by hyperandrogenism and/or chronic anovulation and typically manifests in a host of symptoms, including hirsutism, acne, menstrual disruption, and infertility (Costello *et al.*, 2007). PCOS is also associated with an increased risk of cardiometabolic complications, such as hypertension, dyslipidaemia, insulin resistance (IR), and type 2 diabetes mellitus (T2DM; Azziz *et al.*, 2016). Furthermore, PCOS is also associated with increased psychological morbidity [*e.g.*, increased risk of stress, depression, low self-esteem, poor body image, and reduced health-related quality of life (HRQoL)] (Himelein *et al.*, 2006; Weiner *et al.*, 2004).

The exact PCOS aetiology is yet to be fully defined, but increased body weight, particularly when associated with adiposity, is considered pivotal in its manifestation and severity of symptoms (Sam, 2007). It has been reported that ~90% of women with PCOS also have a body mass index (BMI) > 25 kg/m^2, but encouragingly, even moderate weight loss (*e.g.* 5%) has been shown to elicit clinically meaningful reductions in hyperandrogenism and to improve menstrual regularity (Kiddy *et al.*, 1992; Balen *et al.*, 1995; Holte *et al.*, 1995; Legro, 2000; Sir-Petermann *et al.*, 2009). Of note, women with PCOS are also reported to have a greater severity of IR than weight-matched women without PCOS (Dunaif *et al.*, 1989; Norman *et al.*, 2001). Additionally, the increased susceptibility to obesity that is associated with PCOS (Glueck *et al.*, 2005) further exacerbates IR and the accompanying metabolic (Legro *et al.*, 2001; Ehrmann *et al.*, 2006) and reproductive (Balen *et al.*, 1995; Kiddy *et al.*, 1990) dysfunctions. Therefore, women with PCOS exhibit increased risk of impaired glucose tolerance (IGT) and T2DM regardless of weight and age (Legro *et al.*, 1999).

Because no curative treatment for PCOS is currently known, management of women with PCOS and overweight/obesity focuses on weight loss through lifestyle changes, namely increased exercise/physical activity and caloric restriction through improved diet. The ultimate aim of these lifestyle changes is to alleviate PCOS's clinical manifestations and lower the related risk of T2DM

and cardiovascular disease (CVD; National Institute for Health and Care Excellence, 2013). Exercise interventions have been shown to be beneficial to the health of different IR populations, independent of weight loss (Boulé *et al.*, 2001; Thomas *et al.*, 2006; Yang *et al.*, 2014), therefore, the incorporation of moderate-intensity exercise in the treatment of PCOS may be favourable. The current evidence tends to support this notion; although most studies incorporating exercise interventions in women with PCOS report little or no weight loss (Azziz *et al.*, 2016), exercise has been shown to have favourable effects on IR, body fat levels and distribution (*i.e.* central or peripheral adiposity), and CVD risk in these patients (Harrison *et al.*, 2011).

3.1.1 Rationale for a systematic review

In contrast to available published systematic reviews at the time of protocol registration (Harrison *et al.*, 2011; Moran et al., 2011; Domecq et al., 2013; Haqq et al., 2015), the current systematic review proposed to make a greater number of comparisons (Section 3.2) and include more outcomes than similar previous reviews. Although some components of the current systematic review may draw parallels with previous lifestyle reviews, it is notable that the most recent prior searches were completed in 2009; it is therefore reasonable to assume that there is a substantial body of new literature that has not been subjected to analysis. Furthermore, no previous review has made all of the comparisons that are being made in the current review (Table 3.1).

From the PROSPERO website, at the time of protocol registration four additional systematic reviews were identified, and their working status was classified as ongoing. Two of these were considerably over their anticipated completion date (Mani et al., 2013; Lundgren, Moholdt and Riphagen, 2015), whilst the literature searching for the current review did not locate any related evidence of publication or dissemination of findings from these authors. A third study (Santos *et al.*, 2017) stipulated that their ongoing review is an update of previous work, but again, a final published manuscript, or indeed the original work was not located in this literature search. The fourth study Benham *et al.* (2018) has since been published and an overview of this study is presented in Appendix 7.2. Whilst there may be some overlap between the outcomes assessed in the current review and those of Santos *et al.*

(2017) and Benham *et al.* (2018), the focus of the review and the primary aims are different. The primary aim of Santos *et al.* (2017) and Benham *et al.* (2018) was to investigate the effect that exercise may specifically have upon reproductive health when compared to usual care or a no-exercise control group.

The current review presented here aimed to directly measure the effect of exercise in PCOS women and this was done in part by comparing intervention groups to a control or no treatment group. Where this differs to Santos *et al.* (2017) and Benham *et al.* (2017) is that studies including a third arm or combined treatment arms were included providing that the effect of exercise can be isolated. This provided an opportunity to complete a sub-analysis of treatment variations, where feasible; for example, interventions that contain dietary elements compared to those that do not, home-based or unsupervised programmes compared to those that are supervised. Also, studies that include behavioural components, versus those that have none, can be contrasted and have their varying effectiveness assessed. It is assumed that varying methodologies will influence participant adherence and attrition rates and this was also investigated.

In addition, the current systematic review also completed subgroup analyses where data were available. Sub-analyses to assess the effects of exercise mode (*e.g.*, aerobic exercise versus resistance training), duration, frequency, and intensity, method of intervention delivery (*e.g.*, supervised versus unsupervised), and participant BMI upon study entry (*e.g.*, ≤ 24.9 kg/m^2 versus ≥ 25.0 kg/m^2) were completed. There were insufficient data to complete sub-analysis on the use of pharmacological treatments, or the effects of adding a behaviour change component. However, intervention adherence and attrition rates were reported.

Where previous systematic reviews have only completed risk of bias over a score-based assessment, there may be methodological concerns. Risk of bias tools typically give equal weighting to each limitation they measure; in reality, different limitations may affect validity to varying degrees. The Cochrane Collaboration recommends using an approach that defines the quality of a body of evidence as the extent to which a reviewer can be confident that an estimate of effect or association is close to the quantity of specific interest (Schünemann *et al.*, 2011). This is not completely apparent in

previous reviews, so the current review addressed this by offering a quality assessment of included studies using the Grades of Recommendation, Assessment, Development and Evaluation (GRADE) assessment tool (Guyatt *et al.,* 2008). Furthermore, assessment of heterogeneity was completed using the I^2 statistic, as is also recommended by the Cochrane collaboration (Deeks, Higgins and Altman, 2011).

Despite some similarities, the current review differs significantly from existing works, including a wide range of outcomes and analyses that have not been included in previous published work or proposed reviews from PROSPERO. Furthermore, by following recommendations from the *Cochrane Handbook for Systematic Reviews of Interventions*, it set the highest standard in terms of process and statistical analysis.

Therefore, the principal objective of the present review is to identify the optimal exercise, or combination of diet and exercise, prescription to elicit maximum benefit in women with PCOS.

3.2. Aims and Hypothesis

This systematic review aimed to analyse the evidence on the effectiveness of exercise compared to:

i. control or usual care;
ii. diet alone; and
iii. exercise combined with diet;

In addition, it also evaluated the effectiveness of exercise combined with diet compared to:

iv. control or usual care; and
v. diet alone.

3.3. Methods

For complete details of methodology, please refer to Chapter 2, Sections 2.1.1 to 2.1.8.

3.4. Results

3.4.1. Search Results

The database searches returned 2,390 articles for assessment; one additional article was also sent after requesting further information from one of the study authors (Costa *et al.*, 2018). After the removal of duplicate papers, 1,908 articles remained and were screened for eligibility based upon their title and abstract. Following removal of ineligible papers, 87 full-text articles were retrieved for detailed evaluation, and an additional 60 of these were excluded with detailed reason (Figure 3.1 and Appendix 7.3)

This left 27 articles meeting the inclusion criteria (Table 2.1). They are: Almenning, *et al.*, 2015; Brown *et al.*, 2009; Bruner *et al.*, 2006; Costa *et al.*, 2018; Guzick *et al.*, 1994; Hoeger *et al.*, 2004; Jedel *et al.*, 2011; Konopka *et al.*, 2015; Leonhardt *et al.*, 2015; Nasrekani *et al.*, 2016; Nybacka *et al.*, 2011; Nybacka *et al.*, 2013; Petrányi and Zaoura-Petrányi, 2011; Sá *et al.*, 2015; Roessler *et al.*, 2013; Saremi *et al.*, 2013; Saremi and Yaghoubi, 2016; Stener-Victorin *et al.*, 2009; Stener-Victorin *et al*., 2012; Stener-Victorin *et al.*, 2013; Thomson *et al.*, 2008; Thomson *et al.*, 2010; Thomson *et al.*, 2012; Thomson *et al.*, 2016; Turan *et al.*, 2015; Vigorito *et al.*, 2007; and Vizza *et al.*, 2016. However, these publications were based on the findings of only 18 trials, because four trials had multiple publications; these are: Stener-Victorin *et al.* (2009) which had four additional papers (Jedel *et al.*, 2011; Leonhardt *et al.*, 2015; Stener-Victorin *et* al., 2012; Stener-Victorin *et al.*, 2013); Thomson *et al.* (2008) that had three additional publications (Thomson *et al.*, 2010; Thomson *et al.*, 2012; Thomson *et al.*, 2016); Nybacka *et al.* (2011) who had one additional publication (Nybacka *et al.*, 2013); and Sá *et al.*, (2015) who also had one additional publication (Costa *et al.*, 2018).

Although included within the qualitative synthesis, one study was excluded from the meta-analysis (Brown *et al.*, 2009) because data were reported as median and range values (attempts to contact the author were unsuccessful).

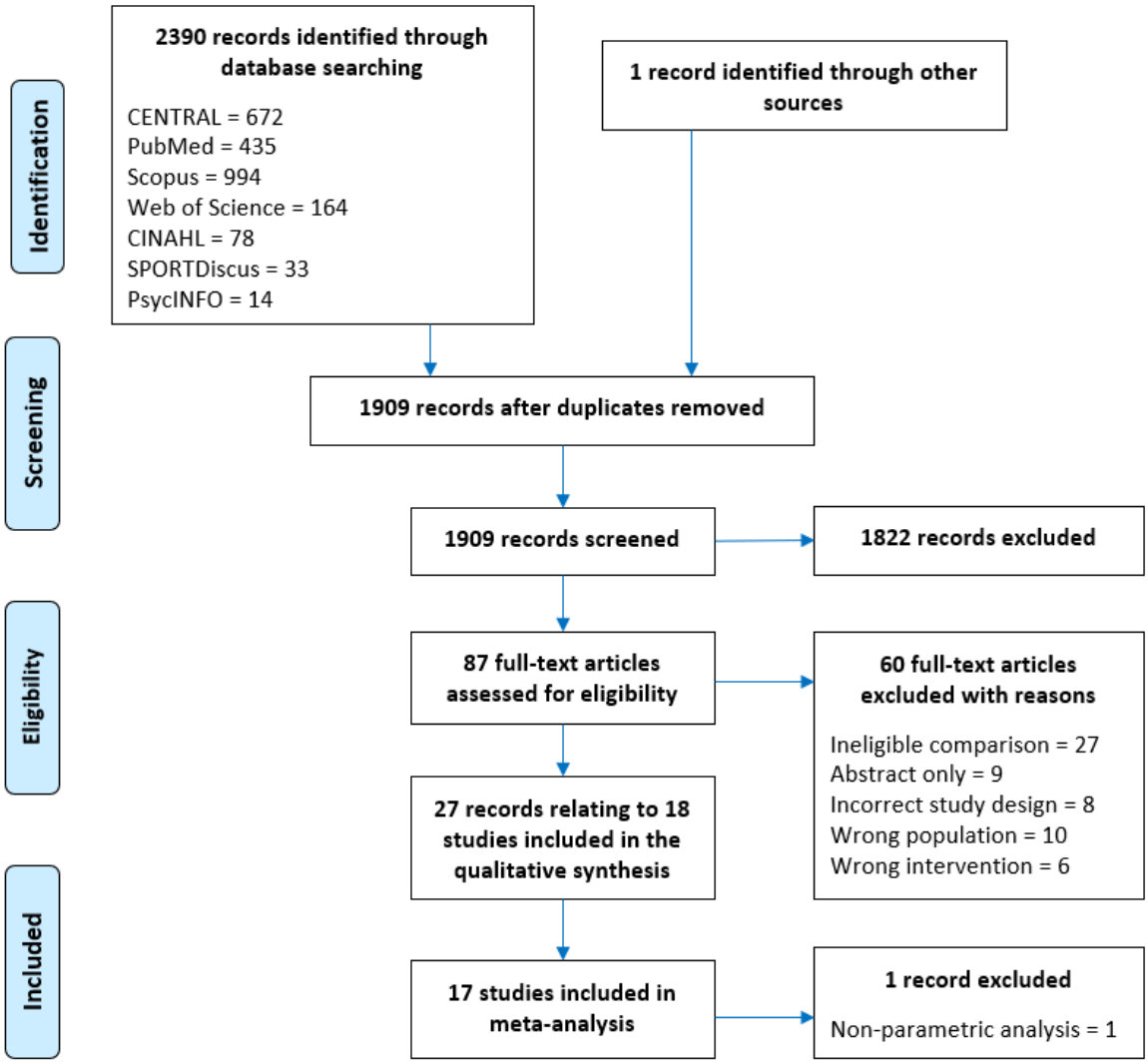


Figure 3.1. Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flow diagram.

3.4.2 Design and attrition of eligible studies

Of the 18 included studies, 16 of them were classified as randomised controlled clinical trials (RCTs), one trial was a quasi-RCT (Petrányi and Zaoura-Petrányi, 2011) and the final study employed a randomised crossover design (Roessler *et al.*, 2013). Study characteristics are presented in Table 3.1.

Twelve trials made the comparison of exercise with usual care, minimal intervention, or control (Almenning, *et al.*, 2015; Brown *et al.*, 2009; Konopka *et al.*, 2015; Nasrekani *et al.*, 2016; Roessler *et al.*, 2013; Sá *et al.*, 2015; Saremi *et al.*, 2013; Saremi and Yaghoubi, 2016; Stener-Victorin *et al.*, 2009; Turan *et al.*, 2015; Vigorito *et al.*, 2007; Vizza *et al.*, 2016). Three trials compared a combined

exercise and diet intervention to diet alone (Thomson *et al.,* 2008; Bruner *et al.,* 2006; Nybacka *et al.,* 2011) and a further three, exercise and diet combined with usual care/minimal intervention/control (Guzick *et al.,* 1994; Hoeger *et al.,* 2004; Petrányi and Zaoura-Petrányi, 2011). Only one trial made the comparison of exercise versus diet, and also of exercise versus exercise and diet combined (Nybacka *et al.,* 2011). The total number of participants included within all trials combined were 758 (exercise/intervention, *n* = 230; control, n = 257; combined treatment arms, *n* = 174; and diet alone, *n* = 54). In addition, a total of 43 additional participants were present in ineligible arms (*i.e.* pharmacological or low-frequency electroacupuncture) from four of the included studies (Hoeger *et al.,* 2004; Petrányi and Zaoura-Petrányi, 2011; Saremi and Yaghoubi, 2016; Stener-Victorin *et al.,* 2009).

Where reported, study attrition ranged from 6% (Turan *et al.,* 2015) up to 50% (Thomson *et al.,* 2008) with a median value of 19.5%; five trials (28%) reported attrition over 20% (Almenning, *et al.,* 2015; Brown *et al.,* 2009; Hoeger *et al.,* 2004; Nybacka *et al.,* 2011; Thomson *et al.,* 2008). Reasons for exercise dropouts included non-exercise related injury (Almenning, *et al.,* 2015; Brown *et al.,* 2009; Roessler *et al.,* 2013; Thomson *et al.,* 2008; Vizza *et al.,* 2016), pregnancy (Almenning, *et al.,* 2015; Hoeger *et al.,* 2004; Roessler *et al.,* 2013; Thomson *et al.,* 2008; Vizza *et al.,* 2016), time (Brown *et al.,* 2009; Roessler *et al.,* 2013; Thomson *et al.,* 2008), work/family commitments (Almenning, *et al.,* 2015; Thomson *et al.,* 2008; Vizza *et al.,* 2016), personal reasons (Nybacka *et al.,* 2011; Sá *et al.,* 2015; Stener-Victorin *et al.,* 2009; Thomson *et al.,* 2008), medical grounds (Brown *et al.,* 2009; Nybacka *et al.,* 2011; Stener-Victorin *et al.,* 2009), and relocation (Thomson *et al.,* 2008). Two trials excluded participants because adherence to intervention was <75% (Turan *et al.,* 2015) or due to a failure to comply with study requirements (Thomson *et al.,* 2008). Eight trials (44%) did not report any attrition (Bruner *et al.,* 2006; Guzick *et al.,* 1994; Konopka *et al.,* 2015; Nasrekani *et al.,* 2016; Petrányi and Zaoura-Petrányi, 2011; Saremi *et al.,* 2013; Saremi and Yaghoubi, 2016; Vigorito *et al.,* 2007).

Table 3.1. Characteristics of studies included in this systematic review.

Study (design)	*N* randomised/ analysed	Intervention Duration (assessment points)	Participant Characteristics (PCOS diagnostic criteria)	Intervention	Outcome measures
Almenning *et al*., 2015 (RCT)	HIIT: 10/8 RT: 11/8 CON: 10/9	10 wks (baseline, 10 wks)	age: 27.2±5.5 y BMI: 26.7±6.0 kg/m^2 (Rotterdam)	HIIT frequency: 3 times/wk HIIT intensity: 2 d/wk, 4 x 4 mins 90-95% HR_{max} /3 x 3 mins ~70% HR_{max}. 1 d/wk, 10 x 1 min 'all-out'/10 x 1 min rest. RT frequency: 3 times/wk RT sets x reps: 3 x 10 RT load: 75% 1-RM	HOMA-IR, FBG, FI, TG, TC, LDL-C, HDL-C VO_2 max, RHR, BW, BMI, WC, BF%, FM, FFM, T, SHBG, FAI, hsCRP
Brown *et al*., 2009 (RCT)	EX: 21/8 CON: 16/12	20-24 wks due to varying length of ramp up phase (baseline, immediately post)	age: 32.3 ± ns y BMI: 33.0 kg/m^2 (NIH)	Exercise: 12 wk moderate-intensity intervention preceded by 8-12 wk ramp-up. Aerobic duration: ~228 mins/wk (≤ 60 bouts) Aerobic intensity: 40-60% VO_2 max	FBG, FI, HOMA-IR, TG, LDL-C, HDL-C, VO_2 max, BW, BMI, WC, FT, SBP, DBP
Bruner *et al*., 2006 (RCT)	EX + DIET: 7/7 DIET: 5/5	12 wks (baseline, 12 weeks)	age: 30.7±4.6 y BMI: 36.6±6.0 kg/m^2 (Rotterdam)	Exercise frequency: 3 times/wk Aerobic intensity: 70-85% HR_{max} Aerobic duration: 30 mins (+10-min warm-up) RT sets x reps: 2-3 x 10-15 RT load: not specified Diet: 1 hour/wk of nutritional counselling	FI, QUICKI, VO_2 max, BW, BMI, WC, T, SHBG, FAI
Guzick *et al*., 1994 (RCT)	EX + DIET: 6/6 CON: 6/6	12 wks (baseline, 12 weeks)	age: 31.7±10.0 y BMI: ns (NIH)	Exercise frequency: 5 times/wk Exercise intensity: 1050-4200 kJ/wk Diet: VLCD (8 wks) with calories increased over final 4 wks (4200-5040 kJ/d). 'Optifast' used to supplement diet.	FBG, FI, BW, WHR, T, SHBG, FT, LH, FSH.

Study (design)	*N* randomised/ analysed	Intervention Duration (assessment points)	Participant Characteristics (PCOS diagnostic criteria)	Intervention	Outcome measures
Hoeger *et al.*, 2004 (RCT)	LS + PLA: 11/6 PLA: 9/7 LS + MF : 9/5 MF: 9/5	48 wks (baseline, 24 wks, 48 wks)	Age: 28.5±5.2 y BMI: 39.0±6.1 kg/m^2 (NIH)	Exercise programme: Individualised to achieve 150 minutes per wk Diet: Individualised healthy balanced meal plan to achieve 500-1000 kcal deficit per day Metformin: 850 mg 2 times/day	BW, T, SHBG, FAI
Konopka *et al.*, 2015 (RCT)	EX: 12/12 CON: 13/13	12 wks (baseline, 12 wks)	Age: 35±5.0 y BMI: 33.0±5.0 kg/m^2 (Rotterdam)	Exercise frequency: 5 times/wk Exercise intensity: 65% VO_2 peak Exercise duration: 60 mins	FBG, FI, HOMA-IR, BMI, BW, FM, FFM, E_2
Nasrekani *et al.*, 2016 (RCT)	EX: 10/10 CON: 10/10	12 wks (baseline, 12 wks)	Age: 30.4±5.9 y BMI: 28.3±6.2 kg/m^2 (Rotterdam)	Exercise frequency: 3 times/wk Exercise intensity: 40-65% HR_{max} Exercise duration: 25-30 mins	VO_2 max, BW, BMI, FSH, LH
Nybacka *et al.* 2011 (RCT)	EX: 19/17 EX + DIET: 19/12 DIET: 19/14	4 months (baseline, 4 months)	Age: 30.8±5.2 y BMI: 36.0±6.2 kg/m^2 (Rotterdam)	Exercise programme: Individualised to meet individuals' capacity, goals and interest Diet: ≥ 600 kcal/day reduction maintaining 55-60% CHO, 25-30% fat and 10-15% protein	FBG, FI, HOMA-IR, BW, BMI, WHR, BF%, FFM, T, SHBG, FT, E_2, FSH, LH
Petrányi *et al.*, 2011 (QRCT)	LS+MF: 29/29 MF: 27/27	6 months (baseline, 6 months)	Age: 29 ± ns y BMI: 27.2±6.9 kg/m^2 (Rotterdam)	Exercise programme: recommendation to increase physical activity levels (specifics unclear) Diet: low glycaemic index diet with caloric restriction for those who are obese. Metformin: 500 mg 3 times/day	BMI, WHR

Study (design)	*N* randomised/ analysed	Intervention Duration (assessment points)	Participant Characteristics (PCOS diagnostic criteria)	Intervention	Outcome measures
Roessler *et al.*, 2013 (Randomised crossover)	EX: 8/7 CON: 9/7	16 wks (baseline, 8 wks, 16 wks)	Age: 31.7±7.9 y BMI: 36.3±7.2 kg/m^2 (Rotterdam)	Exercise frequency: 3 times/wk (2 x cycle, 1 x walk) Exercise intensity: following 2-week ramp, cycling 20-180 secs 80-100% HR_{max} / rest 25-180 secs 45-65% HR_{max}. Walking 3-5 mins 80-90% HR_{max} / 1 min 50-60% HR_{max}. Exercise duration: 45 mins (+10 min warm-up) Control: Group counselling sessions (2 hours, 1 time/wk) focussing on barriers and motivation	VO_2 max, BW, BMI, WC
Sa *et al.*, 2015 (RCT)	EX: 15/14 CON: 15/13	16 wks (baseline, 16 wks)	Age: 26.0±5.0 y BMI: 32.8±4.6 kg/m^2 (Rotterdam)	Exercise frequency: 3 times/wk Exercise intensity: 60-85% HR_{max} Exercise duration: 40 mins (+5 mins)	SBP, DBP, FI, BMI, RHR, VO_2 max, T, FSH, LH
Saremi *et al.*, 2013 (RCT)	EX: 11/11 CON: 11/11	8 wks (baseline, 8wks)	Age: 35.2±4.4 y BMI: 28.3±4.3 kg/m^2 (Rotterdam)	Exercise frequency: 3 times/wk Exercise intensity: 40-65% HR_{max} Exercise duration: 30 mins	FBG, FI, HOMA-IR, TG, TC, LDL-C, HDL-C, VO_2 peak, BW, BMI, BF%, WC, WHR
Saremi *et al.*, 2016 (RCT)	EX + PLA: 10/10 CON : 10/10 EX + CAL: 10/10	8 wks (baseline, 8 wks)	Age: 27.1±5.1 y BMI: 25.5±2.7 kg/m^2 (Rotterdam)	Exercise frequency: 3 times/wk RT sets x reps: 1-2 x 15-20 RT load: 40-60% 1-RM	FBG, FI, HOMA-IR, TG, TC, LDL-C, HDL-C, BW, BMI
Stener-Victorin *et al.*, 2009 (RCT)	EX: 34/22 CON: 17/13 ACU: 33/24	16 wks (baseline, 16 wks, 32 wks)	Age: 30±4.4 y BMI: 28.1±7.3 kg/m^2 (Rotterdam)	Exercise frequency: 3 times/wk Exercise intensity: HR ≥ 120 BPM Exercise duration: 30-45 mins Low-frequency electroacupuncture: 14 x 30 min treatments over 16 wks	SBP, DBP, FBG, FI, HOMA-IR, TG, TC, LDL-C, HDL-C, BMI, WHR, T, FT, SHBG, FAI, LH, FSH, VO_2 max, BMI, E_2

Study (design)	*N* randomised/ analysed	Intervention Duration (assessment points)	Participant Characteristics (PCOS diagnostic criteria)	Intervention	Outcome measures
Thomson *et al.*, 2008 (RCT)	AET + DIET: 31/18 AET + RT + DIET: 33/20 DIET: 30/14	20 wks (baseline, 10 wks, 20 wks)	Age: 29.3±6.8 y BMI: 36.1±4.8 kg/m^2 (Rotterdam)	Exercise frequency: 5 times/wk (3 x aerobic, 2 x RT in combined exercise group) Aerobic intensity: 60-65% HR_{max} progressed to 75-80% HR_{max} by study end Aerobic duration: 25-30 mins progressed to 45 mins by study end RT sets x reps: 3 x 12 RT load: 50-60% 1-RM progressed to 65-75% 1-RM after 2 weeks Diet: energy restricted high protein diet (5000-6000 kJ/day)	SBP, DBP, FBG, FI, HOMA-IR, TG, TC, LDL-C, HDL-C, BW, BF%, FM, FFM, WC, T, SHBG, FAI, PCOS-Q
Turan *et al.*, 2015 (RCT)	EX: 16/14 CON: 16/16	8 wks (baseline, 8 wks)	Age: 24.5 ± 2.8 y BMI: 21.9±3.5 kg/m^2 (Rotterdam)	Exercise frequency: 3 times/wk Exercise duration: 50-60 mins Aerobic intensity: 65-70% HR_{max} RT sets x reps: 1 x 15 RT load: 5-6 on RPE for RT scale	SBP, DBP, FBG, HOMA-IR, FI, TG, TC, HDL-C, LDL-C, BMI, WC, RHR, VO_2 max, T, FT, E_2, LH, FSH
Vigorito *et al.*, 2007 (RCT)	EX: 45/45 CON: 45/45	3 months (baseline, 3 months)	Age: 21.8±2.1 y BMI: 29.4±3.2 kg/m^2 (Rotterdam)	Exercise frequency: 3 times/wk Exercise intensity: 60-70% VO_2 max Exercise duration: 30 mins	SBP, DBP, FBG, FI, TG, TC, LDL-C, HDL-C, VO_2 max, RHR, BMI, WC, E_2, T, FT, SHBG, FAI, LH, FSH, CRP
Vizza *et al.*, 2016 (RCT)	EX: 8/7 CON: 7/6	12 wks (baseline, 12 wks)	Age: 27±5.0 y BMI: 37.8±11.4 kg/m^2	Exercise frequency: 4 times/wk (2 x RT, 2 home-based) RT sets x reps: 2-3 x 8-12 RT load: Progressed with strength gains Home-based: Callisthenics, 3 sets of 10 reps	FBG, FI, HOMA-IR, BW, BMI, WC, FM, FFM, BF%, hsCRP, T, SHBG, FAI, PCOS-Q, SF-36

Key: Design; RCT: randomised controlled trial, QRCT: quasi-randomised controlled trial. *N* randomised: the number of participants randomised into each study arm at the study initiation; analysed is the number of participants included within the analysis; HIIT: high-intensity interval training; RT: resistance training; CON: control group; EX: exercise group; DIET: dietary intervention; LS: lifestyle; PLA: placebo; MF: metformin; ACU: acupuncture; AET: aerobic exercise training; CAL: calcium supplementation. Intervention duration: length of the duration; assessment points: the time-points at which researchers have assessed outcome measures. Participant characteristics presented as mean ± standard deviation (SD) or median in one study [96] for age (in years; y) and BMI (kg/m^2) at study entry; ns: not specified. Diagnostic criteria: the specific criteria used to confirm a PCOS diagnosis; NIH: National Institute of Health (1990) diagnostic criteria; Rotterdam: European Society for Human Reproductive and Embryology/American Society for Reproductive Medicine (2003) diagnostic criteria. Outcome measures refers to the outcomes from each study that are relevant to this systematic review. VO_2 max: maximum oxygen uptake; RHR: resting heart rate; HDL-C: high density lipoprotein cholesterol; LDL-C: low density lipoprotein cholesterol; TC: total cholesterol; TG: triglycerides; FBG: fasting blood glucose; FI: fasting insulin; HOMA-IR: homeostatic assessment of insulin resistance; QUICKI: quantitative insulin sensitivity check index; FM: fat mass; FFM: fat free mass; BF%: body fat percentage; BW: body weight; BMI: body mass index; WC: waist circumference; WHR: waist to hip ratio; SHBG: sex hormone binding globulin; FAI: free androgen index; T: testosterone; FT: free testosterone; E_2: oestradiol; LH: luteinising hormone; FSH: follicle stimulating hormone; SBP: systolic blood pressure; DBP: diastolic blood pressure; hsCRP: high-sensitivity C-reactive protein; d: day; mins: minutes; wk: week; reps: repetitions; RM: maximum number of repetitions; HRmax: maximum heart rate; PCOSQ: PCOS health-related questionnaire; SF-36: Optum36-item Short Form Survey; VLCD: very low calorie diet; CHO: carbohydrate. Studies presented by lead author and year of publication.

3.4.3. Participant characteristics of included studies

Participant characteristics are presented in Table 3.1. Included trials used a range of criteria to diagnose PCOS (Table 1.1). Only three trials (Brown *et al.*, 2009; Guzick *et al.*, 1994; Hoeger *et al.*, 2004) used the NIH diagnostic criteria (Zawadski and Dunaif, 1992), whereas 14 studies (Almenning, *et al.*, 2015; Bruner *et al.*, 2006; Konopka *et al.*, 2015; Nasrekani *et al.*, 2016; Nybacka *et al.*, 2011; Petrányi and Zaoura-Petrányi, 2011; Roessler *et al.*, 2013; Sá *et al.*, 2015; Saremi *et al.*, 2013; Saremi and Yaghoubi, 2016; Stener-Victorin *et al.*, 2009; Thomson *et al.*, 2008; Turan *et al.*, 2015; Vigorito *et al.*, 2007) utilised the Rotterdam consensus criteria (ESHRE/ASRM, 2004). One trial confirmed the PCOS diagnosis via participants' general practitioner/specialist (Vizza *et al.*, 2016), but the criteria used were not explicitly stated. None of these trials specified use of the AE-PCOS definition (Azziz *et al.*, 2008).

Many of the included studies stated clear inclusion/exclusion eligibility criteria for participation. Participants with T2DM, fasting hyperglycaemia, or glucose intolerance were explicitly excluded in nine trials (Brown *et al.*, 2009; Bruner *et al.*, 2006; Konopka *et al.*, 2015; Roessler *et al.*, 2013; Saremi *et al.*, 2013; Stener-Victorin *et al.*, 2009; Thomson *et al.*, 2008; Turan *et al.*, 2015; Vigorito *et al.*, 2007), and nine trials additionally excluded participants with any diagnosed CVD (Bruner *et al.*, 2006; Guzick *et al.*, 1994; Roessler *et al.*, 2013; Saremi *et al.*, 2013; Stener-Victorin *et al.*, 2009; Thomson *et al.*, 2008; Turan *et al.*, 2015; Vigorito *et al.*, 2007; Vizza *et al.*, 2016). Another prerequisite in seven trials (39%) was the activity status of participants upon enrolment; participants had to have a sedentary lifestyle without any recent involvement in any other exercise intervention (Almenning, *et al.*, 2015; Brown *et al.*, 2009; Konopka *et al.*, 2015; Nasrekani *et al.*, 2016; Saremi *et al.*, 2013; Thomson *et al.*, 2008; Vizza *et al.*, 2016).

3.4.4 Intervention and comparison details

Fourteen trials (74%) assessed the effectiveness of an exercise only intervention and six trials (32%) assessed the effectiveness of combined exercise and dietary interventions. Of the included trials, 14 (74%) included intervention arms which consisted of aerobic exercise only, whereas a further three

(16%) combined aerobic exercise with resistance training (Bruner *et al.*, 2006; Thomson *et al.*, 2008; Turan *et al.*, 2015). Of those studies which incorporated some form of aerobic exercise (*n* = 17, 94%), 11 trials (61%) specified either walking, brisk walking or jogging (Almenning *et al.*, 2015; Brown *et al.*, 2009; Bruner *et al.*, 2006; Guzick *et al.*, 1994; Nasrekani *et al.*, 2016; Nybacka *et al.*, 2011; Roessler *et al.*, 2013; Sá *et al.*, 2015; Saremi *et al.*, 2013; Stener-Victorin *et al.*, 2009; Thomson *et al.*, 2008); and seven (39%) incorporated static cycling either on its own or as part of a wider intervention (Almenning *et al.*, 2015; Brown *et al.*, 2009; Bruner *et al.*, 2006; Konopka *et al.*, 2015; Roessler *et al.*, 2013; Stener-Victorin *et al.*, 2009; Vigorito *et al.*, 2007). There were also three single trials which each incorporated either elliptical training (Brown *et al.*, 2009), step training (Turan *et al.*, 2015), or swimming (Nybacka *et al.*, 2011). Five trials (28%) specified that they had allowed participants to self-select training modality from those listed above (Almenning *et al.*, 2015; Brown *et al.*, 2009; Bruner *et al.*, 2006; Nybacka *et al.*, 2011; Stener-Victorin *et al.*, 2009), whereas two trials (Hoeger *et al.*, 2004; Petrányi and Zaoura-Petrányi, 2011) state that they allowed participants to self-select a modality but do not list the exercise choices. Three trials (16%) had intervention arms that were resistance training only (Almenning *et al.*, 2015; Saremi and Yaghoubi, 2016; Vizza *et al.*, 2016). However, one of the trials did not specify the type of exercise undertaken by participants (Petrányi and Zaoura-Petrányi, 2011).

In 10 trials (56%) the modal training session frequency was three times per week (Almenning *et al.*, 2015; Bruner *et al.*, 2006; Nasrekani *et al.*, 2016; Sá *et al.*, 2015; Saremi *et al.*, 2013; Saremi and Yaghoubi, 2016; Stener-Victorin *et al.*, 2009; Thomson *et al.*, 2008; Turan *et al.*, 2015; Vigorito *et al.*, 2007). In contrast, five sessions per week were prescribed in three trials (17%) (Guzick *et al.*, 1994; Konopka *et al.*, 2015; Thomson *et al.*, 2008). In a single trial (Vizza *et al.*, 2016) four training sessions per week were set. Of the remaining four trials, one trial specified a weekly physical activity (PA) time target of 150 minutes per week (Hoeger *et al.*, 2004), another set an exercise dose of 14 kcal/kg/week (Brown *et al.*, 2009), and the final two did not specify any information on the training frequency or volume (Nybacka *et al.*, 2011; Petrányi and Zaoura-Petrányi, 2011).

Eight included trials (44%) set the aerobic exercise intensity based upon a percentage of the participants' maximum heart rate (HR_{max}; Almenning *et al.,* 2015; Bruner *et al.,* 2006; Nasrekani *et al.,* 2016; Roessler *et al.,* 2013; Sá *et al.,* 2015; Saremi *et al.,* 2013; Thomson *et al.,* 2008; Turan *et al.,* 2015), whereas three trials set intensity based upon a percentage of maximal oxygen uptake (VO_2 max; Brown *et al.,* 2009; Konopka *et al.,* 2015; Vigorito *et al.,* 2007). One trial specified that heart rate (HR) during exercise was set at ≥120 beats/min (Stener-Victorin *et al.,* 2009). Three of the trials that used resistance training prescribed the intensity of each exercise based upon a percentage of a pre-determined one-repetition maximum [either 40-60% (Saremi and Yaghoubi, 2016) or 50-75% (Almenning *et al.,* 2015; Thomson *et al.,* 2008)]. One resistance training intervention set the exercise intensity using a rate of perceived exertion (RPE) of 5-6 out of 10 (Turan *et al.,* 2015). Four trials increased the exercise intensity as the intervention progressed (Nasrekani *et al.,* 2016; Saremi *et al.,* 2013; Saremi and Yaghoubi, 2016; Thomson *et al.,* 2008) and a further six trials did not specify any intensity information for their interventions (Bruner *et al.,* 2006; Guzick *et al.,* 1994; Hoeger *et al.,* 2004; Nybacka *et al.,* 2011; Petrányi and Zaoura-Petrányi, 2011; Vizza *et al.,* 2016).

Eleven trials (61%) prescribed session durations lasting for one hour or less; the individual session durations were ≤30 minutes (Almenning *et al.,* 2015; Nasrekani *et al.,* 2016; Roessler *et al.,* 2013; Saremi *et al.,* 2013; Vigorito *et al.,* 2007), >30-60 minutes (Brown *et al.,* 2009; Konopka *et al.,* 2015; Sá *et al.,* 2015; Stener-Victorin *et al.,* 2009; Vizza *et al.,* 2016), or 20-30 to 45 minutes (Thomson *et al.,* 2008). Only one trial incorporated training sessions of >60 minutes (Bruner *et al.,* 2006). Hoeger and colleagues (2004) specified 150 minutes as a weekly target, whereas another trial used a target distance of 10 miles per week (Guzick *et al.,* 1994). Four trials did not specify any timings, or equivalent information, for their individual session duration (Nybacka *et al.,* 2011; Petrányi and Zaoura-Petrányi, 2011; Saremi and Yaghoubi, 2016; Vizza *et al.,* 2016).

In ten of the included trials (56%), participants were fully supervised at every exercise session (Bruner *et al.,* 2006; Konopka *et al.,* 2015; Nasrekani *et al.,* 2016; Nybacka *et al.,* 2011; Roessler *et al.,* 2013; Sá *et al.,* 2015; Saremi *et al.,* 2013; Saremi and Yaghoubi, 2016; Turan *et al.,* 2015; Vigorito *et al.,* 2007), whereas two trials (11%) utilised a mixed approach with some sessions being

supervised and others classified as home-based with participants exercising independently (Almenning *et al.*, 2015; Vizza *et al.*, 2016). Exercise sessions were completely unsupervised in one (6%) additional trial (Stener-Victorin *et al.*, 2009) with support provided by weekly telephone calls. The remaining five trials (26%) did not report supervision arrangements (Brown *et al.*, 2009; Guzick *et al.*, 1994; Hoeger *et al.*, 2004; Petrányi and Zaoura-Petrányi, 2011; Thomson *et al.*, 2008).

Of the six trials (33%) that incorporated a dietary component, five specified either a daily caloric target (Guzick *et al.*, 1994; Thomson *et al.*, 2008), a reduced caloric intake (Nybacka *et al.*, 2011; Petrányi and Zaoura-Petrányi, 2011), or an individualised caloric deficit (Hoeger *et al.*, 2004). One of these six trials (Bruner *et al.*, 2006) incorporated weekly nutritional counselling sessions in order to educate participants on a range of nutritional topics.

Thirteen of the included trials (72%) had a control arm (Table 3.2; Almenning *et al.*, 2015; Brown *et al.*, 2009; Guzick *et al.*, 1994; Konopka *et al.*, 2015; Nasrekani *et al.*, 2016; Roessler *et al.*, 2013; Sá *et al.*, 2015; Saremi *et al.*, 2013; Saremi and Yaghoubi, 2016; Stener-Victorin *et al.*, 2009; Turan *et al.*, 2015; Vigorito *et al.*, 2007; Vizza *et al.*, 2016). Three of these 13 trials used a wait-list control and either offered participants the same intervention (Roessler *et al.*, 2013; Guzick *et al.*, 1994) or a one-month gym membership (Almenning *et al.*, 2015) upon trial completion. Three of the remaining trials (17%) had a diet only intervention arm as their comparison group (Bruner *et al.*, 2006; Nybacka *et al.*, 2011; Thomson *et al.*, 2008), one trial used a placebo (Hoeger *et al.*, 2004), and another trial used metformin treatment only (Petrányi and Zaoura-Petrányi, 2011).

3.4.5. Characteristics of the outcome measures

All studies incorporated participant assessment at baseline and immediately following completion of the intervention; two other trials incorporated an additional midway assessment (Hoeger *et al.*, 2004; Thomson *et al.*, 2008), one trial added a follow-up assessment 16 weeks post-intervention (Stener-Victorin *et al.*, 2009), and another trial assessed at baseline, crossover and immediately post-

intervention (Roessler *et al.*, 2013). No post-intervention follow-up analysis was possible due to lack of studies.

Seven trials (39%) specified the methods which were used to calculate sample size (Almenning *et al.*, 2015; Brown *et al.*, 2009; Sá *et al.*, 2015; Saremi *et al.*, 2013; Saremi and Yaghoubi, 2016; Stener-Victorin *et al.*, 2009; Vigorito *et al.*, 2007), although only five (28%) of those reported the outcome upon which their calculations were based. The primary outcomes that were used in the sample size calculations were: HOMA-IR (Almenning *et al.*, 2015), VO_2 peak (Sá *et al.*, 2015), total testosterone (Stener-Victorin *et al.*, 2009), insulin sensitivity (Brown *et al.*, 2009), and BMI (Nybacka *et al.*, 2011). Only three of the trials stated specific recruitment targets (Almenning *et al.*, 2015; Sá *et al.*, 2015; Stener-Victorin *et al.*, 2009) and all of those trials achieved their calculated sample size target. The additional outcomes included in each trial are provided in Table 3.1.

3.4.6. Assessment of risk of bias in included studies

The assessment of risk of bias in included studies is presented in the Risk of Bias Graph (Figure 3.2), and detailed study information is provided in Figure 3.3 and Appendix 7.1.

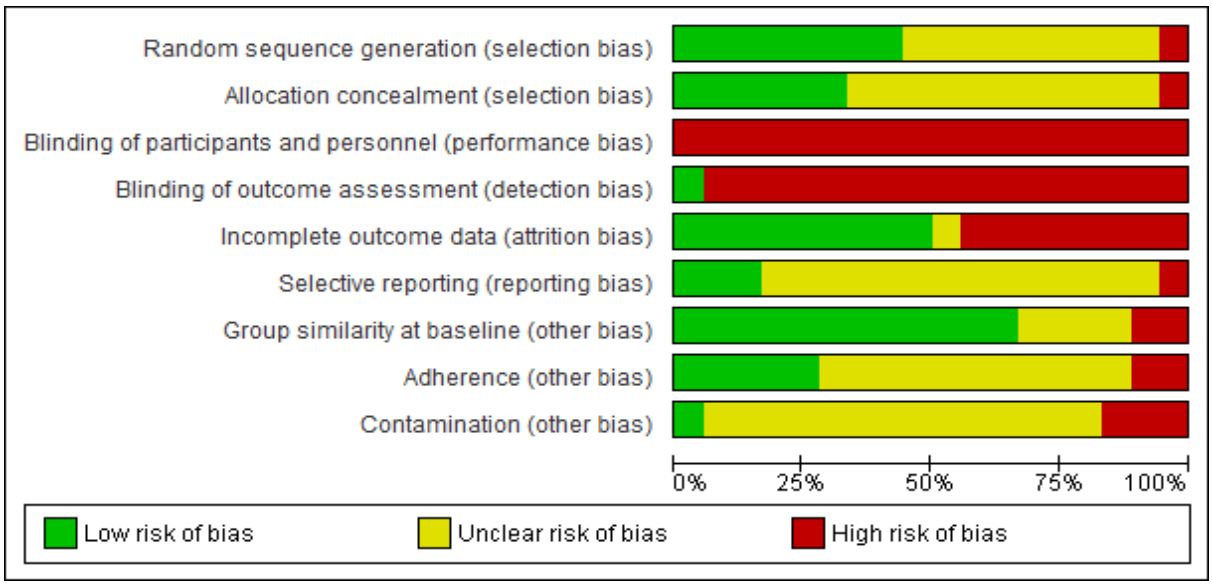


Figure 3.2. Risk of bias judgement for each methodological quality item from the Cochrane Risk of Bias tool, presented as a percentage across all 18 included studies.

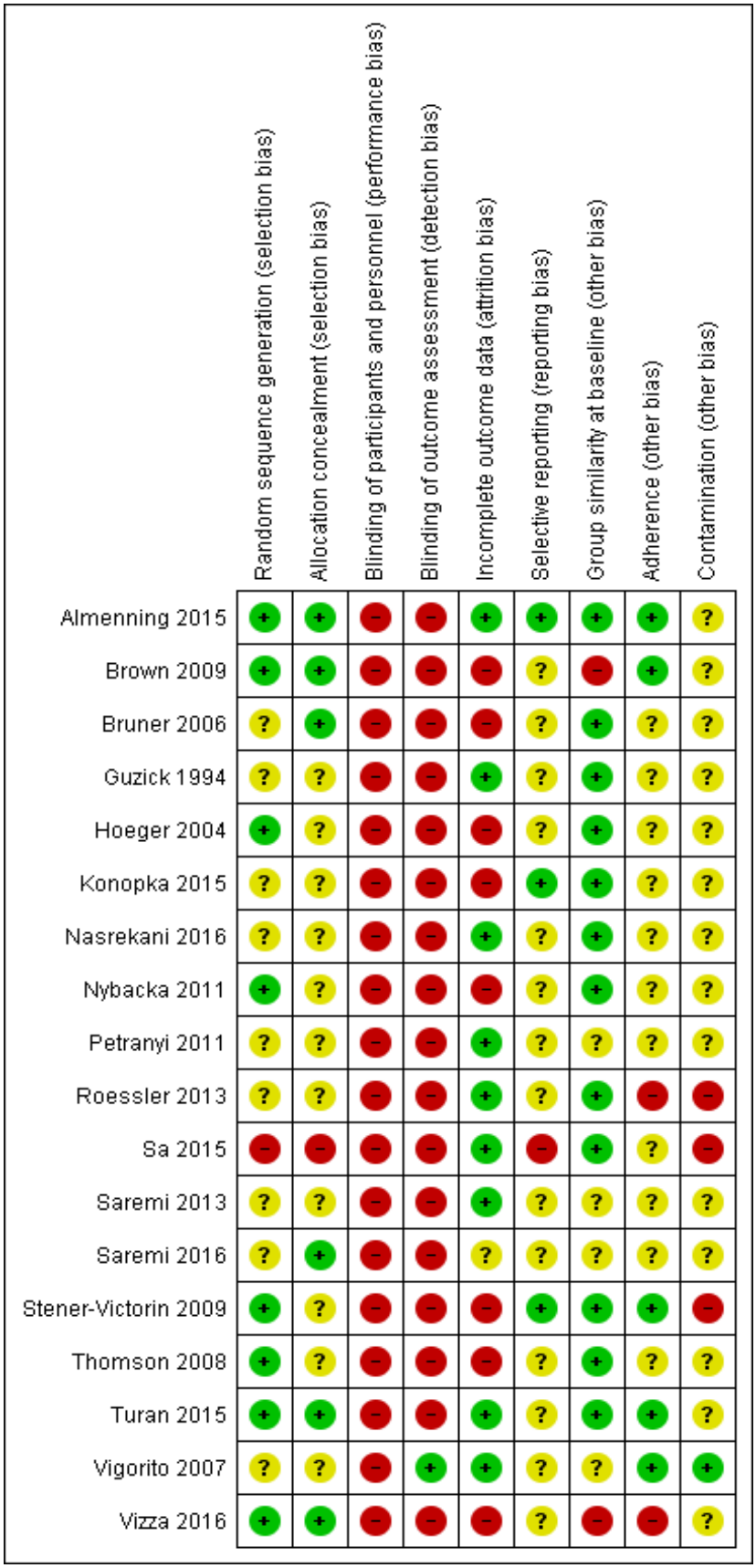


Figure 3.3. Risk of bias items for each included study.

Four of the included trials (22%) were judged to have a low risk of selection bias; that is, they were deemed to have used appropriate methods to generate their sequence for participant randomisation and group allocation concealment (Almenning *et al.*, 2015; Brown *et al.*, 2009; Turan *et al.*, 2015; Vizza *et al.*, 2016). Only one trial was judged to have a high risk of selection bias (Sá *et al.*, 2015) because five participants were allocated to the control group based upon their geographical location. The remaining trials were deemed to have an unclear risk of selection bias due to insufficient reporting of sequence generation or allocation concealment methods. As discussed previously (Section 2.1.3), all trials were judged to be at a high risk of performance bias because of the way exercise interventions are typically structured. Only one trial had a low risk for detection bias (Vigorito *et al.*, 2007) with all remaining trials judged to be at a high risk as it was not stated whether outcome assessors were blinded to participant allocation. One trial (Almenning *et al.*, 2015) used an independent, and blinded, assessor for evaluation of only one outcome (flow mediated dilation).

Eight trials (44%) were judged to be high risk for attrition bias. This was because participant withdrawal rates were >20% (Brown *et al.*, 2009; Hoeger *et al.*, 2004; Nybacka *et al.*, 2011; Stener-Victorin *et al.*, 2009; Thomson *et al.*, 2008), incomplete data due to lab error (Bruner *et al.*, 2006), inappropriate handling of missing data (*i.e.*, last observation carried forward; Vizza *et al.*, 2016), or only a subset of participants completing components of assessment (*i.e.*, hyperinsulinaemic-euglycemic clamp testing; Konopka *et al.*, 2015). A prospective protocol document, or evidence of trial registration was only identified for three trials (Almenning *et al.*, 2015; Konopka *et al.*, 2015; Stener-Victorin *et al.*, 2009), thus making it difficult to judge whether all intended outcomes had been fully reported. Fourteen of the remaining trials (78%) were judged to have an unclear risk of reporting bias, and the final trial (Sá *et al.*, 2015) was judged to be high risk due to incomplete reporting of results.

Eleven trials (61%) were judged to have a low risk of bias based upon similarities between comparative groups at baseline (Almenning, *et al.*, 2015; Bruner *et al.*, 2006; Guzick *et al.*, 1994; Hoeger *et al.*, 2004; Konopka *et al.*, 2015; Nybacka *et al.*, 2011; Roessler *et al.*, 2013; Sá *et al.*, 2015; Stener-Victorin *et al.*, 2009; Thomson *et al.*, 2008; Turan *et al.*, 2015). Of the trials deemed to be

high-risk, one (Vizza *et al.*, 2016) had participants in the intervention group with greater adiposity and less favourable body composition compared to the control group. Another trial (Brown *et al.*, 2009) had an older intervention group that was less hyperandrogenic, hirsute and had lower levels of cardiorespiratory fitness, as well as higher mean BMI, plasma lipids, and IR levels than the control study arm.

Adherence was only reported in seven trials (39%; Almenning, *et al.*, 2015; Brown *et al.*, 2009; Roessler *et al.*, 2013; Stener-Victorin *et al.*, 2009; Turan *et al.*, 2015; Vigorito *et al.*, 2007; Vizza *et al.*, 2016); where reported, the median adherence was 90% with a range from 67% (Roessler *et al.*, 2013) to 103% (Stener-Victorin *et al.*, 2009); participants in the latter study completed more of the intervention than was prescribed. Two included trials (11%) reported intervention adherence below the high-risk threshold of <75% that was outlined in Section 2.1.3 (Roessler *et al.*, 2013; Vizza *et al.*, 2016). Five trials (28%) were judged to have a low risk of adherence bias because intervention adherence was ≥75% (Almenning, *et al.*, 2015; Brown *et al.*, 2009; Stener-Victorin *et al.*, 2009; Turan *et al.*, 2015; Vigorito *et al.*, 2007). Finally, due to an absence of reporting, 14 trials (78%) were judged to have an unclear risk of contamination bias. Only one trial (Vigorito *et al.*, 2007) had a low risk of contamination bias as it was clearly stated that during the intervention period, the control group did not increase PA levels by >4 MET/hr/wk (Waters *et al.*, 2012). In contrast, three trials (17%) had a high risk of contamination bias because comparison groups had either engaged in treatment (Roessler *et al.*, 2013; Stener-Victorin *et al.*, 2009), or control groups had not received their allocated intervention (Sá *et al.*, 2015).

3.5. Effects of interventions

Due to data availability, a meta-analysis was possible only for three comparisons: 1) exercise compared to control; 2) exercise combined with diet compared to a control; and 3) exercise combined with diet compared with diet only.

3.5.1. Exercise versus Control: Primary Outcomes

Eleven trials were included in the exercise versus control meta-analysis (Almenning, *et al.*, 2015; Konopka *et al.*, 2015; Nasrekani *et al.*, 2016; Roessler *et al.*, 2013; Sá *et al.*, 2015; Saremi *et al.*, 2013; Saremi and Yaghoubi, 2016; Stener-Victorin *et al.*, 2009; Turan *et al.*, 2015; Vigorito *et al.*, 2007; Vizza *et al.*, 2016). A summary of key change from baseline findings for the primary outcomes (SBP, DBP, FBG, FI, HOMA-IR, cholesterol, LDL-C, HDL-C and triglycerides), and an overview of the quality of evidence are presented in Table 3.2.

Effect estimates for all included outcomes are presented in Table 3.3. Data are presented for changes from baseline to immediately post-intervention and also for comparisons of values immediately post-intervention. Heterogeneity levels, using the I^2 statistic are also reported.

Table 3.2. Summary of findings for primary outcomes: exercise versus control.

Intervention: exercise; **Comparison**: usual care

Outcomes	**Anticipated absolute effects*** (95% CI)		№ of participants (№ of studies)	Certainty of the evidence (GRADE)	Comments
	Risk with usual care	**Risk with exercise**			
Systolic blood pressure (change from baseline) follow up: range 8 weeks to 16 weeks	The mean systolic blood pressure (change from baseline) ranged from **-2.5 to 1.1 mmHg**	The mean systolic blood pressure (change from baseline) in the intervention group was 2.93 mmHg lower (7.06 lower to 1.2 higher)	158 (4 RCTs)	⨁⨁◯◯ LOW [a,b]	Exercise may result in little to no difference in systolic blood pressure (change from baseline)
Diastolic blood pressure (change from baseline) follow up: range 8 weeks to 16 weeks	The mean diastolic blood pressure (change from baseline) ranged from **-3.1 to 2.9 mmHg**	The mean diastolic blood pressure (change from baseline) in the intervention group was 2.19 mmHg lower (5.23 lower to 0.85 higher)	158 (4 RCTs)	⨁⨁◯◯ LOW [a,b]	Exercise may result in little to no difference in diastolic blood pressure (change from baseline)
Fasting blood glucose (change from baseline) follow up: range 8 weeks to 16 weeks	The mean fasting blood glucose (change from baseline) ranged from **-1.3 to 2.6 mg/dL**	The mean fasting blood glucose (change from baseline) in the intervention group was 1.08 mg/dL lower (2.47 lower to 0.3 higher)	263 (9 RCTs)	⨁⨁◯◯ LOW [c,d]	Exercise may result in little to no difference in fasting blood glucose (change from baseline)
Fasting insulin (change from baseline) follow up: range 8 weeks to 16 weeks	The mean fasting insulin (change from baseline) ranged from **-4.1 to 2.5 µU/ml**	The mean fasting insulin (change from baseline) in the intervention group was 2.44 µU/ml lower (4.42 lower to 0.64 lower)	263 (9 RCTs)	⨁◯◯◯ VERY LOW [e,f,g]	Exercise may reduce fasting insulin (change from baseline), but we are very uncertain
HOMA-IR (change from baseline) follow up: range 8 weeks to 16 weeks	The mean HOMA-IR (change from baseline) ranged from **-0.4 to 0.7**	The mean HOMA-IR (change from baseline) in the intervention group was 0.57 lower (0.99 lower to 0.14 lower)	173 (8 RCTs)	⨁◯◯◯ VERY LOW [d,e,h]	Exercise may reduce HOMA-IR (change from baseline), but we are very uncertain
Total cholesterol (change from baseline) follow up: range 8 weeks to 16 weeks	The mean total cholesterol (change from baseline) ranged from **-8.85 to 6.85 mg/dL**	The mean total cholesterol (change from baseline) in the intervention group was 6.48 mg/dL lower (10.5 lower to 2.45 lower)	225 (7 RCTs)	⨁⨁◯◯ LOW [g,i]	Exercise may reduce total cholesterol (change from baseline) slightly

Intervention: exercise; **Comparison**: usual care

Outcomes	**Anticipated absolute effects*** (95% CI)		№ of participants (№ of studies)	Certainty of the evidence (GRADE)	Comments
	Risk with usual care	**Risk with exercise**			
LDL-C (change from baseline) follow up: range 8 weeks to 16 weeks	The mean LDL-C (change from baseline) ranged from **-17.7 to 7.03 mg/dL**	The mean LDL-C (change from baseline) in the intervention group was 7.51 mg/dL lower (10.01 lower to 5.02 lower)	225 (7 RCTs)	⨁⨁◯◯ LOW [g,i]	Exercise may reduce LDL-C (change from baseline) slightly
HDL-C (change from baseline) follow up: range 8 weeks to 16 weeks	The mean HDL-C (change from baseline) ranged from **-17.7 to 3.5 mg/dL**	The mean HDL-C (change from baseline) in the intervention group was 0.01 mg/dL lower (1.91 lower to 1.89 higher)	225 (7 RCTs)	⨁⨁◯◯ LOW [d,i]	Exercise may result in little to no difference in HDL-C (change from baseline)
Triglycerides (change from baseline) follow up: range 8 weeks to 16 weeks	The mean triglycerides (change from baseline) ranged from **-1.0 to 8.9 mg/dL**	The mean triglycerides (change from baseline) in the intervention group was 4.78 mg/dL lower (7.52 lower to 2.05 lower)	225 (7 RCTs)	⨁⨁◯◯ LOW [g,i]	Exercise likely results in a small effect that may not be an important (or unimportant) reduction in triglycerides (change from baseline)

***The risk in the intervention group** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI). **CI:** Confidence interval; **MD:** Mean difference

GRADE Working Group grades of evidence

- **High certainty:** We are very confident that the true effect lies close to that of the estimate of the effect
- **Moderate certainty:** We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
- **Low certainty:** Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
- **Very low certainty:** We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Explanations

a) Three of the four trials had a high or unclear risk of selection bias, detection bias, and reporting bias; all were at high risk of performance bias; two were at high or unclear risk of attrition bias; and all were at a high or unclear risk of contamination. Therefore, we downgraded by one level.
b) Small number of participants, wide confidence intervals for three of the four trials, and null/negligible effect and appreciable benefit included in the confidence interval for the mean difference. Therefore, we downgraded by one level.
c) Most trials were at an unclear or high risk of selection bias, detection bias, and reporting bias; and all trials were at a high or unclear risk of contamination and low adherence. Therefore, we downgraded by one level.
d) Small number of participants and null/negligible effect and appreciable benefit included in the confidence interval for the mean difference. Therefore, we downgraded by one level.
e) Most trials were at an unclear or high risk of selection bias, detection bias, attrition bias, and reporting bias; and most trials were at a high or unclear risk of contamination and low adherence. Therefore, we downgraded by one level.
f) Considerable heterogeneity was observed. Therefore, we downgraded by one level.
g) Small number of participants and wide confidence intervals in the included trials. Therefore, we downgraded by one level.
h) Considerable heterogeneity was observed and there was minimal or no overlap of confidence intervals. Therefore, we downgraded by one level.
i) Most trials were at an unclear or high risk of selection bias, detection bias, and reporting bias; and all trials were at a high or unclear risk of contamination. Therefore, we downgraded by one level.

Table 3.3. Effect estimates and heterogeneity for change from baseline to post-intervention scores and immediately post-intervention values, for all outcomes analysed in the exercise versus control comparison.

Outcome	Studies	Change from baseline					Immediately post-intervention values				
		N	MD	Lower 95% CI	Upper 95% CI	I^2 (%)	*N*	MD	Lower 95% CI	Upper 95% CI	I^2 (%)
SBP (mmHg)	5,8,10-11	158	-2.93	-7.06	1.20	50	158	2.02	-6.82	10.86	87
DBP (mmHg)	5,8,10-11	158	-2.19	-5.23	0.85	46	158	-0.82	-3.49	1.84	31
FBG (mg/dL)	1,2•,5-8,10-12	263	-1.08	-2.47	0.30	16	238	-1.69	-4.35	0.97	37
FI (µIU/mL)	1,2•,5-8,10-12	263	-2.44**	-4.24	-0.64	91	238	-2.11**	-3.49	-0.73	40
HOMA-IR	1, 2•,5-8,10,12	173	-0.57**	-0.99	-0.14	87	148	-0.22	-0.80	0.36	69
TC (mg/dL)	1,5-8,10-11	225	-5.88**	-9.92	-1.83	35	225	-6.35**	-10.76	-1.95	0
LDL-C (mg/dL)	1,5-8,10-11	225	-7.39***	-9.83	-4.95	0	225	-6.68**	-11.66	-1.70	0
HDL-C (mg/dL)▲	1,5-8,10-11	225	0.29	-1.46	2.04	52	225	1.87	-1.59	5.33	65
TG (mg/dL)	1,5-8,10-11	225	-4.78***	-7.52	-2.05	3	225	-1.97	-7.36	3.42	18
VO_2 max (ml/kg/min)	1,3,5-6,8•,11	229	3.84***	2.87	4.81	17	184	5.01***	3.48	6.54	42
RHR (bpm)	1,8,10-11	156	-2.65	-5.55	0.25	51	156	-3.26***	-4.93	-1.59	0
BMI (kg/m^2)	1,2•,3,4-8,10-12	331	-0.49	-1.04	0.06	66	272	-1.02**	-1.81	-0.23	0
Body Mass (kg)	1,2•,3,4,6-7,12	139	-1.25	-3.27	0.76	33	128	-0.48	-4.86	3.91	0
WC (cm)	1,4-6,10-12	221	-2.62***	-4.13	-1.11	53	221	-2.33	-5.23	0.58	15
WHR	8,11	101	-0.03	-0.08	0.02	0	101	-0.04	-0.08	0.01	19
Body Fat (%)	1,6,12	60	-1.39*	-2.61	-0.18	30	60	-3.28	-7.39	0.83	22
Fat Mass (kg)	1,2•,12	63	-1.70	-3.93	0.53	70	38	5.14	-14.39	24.68	65
FFM (kg)	1,2•,12	63	0.46	-0.89	1.81	58	38	4.99	-7.31	17.28	75
Testosterone (nmol/L)	1,8,10-12	203	-0.09	-0.24	0.06	0	169	-0.08	-0.35	0.19	37
SHBG (nmol/L)	1,8,11-12	173	7.51	-8.01	23.04	89	139	4.03	-18.57	26.63	66
Free T (pg/mL)	8,10	74	-0.43	-1.74	0.88	76	41	0.33	-0.10	0.77	0
FAI	1,8,11-12	139	0.24	-0.55	1.04	0	139	0.68	-1.09	2.44	46
FG	8,11	135	-0.63	-2.08	0.81	0	101	-0.75	-2.03	0.54	0
Oestradiol (pmol/L)	2•,8•,10-11	190	-13.94	-54.53	26.64	65	120	0.27	-11.27	11.80	0
DHEA-S (µmol/L)	1,8	70	-0.60	-1.58	0.39	0	36	-0.20	-1.87	1.46	0
LH (IU/L)	3,8,10-11	185	-0.30	-2.54	1.95	72	151	-0.66	-2.39	1.06	43
FSH (IU/L)	3,8,10-11	185	0.23	-0.08	0.53	0	151	-0.01	-0.40	0.37	0

Outcome	Studies	Change from baseline					Immediately post-intervention values				
		N	MD	Lower 95% CI	Upper 95% CI	I^2 (%)	*N*	MD	Lower 95% CI	Upper 95% CI	I^2 (%)
LH/FSH ratio	8,10	41	-0.02	-0.38	0.33	0	41	0.1	-0.22	0.86	37
PG (nmol/L)	2,11	115	-0.72	-2.53	1.09	74	-	-	-	-	-
Prolactin (ng/mL)	3,11	13	-0.05	-0.71	0.61	0	13	0.20	-0.27	0.68	0
hsCRP (mg/L)	1,12	38	-0.41	-1.19	0.37	0	38	0.67	-1.31	2.65	0
AMH (ng/mL)	1,6-7	67	-0.67	-1.65	0.1	0	67	0.48	-1.89	2.84	0
Adiponectin (µg/mL)	1,8	70	-0.20	-1.04	0.64	0	-	-	-	-	-

Studies: 1: Almenning *et al.*, 2015; 2: Konopka *et al.*, 2015; 3: Nasrekani *et al.*, 2016; 4: Roessler *et al.*, 2013; 5: Sa *et al.*, 2015; 6: Saremi *et al.*, 2013; 7: Saremi and Yaghoubi, 2016; 8: Stener-Victorin *et al.*, 2009; 9: Thomson *et al.*, 2008; 10: Turan *et al.*, 2015; 11: Vigorito *et al.*, 2007; 12: Vizza *et al.*, 2016. **Key:** 95% CI: 95% confidence intervals; SBP: systolic blood pressure; DBP: diastolic blood pressure; FBG: fasting blood glucose; FI: fasting insulin; HOMA-IR: homeostatic model of assessment - insulin resistance; TC: total cholesterol; LDL-C: low-density lipoprotein cholesterol; HDL-C: high-density lipoprotein cholesterol; TG: triglycerides; RHR: resting heart rate; BMI: body mass index; WC: waist circumference; WHR: waist to hip ratio; FFM: fat free mass; SHBG: sex hormone binding globulin; Free T: free testosterone; FAI: free androgen index; FG: Ferriman-Gallwey score; DHEA-S: dehydroepiandrosterone sulfate; LH: luteinising hormone; FSH: follicle stimulating hormone; PG: progesterone; hsCRP: high-sensitivity C-reactive protein; AMH: anti-Müllerian hormone. ▲: positive values favour exercise over control. •: Study only included in the change from baseline analysis; statistically significant effects denoted by: * $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$; *N:* number or participants included within analysis. Effect estimates are reported as mean differences (MD) and 95% confidence intervals, between exercise and usual care groups. Heterogeneity reported using I^2 statistic.

3.5.2. Blood Pressure

Four eligible trials (158 participants) assessed changes in blood pressure (Sá *et al.*, 2015; Stener-Victorin *et al.*, 2009; Turan *et al.*, 2015; Vigorito *et al.*, 2007). There was no significant effect of exercise on systolic blood pressure (SBP) or diastolic blood pressure (DBP) for either change scores or post-intervention values compared with control (Table 3.3). When assessing the quality of the evidence, the result of both SBP and DBP was low-quality evidence due to imprecision (small number of participants, and a null and appreciable effect were included in the 95% CI for the MD), and high or unclear risk of selection bias, detection bias, reporting bias, attrition bias, and contamination (Table 3.2).

In subgroup analyses (Table 3.4), there was evidence of effects for supervised interventions (MD: -4.42 mmHg, 95% CI: -8.32 to -0.51; 3 trials, 147 participants, I^2 = 31%) for SBP change scores compared with control. No effects were found in subgroup analysis of SBP post-intervention values or in any DBP subgroup analysis.

3.5.3. Fasting Blood Glucose

Based on data from nine trials (263 participants), there was no effect of exercise on fasting blood glucose (FBG) change from baseline values compared with control (Table 3.3: Almenning, *et al.*, 2015; Konopka *et al.*, 2015; Sá *et al.*, 2015; Saremi *et al.*, 2013; Saremi and Yaghoubi, 2016; Stener-Victorin *et al.*, 2009; Turan *et al.*, 2015; Vigorito *et al.*, 2007; Vizza *et al.*, 2016). Due to unreported data, one trial was removed (Konopka *et al.*, 2015) for the comparison of absolute post-intervention values (leaving 238 participants in this analysis); there was still no observed effect. There was also no effect of exercise for any of the subgroup analyses (Table 3.4). These findings were judged to be low-quality evidence due to an unclear or high-risk of selection, detection, and reporting bias, contamination, low adherence, small number of participants, and the incorporation of a null or negligible effect and appreciable benefit included in the confidence interval for the mean difference (Table 3.2).

3.5.4. Fasting Insulin

Meta-analysis of data from 263 participants in nine trials (Almenning, *et al.*, 2015; Konopka *et al.*, 2015; Sá *et al.*, 2015; Saremi *et al.*, 2013; Saremi and Yaghoubi, 2016; Stener-Victorin *et al.*, 2009; Turan *et al.*, 2015; Vigorito *et al.*, 2007; Vizza *et al.*, 2016) revealed a favourable effect of exercise on the change from baseline values of FI compared with control (MD: -2.44 µIU/mL, 95% CI: -4.24 to -0.64; Figure 3.3), but there was evidence of considerable heterogeneity (I^2 = 91%). Similarly, statistically significant lowering effects of exercise versus control were found for FI post-intervention values (MD: -2.11 µIU/mL, 95% CI: -3.49 to -0.73; 8 trials, 238 participants, I^2 = 40%). Applying GRADE guidelines, the quality of the evidence was deemed as very low-quality evidence (Table 3.2) due to unclear or high risk judgements of randomisation or allocation procedures, lack of (assessor) blinding, high rate of incomplete outcome data, unclear reporting of outcomes and contamination, low adherence, considerable heterogeneity in the effects of individual studies, small numbers of participants, and wide confidence interval for the mean difference.

In sensitivity analyses, when only trials with larger sample sizes ($n \geq 30$ total participants) were included, although slightly reduced, the observed effect on change from baseline FI remained (MD: -1.09 µIU/mL, 95% CI: -1.64 to -0.53; 2 trials, 120 participants, I^2 = 7%). Similarly, when analysis included only studies with a low risk of bias, an effect of exercise was also found (MD: -3.18 µIU/mL, 95% CI: -5.63 to -0.74; 187 participants, 5 trials, I^2 = 95%). Likewise, post-intervention FI effects remained when small trials (MD: -1.73 µIU/mL, 95% CI: -3.00 to -0.47; 2 trials, 160 participants, I^2 = 5%) and trials with a high risk of bias (MD: -2.10 µIU/mL, 95% CI: -3.04 to -1.17; 5 trials, 187 participants, I^2 = 0%) were removed.

To identify the potential source of heterogeneity in the FI change from baseline analysis, the study by Saremi and Yaghoubi (2016) was removed because it was the greatest outlier. This reduced the I^2 statistic to a level that may not be important (18%) and the effect was maintained (MD: -1.54 µIU/mL, 95% CI: -2.36 to -0.71). The results of the removed trial may have varied due to the mode of exercise used (resistance training) or the use of a placebo.

A statistical effect of exercise versus control on FI was shown in multiple subgroups (Table 3.4). A change in FI from baseline to immediately post-intervention was observed in studies with participants who had a BMI that was 25-29.9 kg/m^2 (MD: -3.25 µIU/mL, 95% CI: -5.27 to -1.22; 5 trials, 168 participants, I^2 = 75%); in interventions that utilised aerobic based exercise (MD: -2.22 µIU/mL, 95% CI: -3.57 to -0.86; 6 trials, 192 participants, I^2 = 10%); also in interventions that were ≤ 12 weeks in duration (MD: -2.92 µIU/mL, 95% CI: -4.91 to -0.93; 7 trials, 225 participants, I^2 = 93%); and where participants were either fully supervised or had a combination of supervised and unsupervised exercise sessions (MD: -2.54 µIU/mL, 95% CI: -4.82 to -0.26; 6 trials, 214 participants, I^2 = 94%, and MD: -3.08 µIU/mL, 95% CI: -5.63 to -0.53; 2 trials, 38 participants, I^2 = 17%, respectively).

Compared with control, there were favourable effects of exercise on post-intervention FI values for participants with a BMI that was 25-29.9 kg/m^2 (MD: -2.27 µIU/mL, 95% CI: -3.24 to -1.31; 5 trials, 168 participants, I^2 = 0%); in interventions that were aerobic exercise-based (MD: -2.48 µIU/mL, 95% CI: -3.92 to -1.04; 5 trials, 167 participants, I^2 = 10%); ≤ 12 weeks duration (MD: -1.80 µIU/mL, 95% CI: -3.18 to -0.42; 6 trials, 200 participants, I^2 = 32%); and supervised (MD: -2.39 µIU/mL, 95% CI: -3.62 to -1.17; 5 trials, 189 participants, I^2 = 30%).

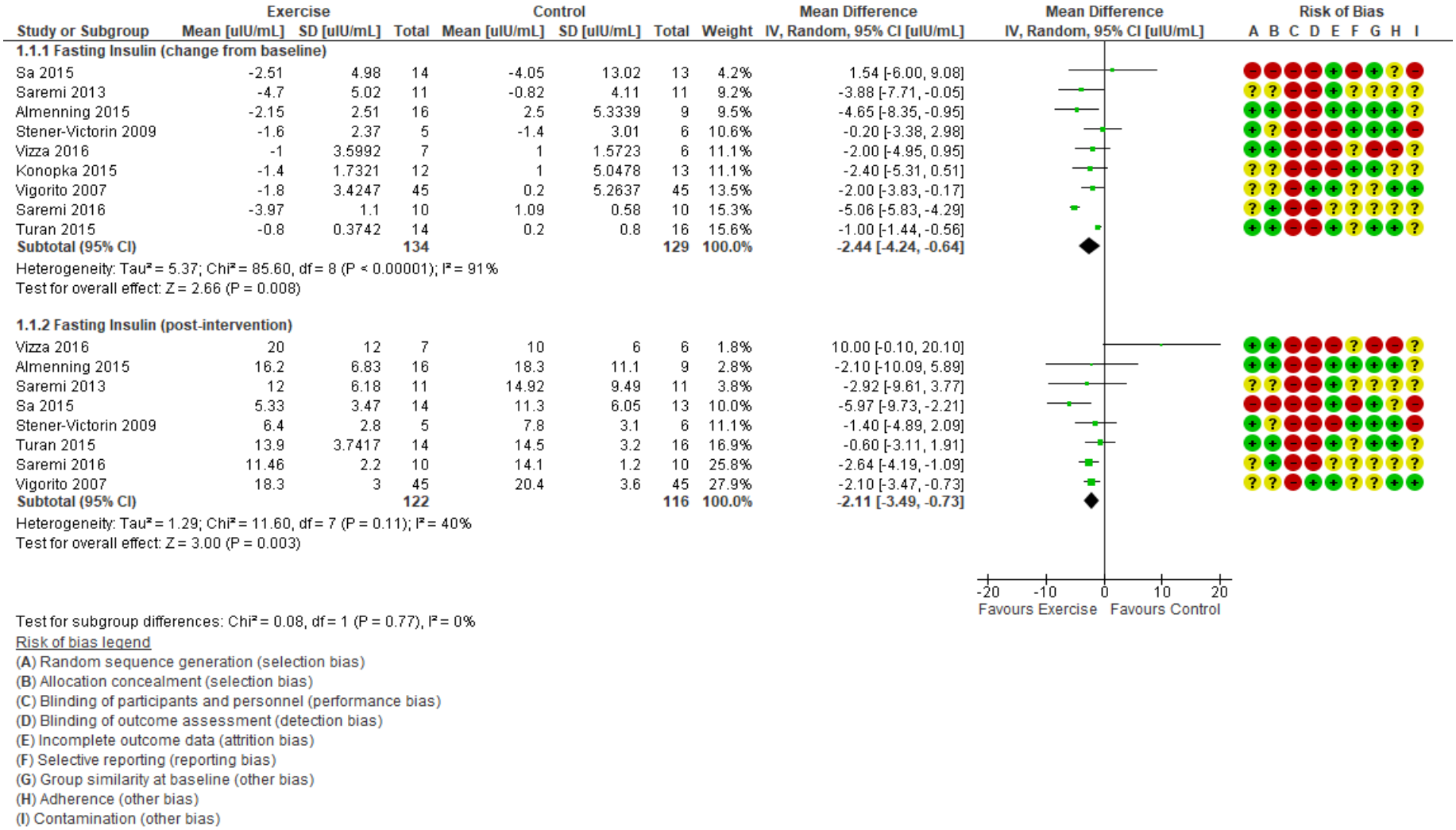


Figure 3.4. Forest plot of comparison: Exercise vs. Control, outcome: fasting insulin plasma levels (µIU/mL).

3.5.5. HOMA-IR

For the eight studies that were included (Almenning, *et al.*, 2015; Konopka *et al.,* 2015; Sá *et al.*, 2015; Saremi *et al.,* 2013; Saremi and Yaghoubi, 2016; Stener-Victorin *et al.,* 2009; Turan *et al.*, 2015; Vizza *et al.*, 2016), there was evidence of greater reductions in HOMA-IR change scores for exercise when compared to control (MD: -0.57, 95% CI: -0.99 to -0.14; 173 participants, I^2 = 87%; Table 3.3; Figure 3.5), but this was not the case and no effect of exercise was evident when the comparison of post-intervention HOMA-IR was completed (Konopka *et al.,* 2015 removed due to lack of data). In a sensitivity analysis that included only trials judged to have a low risk of bias, the effect of exercise was maintained (MD: -0.81, 95% CI: -1.40 to -0.21; 97 participants, 4 trials, I^2 = 77%). Only one trial had a sample size of ≥ 30 participants (Turan *et al.,* 2015), so a corresponding sensitivity analysis was not completed. The quality of the evidence was rated as very low-quality; this was due to unclear or high risk of selection, detection, attrition, and reporting bias, contamination, low adherence, considerable heterogeneity with minimal or no overlap of confidence intervals, a small number of participants, and a null or negligible effect and appreciable benefit included in the confidence interval for the mean difference (Table 3.2).

In the investigation of heterogeneity, the most extreme value was removed (Almenning, *et al.*, 2015); this had a negligible effect on the degree of heterogeneity (I^2 = 89%), but a small effect was still evident (MD: -0.50, 95% CI: -0.96 to -0.05). Similarly, the I^2 statistic remained representative of at least substantial heterogeneity in sub-analyses; the lowest reported value was 60% which was observed in the aerobic exercise intervention subgroup (Table 3.4).

In subgroup analyses statistical effects on HOMA-IR were revealed for change from baseline to immediately post-intervention values for aerobic exercise interventions (MD: -0.73, 95% CI: -1.24 to -0.21; 5 trials, 102 participants, I^2 = 60%); interventions which were ≤ 12 weeks in duration (MD: -0.69, 95% CI: -1.13 to -0.26; 6 trials, 135 participants, I^2 = 89%); for interventions that had a supervised delivery (MD: -0.80, 95% CI: -1.19 to -0.42; 5 trials, 124 participants, I^2 = 76%); and also for participants in the BMI 25-29.9 kg/m^2 subgroup (MD: -0.83, 95% CI: -1.39 to -0.26; 4 trials, 78

participants, I^2 = 75%). There were no statistical effects observed in the post-intervention subgroup analyses (Table 3.4).

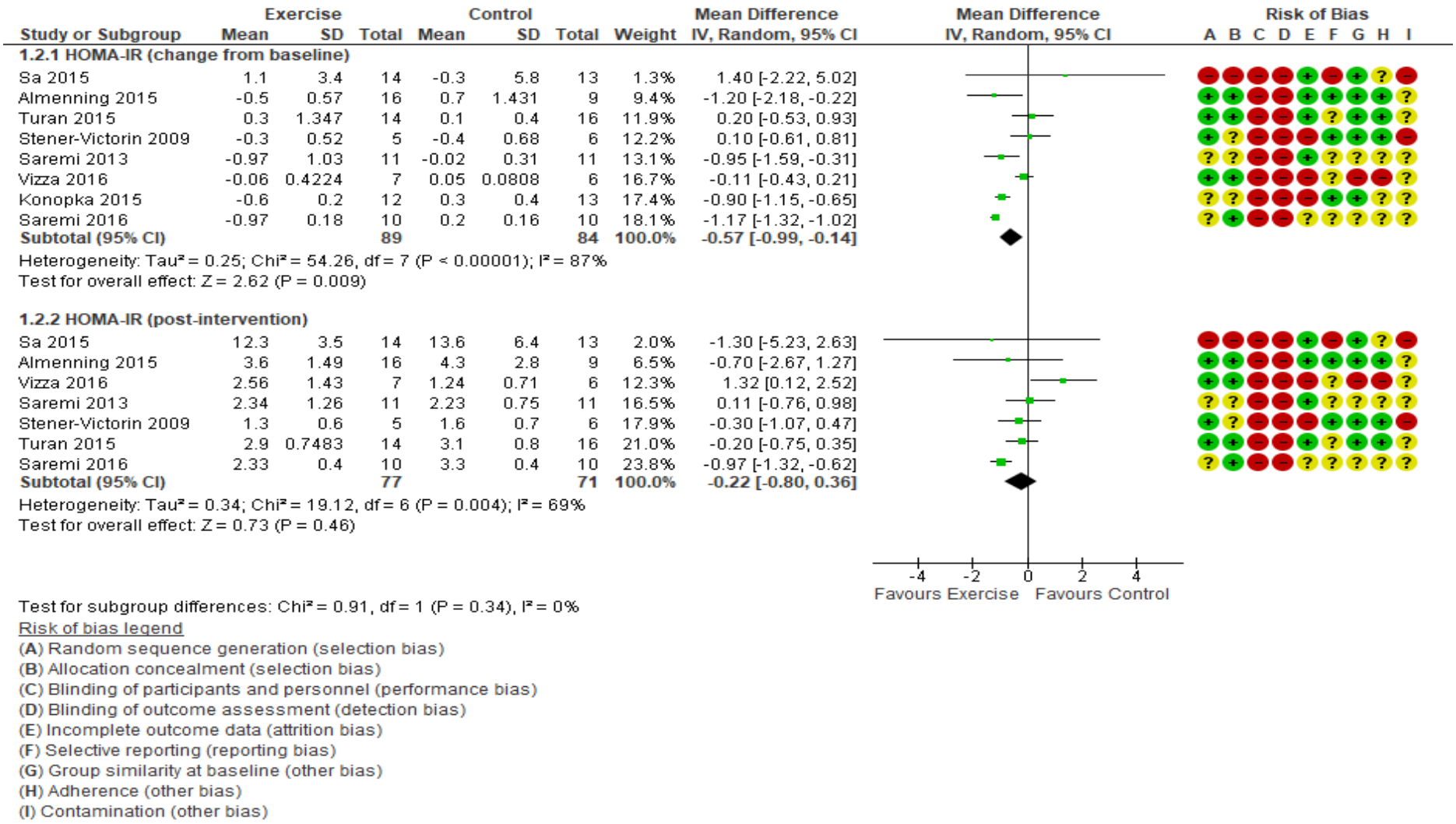


Figure 3.5. Forest plot of comparison: Exercise vs. Control, outcome: Homeostatic Model Assessment for Insulin Resistance (HOMA-IR).

Table 3.4. Summary of effect estimates and heterogeneity from sub-group analyses in blood pressure and glucose homeostasis related outcomes.

Outcome	Sub-analysis	Sub-group	Change from baseline			Post-intervention		
			Trials (*N*)	Effect Estimate MD (95% CI)	I^2 (%)	Trials (*N*)	Effect Estimate MD (95% CI)	I^2 (%)
SBP (mmHg)	BMI at entry	25-29.9 kg/m^2	2 (101)	-1.80 (-8.53 to 4.93)	65	2 (101)	0.19 (-12.94 to 13.31)	82
	Intervention	Aerobic exercise	3 (128)	-3.71 (-8.88 to 1.47)	60	3 (128)	-1.41 (-8.65 to 5.82)	65
	Duration	≤12 weeks	2 (120)	-3.03 (-7.54 to 1.47)	27	2 (120)	1.90 (-13.19 to 16.99)	95
		>12 weeks	2 (38)	-2.91 (-12.41 to 6.60)	79	2 (38)	2.06 (-8.29 to 12.41)	59
	Format	Supervised	3 (147)	-4.42 (-8.32 to -0.51) *	31	3 (147)	0.45 (-10.04 to 10.94)	90
DBP (mmHg)	BMI at entry	25-29.9 kg/m^2	2 (101)	-1.19 (-3.11 to 0.73)	0	2 (101)	-1.31 (-3.09 to 0.47)	0
	Intervention	Aerobic exercise	3 (128)	-2.67 (-6.50 to 1.17)	62	3 (128)	-1.41 (-3.12 to 0.29)	0
	Duration	≤12 weeks	2 (120)	-1.30 (-3.31 to 0.71)	0	2 (120)	0.96 (-4.33 to 6.26)	68
		>12 weeks	2 (38)	-3.95 (-11.78 to 3.89)	77	2 (38)	-3.16 (-7.54 to 1.22)	0
	Format	Supervised	3 (147)	-3.03 (-7.36 to 1.30)	60	3 (147)	-0.22 (-3.50 to 3.07)	43
FBG (mg/dL)	BMI at entry	25-29.9 kg/m^2	5 (168)	-0.79 (-2.08 to 0.50)	0	5 (168)	-1.59 (-5.29 to 2.11)	62
		≥ 30 kg/m^2	3 (65)	-0.87 (-8.95 to 7.22)	58	2 (40)	-1.21 (-8.83 to 6.41)	0
	Intervention	Aerobic exercise	6 (192)	-0.70 (-2.46 to 1.05)	21	5 (167)	-0.83 (-2.80 to 1.13)	0
		Resistance exercise	3 (50)	-1.01 (-3.37 to 1.34)	11	3 (50)	-3.81 (-13.74 to 6.11)	76
	Duration	≤12 weeks	7 (225)	-1.47 (-3.03 to 0.10)	18	6 (200)	-2.18 (-5.82 to 1.46)	53
		>12 weeks	2 (38)	0.38 (-2.42 to 3.19)	0	2 (38)	-0.33 (-4.29 to 3.64)	0
	Format	Supervised	6 (214)	-1.75 (-4.06 to 0.56)	40	5 (189)	-3.04 (-7.59 to 1.52)	60
		Mixed Delivery	2 (38)	-0.73 (-3.05 to 1.58)	0	2 (38)	0.00 (-4.91 to 4.91)	0
FI (µIU/mL)	BMI at entry	25-29.9 kg/m^2	5 (168)	-3.25 (-5.27 to -1.22) **	75	5 (168)	-2.27 (-3.24 to -1.31) ***	0
		≥ 30 kg/m^2	3 (65)	-1.94 (-3.94 to 0.06)	0	2 (40)	1.30 (-14.29 to 16.89)	88
	Intervention	Aerobic exercise	6 (192)	-2.22 (-3.57 to -0.86) ***	10	5 (167)	-2.48 (-3.92 to -1.04) ***	10
		Resistance exercise	3 (50)	-3.99 (-5.97 to -2.00) ***	54	3 (50)	-0.24 (-6.99 to 6.51)	68
	Duration	≤12 weeks	7 (225)	-2.92 (-4.91 to -0.93) **	93	6 (200)	-1.80 (-3.18 to -0.42) **	32
		>12 weeks	2 (38)	0.06 (-2.87 to 2.99)	0	2 (38)	-3.63 (-8.11 to 0.85)	67
	Format	Supervised	6 (214)	-2.54 (-4.82 to -0.26) *	94	5 (189)	-2.39 (-3.62 to -1.17) ***	30
		Mixed Delivery	2 (38)	-3.08 (-5.63 to -0.53) *	17	2 (38)	3.54 (-8.29 to 15.37)	71
HOMA-IR	BMI at entry	25-29.9 kg/m^2	4 (78)	-0.83 (-1.39 to -0.26) **	75	4 (78)	-0.51 (-1.10 to 0.07)	55
		≥ 30 kg/m^2	3 (65)	-0.43 (-1.19 to 0.32)	87	2 (40)	0.71 (-1.47 to 2.88)	55

Outcome	Sub-analysis	Sub-group	Change from baseline			Post-intervention		
			Trials (*N*)	Effect Estimate MD (95% CI)	I^2 (%)	Trials (*N*)	Effect Estimate MD (95% CI)	I^2 (%)
	Intervention	Aerobic exercise	5 (102)	-0.73 (-1.24 to -0.21) **	60	4 (77)	-0.15 (-0.70 to 0.40)	0
		Resistance exercise	3 (50)	-0.74 (-1.58 to 0.10)	94	3 (50)	-0.24 (-1.89 to 1.41)	85
	Duration	≤12 weeks	6 (135)	-0.69 (-1.13 to -0.26) **	89	5 (110)	-0.14 (-0.88 to 0.59)	78
	Format	Supervised	5 (124)	-0.80 (-1.19 to -0.42) ***	76	4 (99)	-0.46 (-1.09 to 0.17)	66
		Mixed Delivery	2 (38)	-0.55 (-1.60 to 0.50)	77	2 (38)	0.47 (-1.49 to 2.42)	66

Key: Outcome: outcome where sub-analysis was completed. Sub-analysis: how the studies were categorised for analysis; BMI at entry: mean baseline BMI of participants; Intervention: type of intervention; Duration: length of exercise intervention; Format: delivery mode of intervention. Sub-group: groups each study was classified into. Trials: number of studies included within sub-analysis, *N*: number or participants included within sub-analysis. Effect estimates are reported as mean difference (MD), and 95% confidence intervals, between exercise and control groups. Significant evidence of effect denoted by: * $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$. Heterogeneity reported using I^2 statistic: 0-40% might not be important; 30-60% may represent moderate heterogeneity; 50-90% may represent substantial heterogeneity; 75-100% may represent considerable heterogeneity. SBP: systolic blood pressure; DBP: diastolic blood pressure; FBG: fasting blood glucose; FI: fasting insulin; BMI: body mass index.

3.5.6. Circulating lipids

Seven trials (225 participants) were included in the analysis of all outcomes related to participants' lipid profile (Almenning, *et al.,* 2015; Sá *et al.,* 2015; Saremi *et al.,* 2013; Saremi and Yaghoubi, 2016; Stener-Victorin *et al.,* 2009; Turan *et al.,* 2015; Vigorito *et al.,* 2007). A statistical effect of exercise compared to control change scores was observed for total cholesterol (MD: -5.88 mg/dL, 95% CI: -9.92 to -1.83; I^2 = 35%), LDL-C (MD: -7.39 mg/dL, 95% CI: -9.83 to -4.95; I^2 = 0%), and triglycerides (MD: -4.78 mg/dL, 95% CI: -7.52 to -2.05; I^2 = 3%), but not for HDL-C (Figure 3.6). Analysis of post-intervention values for lipid-related outcomes revealed an effect on total cholesterol (MD: -6.35 mg/dL, 95% CI: -10.76 to -1.95; I^2 = 0%) and LDL-C (MD: -6.68 mg/dL, 95% CI: -11.66 to -1.70; I^2 = 0%) only (Table 3.3). These results were rated as low-quality evidence (Table 3.2) due to high or unclear risk of selection bias, detection bias, reporting bias, contamination, imprecision due to small number of participants and also wide confidence intervals in the included trials.

In sensitivity analyses, the favourable effects of exercise versus control on change scores for total cholesterol (MD: -5.94 md/dL, 95% CI: -10.32 to -1.55; 5 trials, 187 participants, I^2 = 40%), LDL-C (MD: -6.60 mg/dL, 95% CI: -9.88 to -3.32; 5 trials, 187 participants, I^2 = 14%), and triglycerides (MD: -5.97 mg/dL, 95% CI: -10.91 to -1.03; 5 trials, 187 participants, I^2 = 33%), were maintained in studies with a low risk of bias, and also in larger trials (MD: -3.74 mg/dL, 95% CI: -6.13 to -1.35; 120 participants, 2 trials, I^2 = 0%; MD: -8.58, 95% CI: -11.44 to -5.71; 120 participants, 2 trials, I^2 = 0%; and MD: -3.62 mg/dL, 95% CI: -6.22 to -1.02; 120 participants, 2 trials, I^2 = 0%, respectively). Sensitivity analyses for LDL-C post-intervention values showed a retained effect when trials with a high risk of bias were excluded (MD: -8.64 mg/dL, 95% CI: -16.30 to -0.98; 5 trials, 187 participants, I^2 = 22%), but not when smaller trials were removed.

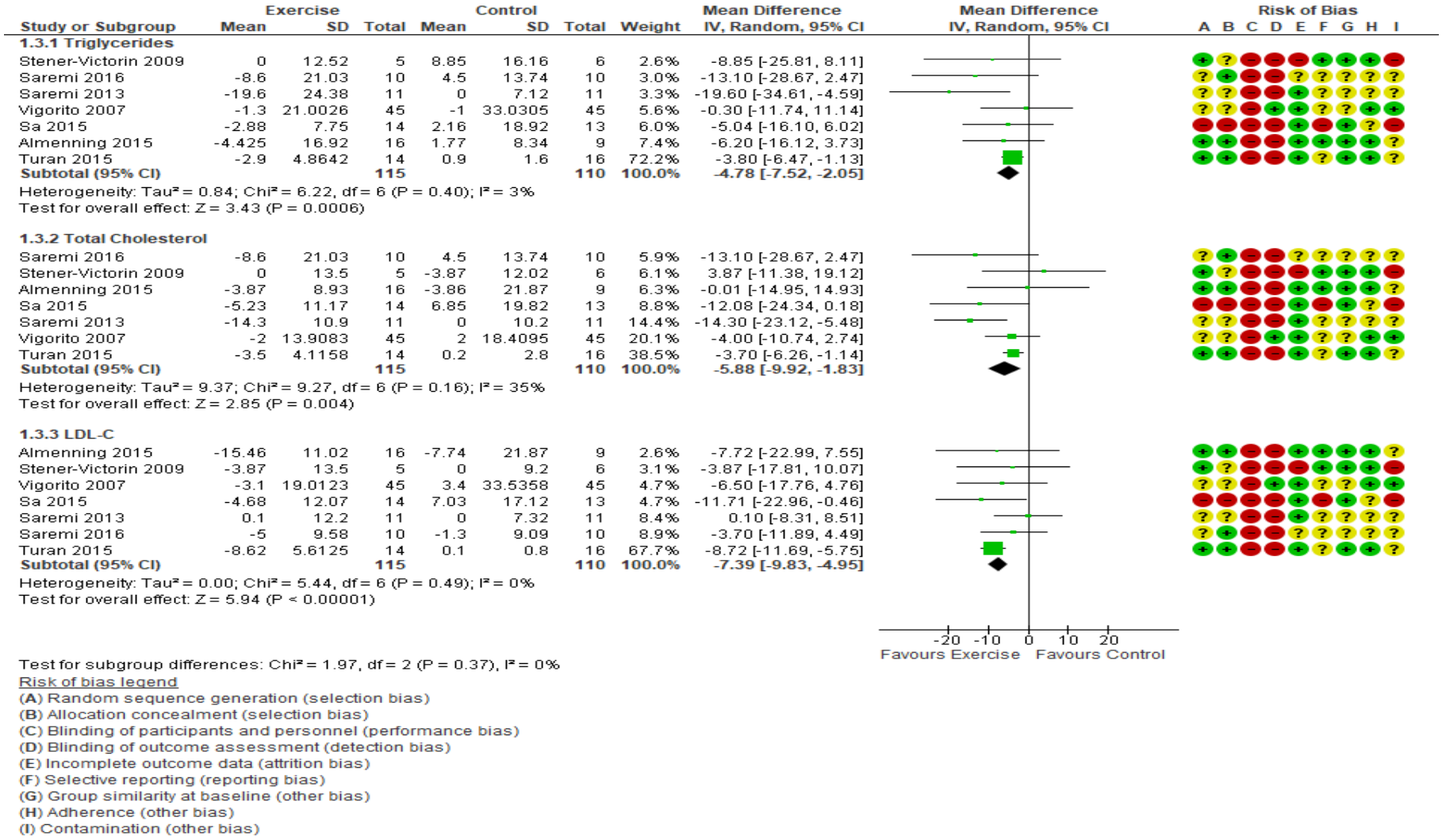


Figure 3.6. Forest plot of comparison: Exercise vs. Control, change from baseline to immediately post-intervention analysis of outcomes related to lipid profile [circulating triglycerides, total cholesterol, and low-density lipoprotein cholesterol (LDL-C) levels; all in mg/dL].

Subgroup analyses of total cholesterol change (Table 3.5) revealed statistical effects for interventions that were ≤ 12 weeks duration (MD: -5.94 mg/dL, 95% CI: -10.32 to -1.55; 5 trials, 187 participants, I^2 = 37%) or where participants were supervised (MD: -7.25 mg/dL, 95% CI: -11.92 to -2.58; 5 trials, 189 participants, I^2 = 48%). There was also evidence of an effect in subgroup analysis for change from baseline (MD: -6.68 mg/dL, 95% CI: -13.00 to -0.35; 5 trials, 167 participants, I^2 = 39%) and post-intervention total cholesterol values (MD: -6.90 mg/dL, 95% CI: -11.90 to -1.90; 5 trials, 167 participants, I^2 = 0%) when interventions were aerobic in nature. Subgroup analysis of post-intervention total cholesterol also revealed a statistical effect when interventions were > 12 weeks in length (MD: -9.92 mg/dL, 95% CI: -17.81 to -2.04; 2 trials, 38 participants, I^2 = 0%) or when participants were supervised (MD: -6.76 mg/dL, 95% CI: -11.27 to -2.26; 5 trials, 189 participants, I^2 = 0%).

In subgroup analyses for change from baseline LDL-C, a statistically favourable exercise effect was found when interventions were ≤ 12 weeks duration (MD: -6.60 mg/dL, 95% CI: -9.88 to -3.32; 5 trials, 187 participants, I^2 = 13%), or interventions were supervised (MD: -6.70 mg/dL, 95% CI: -10.29 to -3.12; 5 trials, 189 participants, I^2 = 23%). Subgroup analysis for post-intervention LDL-C circulating levels revealed favourable exercise effects in participants with BMI of 25-29.9 kg/m^2 (MD: -9.54 mg/dL, 95% CI: -18.71 to -0.36; 5 trials, 168 participants, I^2 = 22%), and interventions of ≤ 12 weeks duration (MD: -8.64 mg/dL, 95% CI: -16.30 to -0.98; 5 trials, 187 participants, I^2 = 22%), that were supervised (MD: -7.58 mg/dL, 95% CI: -13.73 to -1.43; 5 trials, 187 participants, I^2 = 24%) or aerobic in nature (MD: -5.87 mg/dL, 95% CI: -11.68 to -0.07; 5 trials, 167 participants, I^2 = 0%; Table 3.5).

For HDL-C, only subgroup analyses of resistance training interventions showed statistically significant effects; resistance training had a negative effect on change from baseline levels (MD: -2.19 mg/dL, 95% CI: -4.21 to -0.18; 2 trials, 37 participants, I^2 = 0%), but a positive effect on post-intervention values (MD: 7.29 mg/dL, 95% CI: 1.11 to 13.46; 2 trials, 37 participants, I^2 = 17%; Table 3.5). No effects of exercise were found in other HDL-C subgroup analyses.

Compared with control, exercise had a statistically favourable effect on circulating triglyceride values in the following subgroups: BMI: 25-29.9 kg/m^2 (MD: -8.17 mg/dL, 95% CI: -14.44 to -1.89; 5 trials, 167 participants, I^2 = 13%); aerobic exercise interventions (MD: -6.80 mg/dL, 95% CI: -13.12 to -0.48; 5 trials, 167 participants, I^2 = 5%); ≤12 weeks duration (MD: -6.06 mg/dL, 95% CI: -10.82 to -1.31; 5 trials, 187 participants, I^2 = 30%); and when interventions were supervised (MD: -5.91 mg/dL, 95% CI: -10.75 to -1.06; 5 trials, 189 participants, I^2 = 29%; Table 3.5). Analysis of triglyceride post-intervention values only revealed an effect of exercise in trials that were > 12 weeks (MD: -13.85 mg/dL, 95% CI: -26.33 to -1.36; 2 trials, 38 participants, I^2 = 0%).

Table 3.5. Summary of effect estimates and heterogeneity from sub-group analyses in lipidemic related outcomes.

Outcome	Sub-analysis	Sub-group	Change from baseline			Post-intervention		
			Trials (*N*)	Effect Estimate MD (95% CI)	I^2 (%)	Trials (*N*)	Effect Estimate MD (95% CI)	I^2 (%)
Triglycerides (mg/dL)	BMI at entry	25-29.9 kg/m^2	5 (168)	-8.17 (-14.44 to -1.89) **	13	5 (168)	3.04 (-4.97 to 11.05)	0
	Intervention	Aerobic exercise	5 (167)	-6.80 (-13.12 to -0.48) *	5	5 (167)	-2.89 (-14.44 to 8.65)	41
		Resistance exercise	2 (37)	-9.91 (-22.32 to 2.49)	0	2 (37)	6.05 (-12.08 to 24.19)	0
	Duration	≤12 weeks	5 (187)	-6.06 (-10.82 to -1.31) **	30	5 (187)	-1.10 (-4.73 to 2.54)	0
		>12 weeks	2 (38)	-6.18 (-15.44 to 3.09)	0	2 (38)	-13.85 (-26.33 to -1.36) *	0
	Format	Supervised	5 (189)	-5.91 (-10.75 to -1.06) *	29	5 (189)	-2.49 (-6.77 to 1.79)	7
TC (mg/dL)	BMI at entry	25-29.9 kg/m^2	5 (168)	-6.30 (-12.81 to 0.21)	41	5 (168)	-4.16 (-10.31 to 2.00)	0
	Intervention	Aerobic exercise	5 (167)	-6.68 (-13.00 to -0.35) *	39	5 (167)	-6.90 (-11.90 to -1.90) **	0
		Resistance exercise	2 (37)	-9.72 (-21.67 to 2.22)	0	2 (37)	6.47 (-16.70 to 29.63)	0
	Duration	≤12 weeks	5 (187)	-5.94 (-10.32 to -1.55) **	37	5 (187)	-4.74 (-10.05 to 0.57)	0
		>12 weeks	2 (38)	-4.78 (-20.35 to 10.80)	61	2 (38)	-9.92 (-17.81 to -2.04) **	0
	Format	Supervised	5 (189)	-7.25 (-11.92 to -2.58) **	48	5 (189)	-6.76 (-11.27 to -2.26) **	0
LDL-C (mg/dL)	BMI at entry	25-29.9 kg/m^2	5 (168)	-3.41 (-8.05 to 1.24)	0	5 (168)	-9.54 (-18.71 to -0.36) *	22
	Intervention	Aerobic exercise	5 (167)	-4.17 (-9.23 to 0.90)	0	5 (167)	-5.87 (-11.68 to -0.07) *	0
		Resistance exercise	2 (37)	-6.50 (-16.32 to 3.32)	22	2 (37)	-13.57 (-38.44 to 11.29)	45
	Duration	≤12 weeks	5 (187)	-6.60 (-9.88 to -3.32) ***	13	5 (187)	-8.64 (-16.30 to -0.98) *	22
		>12 weeks	2 (38)	-8.62 (-17.37 to 0.14)	0	2 (38)	-5.05 (-12.97 to 2.86)	0
	Format	Supervised	5 (187)	-6.70 (-10.29 to -3.12) ***	23	5 (187)	-7.58 (-13.73 to -1.43) *	24
HDL-C▲ (mg/dL)	BMI at entry	25-29.9 kg/m^2	5 (168)	0.99 (-2.89 to 4.88)	61	5 (168)	3.36 (-3.33 to 10.05)	62
	Intervention	Aerobic exercise	5 (167)	2.69 (-1.47 to 6.86)	59	5 (167)	1.04 (-3.06 to 5.15)	29
		Resistance exercise	2 (37)	-2.19 (-4.21 to -0.18) *	0	2 (37)	7.29 (1.11 to 13.46) *	17
	Duration	≤12 weeks	5 (187)	-0.10 (-2.27 to 2.08)	57	5 (187)	2.83 (-2.73 to 8.40)	76
		>12 weeks	2 (38)	2.93 (-3.96 to 9.82)	64	2 (38)	1.25 (-1.42 to 3.92)	0
	Format	Supervised	5 (189)	-0.32 (-1.87 to 1.23)	45	5 (189)	1.93 (-1.86 to 5.72)	75

Key: Outcome: outcome where sub-analysis was completed. Sub-analysis: how the studies were categorised for analysis; BMI at entry: mean baseline BMI of participants; Intervention: type of intervention; Duration: length of exercise intervention; Format: delivery mode of intervention. Sub-group: groups each study was classified into. Trials: number of studies included within sub-analysis, *N*: number or participants included within sub-analysis. Effect estimates are reported as mean difference (MD), and 95% confidence intervals, between exercise and control groups. Significant evidence of effect denoted by: * $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$. Heterogeneity reported using I^2

statistic: 0-40% might not be important; 30-60% may represent moderate heterogeneity; 50-90% may represent substantial heterogeneity; 75-100% may represent considerable heterogeneity. ▲: positive values favour exercise over control. TC: total cholesterol; LDL-C: low-density lipoprotein cholesterol; HDL-C: high-density lipoprotein cholesterol.

3.5.7. Exercise versus Control: Secondary outcomes

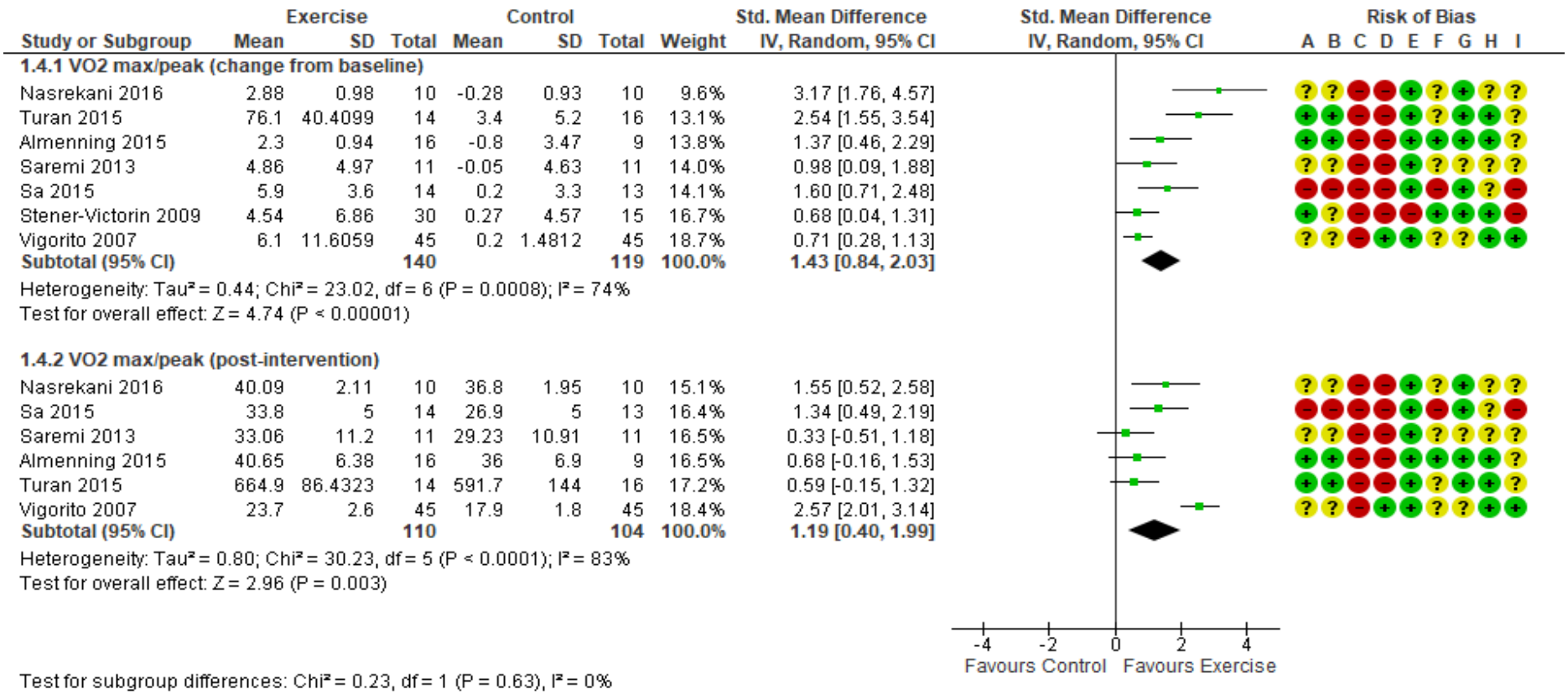
In addition to the primary outcomes, effect estimates were calculated for all outcomes that were reported within the included studies. Where more than one study reported an outcome, meta-analyses have been completed (Table 3.3); single study outcomes have also been reported qualitatively within the results.

3.5.8. Maximal or peak oxygen uptake

Seven trials were included in the meta-analysis of maximal (VO_2 max) or peak (VO_2 peak) values (Almenning, *et al.*, 2015; Nasrekani *et al.*, 2016; Sá *et al.*, 2015; Saremi *et al.,* 2013; Stener-Victorin *et al.,* 2009; Turan *et al.,* 2015; Vigorito *et al.,* 2007). A large statistical effect of exercise versus control was reported for both change scores (SMD: 1.43, 95% CI: 0.84 to 2.03; 259 participants, 7 trials, I^2 = 74%) and comparison of post-intervention (SMD: 1.19, 95% CI: 0.40 to 1.99; I^2 = 83%) VO_2 max/peak values (Figure 3.7). When studies that only reported relative VO_2 max/peak values (*i.e.* expressed as ml/kg/min) were included (Figure 3.8), the effect of exercise was maintained in both change from baseline values (MD: 3.84 ml/kg/min, 95% CI: 2.87 to 4.81; 6 trials, 229 participants, I^2 = 17%) and post-intervention values and (MD: 5.01 ml/kg/min, 95% CI: 3.48 to 6.54; 5 trials, 184 participants, I^2 = 42%).

Sensitivity analyses were completed for the SMD VO_2 max/peak change from baseline values. An effect was maintained when small trials (SMD: 1.21, 95% CI: 0.29 to 2.12; 3 trials, 165 participants, I^2 = 83%) and those with a high risk of bias (SMD: 1.63, 95% CI: 0.78 to 2.48; 5 trials, 187 participants, I^2 = 80%) were removed. SMD was also used to complete sensitivity analysis for sample size in the comparison of post-intervention values; 2 trials (Turan *et al.,* 2015; Vigorito *et al.,* 2007), involving 120 participants were included, but the effect was lost. However, an effect remained when the trials with a high risk of bias were removed (SMD: 1.16, 95% CI: 0.21 to 2.12; 5 trials, 187 participants, I^2 = 87%).

When only studies incorporating relative VO_2 max/peak change scores were considered, small studies were removed in sensitivity analysis and the effect of exercise was maintained (MD: 1.21 ml/kg/min, 95% CI: 0.29 to 2.12, 165 participants, 3 trials, $I^2 = 83\%$). An effect also remained when studies with a high risk of bias were removed (MD: 3.35 ml/kg/min, 95% CI: 2.59 to 4.10; 157 participants, 4 trials, $I^2 = 0\%$). All trials in the post-intervention relative VO_2 max/peak analysis were considered low risk of bias, so a sensitivity analysis was not necessary as it duplicated the primary analysis.



Test for subgroup differences: Chi² = 0.23, df = 1 (P = 0.63), I² = 0%

Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Group similarity at baseline (other bias)
(H) Adherence (other bias)
(I) Contamination (other bias)

Figure 3.7. Forest plot of comparison: exercise vs. control, standardised mean difference; outcome: VO2 max/peak.

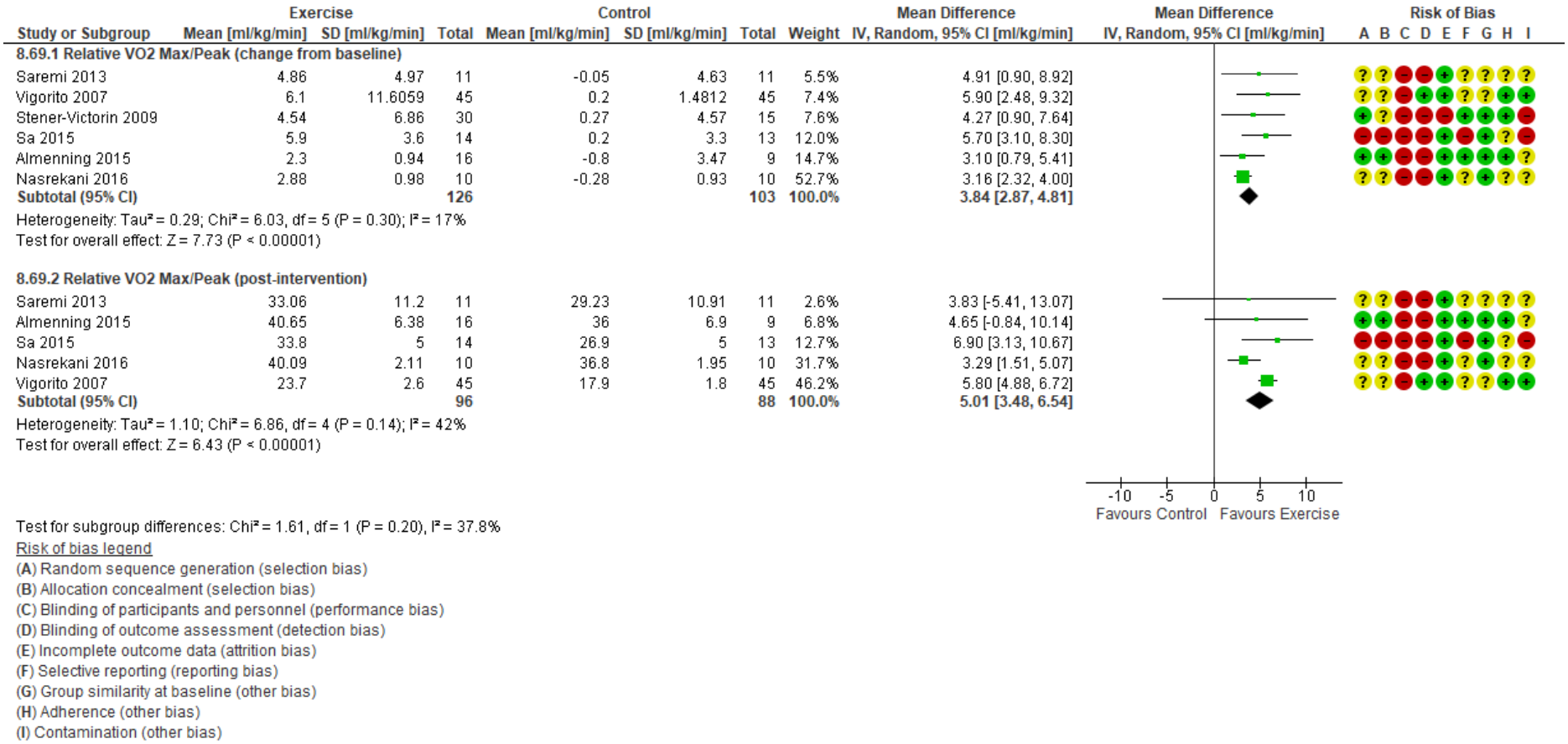


Figure 3.8. Forest plot of comparison: exercise vs. control, mean difference; outcome: relative VO2 max/peak (ml/kg/min).

For ease of interpretation, we performed subgroup analyses only on the relative VO_2 max/peak (*i.e.* ml/kg/min) data (Table 3.6). Subgroup analysis of the change from baseline relative VO_2 max/peak values revealed statistical improvements with aerobic exercise (MD: 4.11 ml/kg/min, 95% CI: 3.07 to 5.14; 221 participants, 6 trials; I^2 = 21%), for interventions that were short (MD: 3.35 ml/kg/min, 95% CI: 2.59 to 4.10; 157 participants, 4 trials; I^2 = 0%) or long (MD: 5.17 ml/kg/min, 95% CI: 3.11 to 7.23; 72 participants, 2 trials; I^2 = 0%) in duration and also for participants that had a BMI of 25-29.9 kg/m^2 (MD: 3.39 ml/kg/min, 95% CI: 2.66 to 4.13; 202 participants, 5 trials; I^2 = 0%). The post-intervention pooled analysis showed an effect of exercise on relative VO_2 max/peak in four subgroups: participants with a BMI of 25-29.9 kg/m^2 or ≤ 12 weeks in duration (both subgroups: MD: 4.70 ml/kg/min, 95% CI: 2.90 to 6.49; 157 participants, 4 trials; I^2 = 51%), and for aerobic exercise (MD: 5.05 ml/kg/min, 95% CI: 3.53 to 6.56; 176 participants, 5 trials; I^2 = 41%), and supervised (MD: 5.04 ml/kg/min, 95% CI: 3.25 to 6.82; 159 participants, 4 trials; I^2 = 56%) interventions.

Only one trial reported data from a 16-week post-intervention follow-up (Stener-Victorin *et al.*, 2009). They reported that a 12% increase in VO_2 max (4.11 ± 5.20 ml/kg/min), that was statistically different (P = .001) from baseline was still evident in the exercise group at follow-up. The corresponding change for control (+7%) was not statistically significant, and there were no significant differences between groups.

3.5.9. Resting heart rate

A meta-analysis of four included trials (Almenning, *et al.*, 2015; Stener-Victorin *et al.*, 2009; Turan *et al.*, 2015; Vigorito *et al.*, 2007) showed no effect of exercise on the change from baseline for resting heart rate (RHR) values (Table 3.3). However, in these trials, RHR post-intervention values were statistically lower in the exercise interventions when compared to control (MD: -3.26 beats/min, 95% CI: -4.93 to -1.59; 156 participants, I^2 = 0%). When the post-intervention sensitivity analyses were completed, this effect remained when only larger trials were considered (MD: -3.18 beats/min,

95% CI: -5.59 to -0.77, 3 trials, 145 participants, I^2 = 16%), and also in those with a low risk of bias (MD: -3.53 beats/min, 95% CI: -5.28 to -1.78; 2 trials, 120 participants, I^2 = 0%).

When subgroup analyses were completed, there was a statistical effect of exercise compared with control on both RHR change from baseline and post-intervention values in interventions that were aerobic exercise-based (MD: -3.32 beats/min, 95% CI: -5.50 to -1.15; 118 participants, 3 trials; I^2 = 0%; and, MD: -3.00 beats/min, 95% CI: -4.72 to -1.28; 118 participants, 3 trials; I^2 = 0% respectively), and in those that were supervised (MD: -4.06 beats/min, 95% CI: -7.42 to -0.70; 120 participants, 2 trials; I^2 = 26%; and, MD: -3.53 beats/min, 95% CI: -5.28 to -1.78; 120 participants, 2 trials; I^2 = 0% respectively). Post-intervention subgroup analysis also revealed effects in interventions of ≤ 12 weeks (MD: -3.18 beats/min, 95% CI: -5.59 to -0.77; 145 participants, 3 trials; I^2 = 16%) and when participants had a BMI 25-29.9 kg/m^2 at study entry (MD: -3.04 beats/min, 95% CI: -4.76 to -1.32; 126 participants, 3 trials; I^2 = 10%); these effects were not observed in the change form baseline data (Table 3.6).

Table 3.6. Summary of effect estimates and heterogeneity from sub-group analyses in cardiorespiratory, anthropometric and body composition related outcomes.

Outcome	Sub-analysis	Sub-group	Change from baseline			Post-intervention		
			Trials (*N*)	Effect Estimate MD (95% CI)	I^2 (%)	Trials (*N*)	Effect Estimate MD (95% CI)	I^2 (%)
VO_2 max/peak▲ (ml/kg/min)	BMI at entry	25-29.9 kg/m^2	5 (202)	3.39 (2.66 to 4.13) ***	0	4 (157)	4.70 (2.90 to 6.49) ***	51
	Intervention	Aerobic exercise	6 (221)	4.11 (3.07 to 5.14) ***	21	5 (176)	5.05 (3.53 to 6.56) ***	41
	Duration	≤12 weeks	4 (157)	3.35 (2.59 to 4.10) ***	0	4 (157)	4.70 (2.90 to 6.49) ***	51
	Format	Supervised	4 (159)	4.43 (2.76 to 6.10) ***	47	4 (159)	5.04 (3.25 to 6.82) ***	56
Resting Heart Rate (beats/min)	BMI at entry	25-29.9 kg/m^2	3 (126)	-1.76 (-4.41 to 0.89)	58	3 (126)	-3.04 (-4.76 to -1.32) ***	10
	Intervention	Aerobic exercise	3 (118)	-3.32 (-5.50 to -1.15) **	0	3 (118)	-3.00 (-4.72 to -1.28) ***	0
	Duration	≤12 weeks	3 (145)	-2.54 (-5.90 to 0.81)	66	3 (145)	-3.18 (-5.59 to -0.77) **	16
	Format	Supervised	2 (120)	-4.06 (-7.42 to -0.70) *	26	2 (120)	-3.53 (-5.28 to -1.78) ***	0
BMI (kg/m^2)	BMI at entry	≤ 24.9 kg/m^2	1 (30)	-0.14 (-1.23 to 0.95)	NA	1 (30)	0.00 (-3.05 to 3.05)	0
		25-29.9 kg/m^2	6 (222)	-0.01 (-0.47 to 0.45)	30	6 (188)	-1.04 (-1.89 to -0.19) *	0
		≥ 30 kg/m^2	4 (79)	-1.34 (-1.86 to -0.82) ***	0	3 (54)	-0.70 (-5.55 to 4.15)	38
	Intervention	Aerobic exercise	8 (260)	-0.78 (-1.38 to -0.18) **	57	7 (201)	-1.45 (-2.59 to -0.32) **	0
		Resistance exercise	3 (50)	0.50 (0.00 to 1.00) *	0	3 (50)	-0.10 (-2.76 to 2.55)	23
	Duration	≤12 weeks	8 (245)	-0.43 (-1.22 to 0.35)	64	7 (220)	-0.91 (-1.73 to -0.08) *	0
		>12 weeks	3 (86)	-0.61 (-1.61 to 0.38)	77	3 (52)	-2.42 (-5.28 to 0.45)	0
	Format	Supervised	8 (248)	-0.65 (-1.42 to 0.12)	74	7 (223)	-1.06 (-1.87 to -0.25) **	0
		Mixed Delivery	2 (38)	0.19 (-1.56 to 1.93)	0	2 (38)	1.82 (-5.85 to 9.50)	45
Body Mass (kg)	BMI at entry	25-29.9 kg/m^2	3 (67)	-0.40 (-3.02 to 2.21)	0	3 (67)	-0.55 (-7.88 to 6.78)	0
		≥ 30 kg/m^2	3 (52)	-4.07 (-6.46 to -1.67) ***	0	2 (27)	13.88 (-16.21 to 43.97)	61
	Intervention	Aerobic exercise	5 (98)	-1.88 (-4.08 to 0.32)	38	4 (73)	-1.54 (-8.26 to 5.17)	0
		Resistance exercise	3 (50)	0.62 (-1.27 to 2.51)	0	3 (50)	3.99 (-9.39 to 17.36)	50
	Duration	≤12 weeks	6 (125)	-1.23 (-3.45 to 0.98)	44	5 (100)	-0.59 (-5.10 to 3.92)	0
	Format	Supervised	5 (101)	-1.61 (-4.21 to 0.99)	49	4 (76)	-1.00 (-5.72 to 3.72)	0
		Mixed Delivery	2 (38)	0.26 (-3.22 to 3.74)	0	2 (38)	11.85 (-21.86 to 45.56)	71
Waist Circumference (cm)	BMI at entry	25-29.9 kg/m^2	3 (137)	-2.21 (-4.25 to -0.16) *	0	3 (137)	-2.02 (-3.39 to -0.65) **	0
		≥ 30 kg/m^2	3 (54)	-4.18 (-7.86 to -0.50) *	62	3 (54)	1.38 (-14.27 to 17.04)	69
	Intervention	Aerobic exercise	5 (170)	-3.30 (-6.10 to -0.51) *	50	5 (170)	-2.22 (-3.56 to -0.87) ***	0

Outcome	Sub-analysis	Sub-group	Change from baseline			Post-intervention		
			Trials (*N*)	Effect Estimate MD (95% CI)	I^2 (%)	Trials (*N*)	Effect Estimate MD (95% CI)	I^2 (%)
Body Fat (%)		Resistance exercise	2 (30)	-2.40 (-4.04 to -0.75) **	0	2 (30)	10.31 (-13.73 to 34.35)	62
	Duration	≤12 weeks	5 (180)	-1.69 (-2.38 to -0.99) ***	0	5 (180)	-1.73 (-4.25 to 0.78)	8
		>12 weeks	2 (41)	-5.19 (-11.43 to 1.05)	52	2 (41)	-6.86 (-14.02 to 0.30)	0
	Format	Supervised	5 (183)	-3.21 (-5.56 to -0.85) **	64	5 (183)	-2.16 (-3.50 to -0.82) **	0
		Mixed Delivery	2 (38)	-2.09 (-4.36 to 0.19)	28	2 (38)	8.80 (-17.70 to 35.29)	72
	BMI at entry	25-29.9 kg/m^2	2 (47)	-1.60 (-3.68 to 0.47)	59	2 (47)	-4.51 (-8.10 to -0.92) **	0
	Intervention	Aerobic exercise	2 (39)	-1.36 (-3.73 to 1.01)	76	2 (39)	-4.99 (-8.73 to -1.25) **	0
		Resistance exercise	2 (30)	-0.95 (-2.02 to 0.13)	0	2 (30)	0.47 (-5.95 to 6.88)	0
	Format	Mixed Delivery	2 (38)	-0.81 (-2.03 to 0.42)	0	2 (38)	-0.88 (-6.32 to 4.56)	0

Key: Outcome: outcome where sub-analysis was completed. Sub-analysis: how the studies were categorised for analysis; BMI at entry: mean baseline BMI of participants; Intervention: type of intervention; Duration: length of exercise intervention; Format: delivery mode of intervention. Sub-group: groups each study was classified into. Trials: number of studies included within sub-analysis, *N*: number or participants included within sub-analysis. Effect estimates are reported as mean difference (MD), and 95% confidence intervals, between exercise and control groups. Significant evidence of effect denoted by: * $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$. Heterogeneity reported using I^2 statistic: 0-40% might not be important; 30-60% may represent moderate heterogeneity; 50-90% may represent substantial heterogeneity; 75-100% may represent considerable heterogeneity. ▲: positive values favour exercise over control. VO_2 max/peak: relative maximal/peak oxygen uptake; BMI: body mass index.

3.5.10. Body Mass Index and Body Mass

Ten trials were included in the meta-analysis of BMI (Almenning *et al.*, 2015; Nasrekani *et al.*, 2016; Roessler *et al.*, 2013; Sá *et al.*, 2015; Saremi *et al.*, 2013; Saremi and Yaghoubi, 2016; Stener-Victorin *et al.*, 2009; Turan *et al.*, 2015; Vigorito *et al.*, 2007; Vizza *et al.*, 2016). When compared with control, there was a statistical effect of exercise on BMI post-intervention (MD: -1.02 kg/m^2, 95% CI: -1.81 to -0.23; 10 trials, 272 participants, I^2 = 0%), but not change from baseline values (Table 3.3). When sensitivity analyses were completed, trials with a high risk of bias were removed and a post-intervention effect remained (MD: -0.95 kg/m^2, 95% CI: -1.78 to -0.12; 6 trials, 207 participants, I^2 = 0%), but there was no longer an effect when trials with a small number of participants were removed.

Subgroup analysis (Table 3.6) revealed a statistical reduction in BMI change scores with exercise in studies that involved participants with BMI ≥30 kg/m^2 (MD: -1.34 kg/m^2, 95% CI: -1.86 to -0.82; 79 participants, 4 trials; I^2 = 0%). Change from baseline analysis also revealed a statistical decrease in participant BMI when aerobic exercise interventions were used (MD: -0.78 kg/m^2, 95% CI: -1.38 to -0.18; 260 participants, 8 trials; I^2 = 57%). In contrast, a statistical increase was observed when studies incorporated resistance training interventions (MD: 0.50 kg/m^2, 95% CI: 0.00 to 1.00; 50 participants, 3 trials; I^2 = 0%). Furthermore, post-intervention subgroup analyses revealed reductions in BMI for aerobic exercise-based (MD: -1.45 kg/m^2, 95% CI: -2.59 to -0.32; 201 participants, 7 trials; I^2 = 0%), and supervised interventions (MD: -1.06 kg/m^2, 95% CI: -1.87 to -0.25; 223 participants, 7 trials; I^2 = 0%), as well as those that were ≤12 weeks in duration (MD: -0.91 kg/m^2, 95% CI: -1.73 to -0.08; 220 participants, 7 trials; I^2 = 0%), and also in trials where participants had a BMI of 25-29.9 kg/m^2 at study entry (MD: -1.04 kg/m^2, 95% CI: -1.89 to -0.19; 188 participants, 6 trials; I^2 = 0%).

Follow-up reporting (16 weeks post-intervention) of BMI from one trial (Stener-Victorin *et al.*, 2009) showed no statistically significant within-group changes or between-group differences in either exercise or control arms. Stener-Victorin *et al.* (2009) also reported similar findings immediately post-intervention.

When comparisons of body mass change from baseline to immediately post-intervention, and post-intervention only values were made, the meta-analysis revealed no statistical effects of exercise versus control (Table 3.3). However, when body mass subgroup analysis was completed, statistical effects of exercise versus control for change from baseline values were observed for studies involving participants with BMI ≥30 kg/m^2 (MD: -4.07 kg, 95% CI: -6.46 to -1.67; 52 participants, 3 trials; I^2 = 0%). Further body mass analyses revealed no additional statistical effects in any other subgroup (Table 3.6).

3.5.11. Waist and hip circumference and waist-to-hip ratio

Seven trials reported waist circumference (WC) change values (Almenning *et al.*, 2015; Roessler *et al.*, 2013; Sá *et al.*, 2015; Saremi *et al.*, 2013; Turan *et al.*, 2015; Vigorito *et al.*, 2007; Vizza *et al.*, 2016), and when their findings were pooled (Table 3.3), a statistically significant beneficial effect of exercise compared with controls was revealed (MD: -2.62 cm, 95% CI: -4.13 to -1.11; 7 trials, 221 participants, I^2 = 53%). In contrast, there were no observed effects in the analysis for post-intervention WC values. The favourable effect of exercise on WC change values remained when trials with a low risk of bias (MD: -1.51 cm, 95% CI: -2.26 to -0.76; 167 participants, 4 trials, I^2 = 0%), or with larger sample sizes (MD: -1.48 cm, 95% CI: -2.26 to -0.71; 120 participants, 2 trials, I^2 = 0%) were analysed separately in the sensitivity analysis. Furthermore, in the investigation of heterogeneity, when the largest study outlier (Sá *et al.*, 2015) was removed from this analysis, the I^2 was reduced to 0%, whilst an effect remained (MD: -1.68 cm, 95% CI: -2.38 to -0.99).

In subgroup analyses for WC change (Table 3.6), exercise had a statistically beneficial effect in studies involving participants with a BMI of 25-29.9 kg/m^2 and ≥30 kg/m^2 (MD: -2.21 cm, 95% CI: -4.25 to -0.16; 137 participants, 3 trials, I^2 = 0%; and MD: -4.18 cm, 95% CI: -7.86 to -0.50; 54 participants, 3 trials, I^2 = 62%, respectively). Beneficial effects were also observed when interventions had a duration of ≤ 12 weeks (MD: -1.69 cm, 95% CI: -2.38 to -0.99; 180 participants, 5 trials, I^2 = 0%), were aerobic or resistance-based (MD: -3.30 cm, 95% CI: -6.10 to -0.51; 170 participants, 5 trials, I^2 = 50%; and MD: -2.40 cm, 95% CI: -4.04 to -0.75; 30 participants, 2 trials, I^2

= 0%, respectively), and were supervised (MD: -3.21 cm, 95% CI: -5.56 to -0.85; 183 participants, 5 trials, I^2 = 64%). Subgroup analysis of post-intervention WC revealed statistically lower values in exercise interventions that involved aerobic exercise (MD: -2.22 cm, 95% CI: -3.56 to -0.87; 170 participants, 5 trials, I^2 = 0%), or were supervised (MD: -2.16 cm, 95% CI: -3.50 to -0.82; 183 participants, 5 trials, I^2 = 0%), and for those where participants had a BMI of 25-29.9 kg/m^2 at study entry (MD: -2.02 cm, 95% CI: -3.39 to -0.65; 137 participants, 3 trials, I^2 = 0%).

Data from two trials (Stener-Victorin *et al.,* 2009; Vigorito *et al.,* 2007) were pooled in the analysis of waist-to-hip ratio (WHR); there was no effect in either change from baseline or post-intervention values analyses (Table 3.3).

3.5.12. Body composition

The pooled MD for change from baseline of body fat percentage (Table 3.3) showed a favourable statistical effect (MD: -1.39%, 95% CI: -2.61 to -0.18; 3 trials, 60 participants, I^2 = 30%), but this was not apparent for comparison of post-intervention values (Almenning *et al.,* 2015; Saremi *et al.,* 2013; Vizza *et al.,* 2016). Trials that were judged to have a high risk of bias were removed, resulting in a disappearance of any statistical effect. There was an insufficient number of large studies for this sensitivity analysis to be performed. Moreover, no effect of exercise versus control for change from baseline, or post-intervention analysis was found for fat mass or fat-free mass outcomes (Table 3.3).

A statistical effect of exercise was observed for body fat percentage change only in interventions lasting ≤ 12 weeks; the main analysis included only trials ≤ 12 weeks in duration (Table 3.3), meaning results of this sub analysis were identical. No other statistical effects were found across any of the other subgroup analyses on body fat percentage change (Table 3.6). However, body fat percentage was statistically lower post-intervention in exercise interventions that included participants with BMI of 25-29.9 kg/m^2 (MD: -4.51%, 95% CI: -8.10 to -0.92; 47 participants, 2 trials, I^2 = 0%), or that utilised aerobic exercise (MD: -4.99%, 95% CI: -8.73 to -1.25; 39 participants, 2 trials, I^2 = 0%). There was no evidence of effect in the subgroup analyses for fat mass or fat-free mass.

3.5.13. Androgenic, hormonal, and inflammatory markers

In pooled analyses of change from baseline and post-intervention values, no statistical effects (either beneficial or harmful) of exercise were found for any of the androgenic/hormonal and biomarkers/variables that are associated with inflammation [*i.e.*, testosterone, free testosterone, FAI, SHBG, Ferriman-Gallwey scores, oestradiol, LH, FSH, LH/FSH ratio, progesterone, prolactin, high-sensitivity C-reactive protein (*hs*CRP), anti-Mullerian hormone (AMH), or adiponectin] when compared with control (Table 3.3). Similarly, no statistical effects were observed for any subgroup in any of these outcomes (Table 3.7).

Table 3.7. Summary of effect estimates and heterogeneity from sub-group analyses in androgenic and inflammatory related outcomes.

Outcome	Sub-analysis	Sub-group	Change from baseline			Post-intervention		
			Trials (*N*)	Effect Estimate MD (95% CI)	I^2 (%)	Trials (*N*)	Effect Estimate MD (95% CI)	I^2 (%)
Total Testosterone (nmol/L)	BMI at entry	25-29.9 kg/m^2	3 (160)	-0.11 (-0.27 to 0.05)	0	3 (126)	-0.03 (-0.46 to 0.41)	68
	Intervention	Aerobic exercise	3 (152)	-0.10 (-0.27 to 0.06)	0	3 (118)	-0.03 (-0.48 to 0.43)	60
		Resistance exercise	2 (30)	0.00 (-0.30 to 0.30)	0	2 (30)	0.04 (-0.29 to 0.36)	0
	Duration	≤12 weeks	4 (158)	-0.02 (-0.23 to 0.19)	0	4 (158)	-0.08 (-0.40 to 0.25)	50
	Format	Supervised	2 (120)	-0.10 (-0.41 to 0.21)	0	2 (120)	-0.30 (-0.51 to -0.09)	0
		Mixed Delivery	2 (38)	0.04 (-0.24 to 0.32)	0	2 (38)	0.11 (-0.33 to 0.55)	45
SHBG▲ (nmol/L)	BMI at entry	25-29.9 kg/m^2	3 (160)	11.16 (-8.39 to 30.71)	92	3 (126)	14.99 (-18.49 to 48.47)	0
	Intervention	Aerobic exercise	3 (152)	9.49 (-13.77 to 32.76)	92	3 (118)	18.97 (-23.25 to 61.19)	69
		Resistance exercise	2 (30)	4.98 (-14.52 to 24.49)	68	2 (30)	-0.79 (-45.26 to 43.67)	67
	Duration	≤12 weeks	3 (128)	2.45 (-1.04 to 5.93)	0	3 (128)	-0.81 (-8.06 to 6.45)	22
	Format	Mixed Delivery	2 (38)	-3.23 (-14.91 to 8.46)	0	2 (38)	4.51 (-56.46 to 85.47)	89
Free Androgen Index	BMI at entry	25-29.9 kg/m^2	3 (126)	0.09 (-0.78 to 0.96)	0	3 (126)	0.06 (-1.13 to 1.26)	10
	Intervention	Aerobic exercise	3 (118)	0.51 (-0.52 to 1.53)	0	3 (118)	0.10 (-1.17 to 1.36)	10
		Resistance exercise	2 (30)	-0.04 (-1.67 to 1.58)	57	2 (30)	1.71 (-3.65 to 7.08)	74
	Duration	≤12 weeks	3 (128)	0.11 (-0.71 to 0.93)	0	3 (128)	0.34 (-1.45 to 2.13)	50
	Format	Mixed Delivery	2 (38)	1.79 (-3.18 to 6.76)	72	2 (38)	1.67 (-3.80 to 7.14)	75
Oestradiol (pmol/L)	BMI at entry	25-29.9 kg/m^2	2 (135)	-41.14 (-141.68 to 59.40)	70	-	-	-
	Intervention	Aerobic exercise	3 (160)	-47.22 (-117.52 to 23.08)	65	-	-	-
	Duration	≤12 weeks	3 (145)	-1.14 (-36.61 to 34.33)	61	2 (120)	0.27 (-11.27 to 11.80)	0
	Format	Supervised	3 (145)	-1.14 (-36.61 to 34.33)	61	2 (120)	0.27 (-11.27 to 11.80)	0
Luteinising Hormone (IU/L)	BMI at entry	25-29.9 kg/m^2	3 (155)	-1.12 (-3.63 to 1.39)	61	3 (121)	0.15 (-0.92 to 1.22)	0
	Intervention	Aerobic exercise	3 (155)	-1.12 (-3.63 to 1.39)	61	3 (121)	0.15 (-0.92 to 1.22)	0
	Duration	≤12 weeks	3 (140)	-0.59 (-3.24 to 2.06)	81	3 (140)	-1.60 (-4.73 to 1.54)	62
	Format	Supervised	3 (140)	-0.59 (-3.24 to 2.06)	81	3 (140)	-1.60 (-4.73 to 1.54)	62
Follicle Stimulating Hormone▲ (IU/L)	BMI at entry	25-29.9 kg/m^2	3 (155)	0.09 (-0.35 to 0.52)	0	3 (121)	0.11 (-0.35 to 0.56)	0
	Intervention	Aerobic exercise	3 (155)	0.09 (-0.35 to 0.52)	0	3 (121)	0.11 (-0.35 to 0.56)	0
	Duration	≤12 weeks	3 (140)	0.19 (-0.13 to 0.51)	0	3 (140)	-0.03 (-0.42 to 0.37)	0
	Format	Supervised	3 (140)	0.19 (-0.13 to 0.51)	0	3 (140)	-0.03 (-0.42 to 0.37)	0

Key: Outcome: outcome where sub-analysis was completed. Sub-analysis: how the studies were categorised for analysis; BMI at entry: mean baseline BMI of participants; Intervention: type of intervention; Duration: length of exercise intervention; Format: delivery mode of intervention. Sub-group: groups each study was classified into. Trials: number of studies included within sub-analysis, *N*: number or participants included within sub-analysis. Effect estimates are reported as mean difference (MD), and 95% confidence intervals, between exercise and control groups. Significant evidence of effect denoted by: * $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$. Heterogeneity reported using I^2 statistic: 0-40% might not be important; 30-60% may represent moderate heterogeneity; 50-90% may represent substantial heterogeneity; 75-100% may represent considerable heterogeneity. ▲: positive values favour exercise over control.; SHBG: sex hormone-binding globulin.

3.5.14. Psychosocial Outcomes

In two trials (57 participants) that assessed psychosocial outcomes using the PCOS-Q, we found no effect of exercise on any PCOS-Q domain compared with control. Three trials (84 participants) used the SF-36 (Sá *et al.*, 2015; Stener-Victorin *et al.*, 2009; Vizza *et al.*, 2016). Data only allowed for change from baseline analysis and no sub-analysis was possible. For SF-36 domains, a favourable effect of exercise versus control was found for physical functioning (MD: 11.81, 95% CI: 2.36 to 21.25; I^2 = 74%), general health (MD: 10.05, 95% CI: 3.89 to 16.20; I^2 = 0%), social functioning (MD: 11.75, 95% CI: 2.56 to 20.95; I^2 = 6%), and mental health (MD: 11.70, 95% CI: 1.27 to 22.13; I^2 = 47%) domains (Figure 3.9). When the change values for SF-36 total score (*i.e.* all domain scores combined) were analysed, a statistical effect was evident (MD: 7.78, 95% CI: 4.41 to 11.14; I^2 = 50%).

There were insufficient data to complete sensitivity analyses; however, all three trials were judged to have a high risk of bias in at least one domain (Figure 3.9), and only one trial had a sample size ≥ 30 (Vizza *et al.*, 2016). Heterogeneity was investigated in the physical functioning domain; the largest outlier was removed (Sá *et al.*, 2015) and the I^2 was reduced to 33%, whilst an effect was maintained (MD: 7.23, 95% CI: 1.66 to 12.80). The same trial was removed in the general health analysis, resulting in a reduction in I^2 to 0%, and a preserved effect (MD: 7.97, 95% CI: 1.07 to 4.88). When the greatest outliers were removed from the social functioning (Vizza *et al.*, 2016) and mental health (Stener-Victorin *et al.*, 2009) domains, both I^2 values were reduced to 0%, but the effect only remained in the mental health domain (MD: 17.84, 95% CI: 7.33 to 28.36).

3.5.15. Additional Outcomes

Six trials (Almenning *et al.*, 2015; Sá *et al.*, 2015; Stener-Victorin *et al.*, 2009; Turan *et al.*, 2015; Vigorito *et al.*, 2007; Vizza *et al.*, 2016) also reported a range of additional outcomes; these outcomes were unique to these studies and so they could not be included in the meta-analysis. Key findings from these trials are presented in Table 3.8.

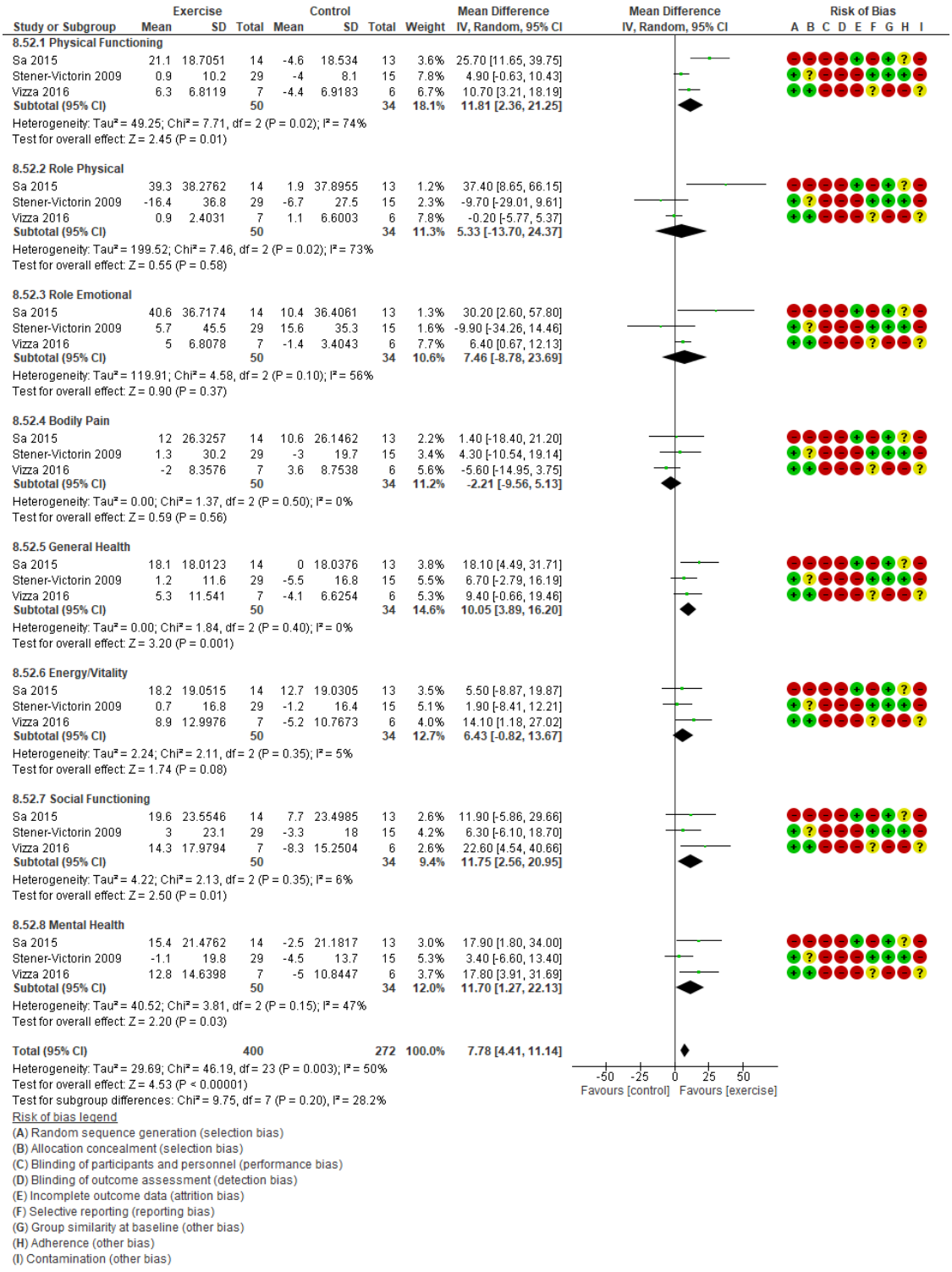


Figure 3.9. Forest plot of comparison: exercise vs. control, change from baseline; outcome: SF-36 Domains.

Table 3.8. Exercise versus control: summary of findings regarding study outcomes that were reported only by a corresponding single trial.

Trial	Statistical significance	Outcomes
Almenning *et al*., 2015	No statistically significant findings	HR recovery; Leptin
	Statistically significant findings	➢ Nitric oxide bio-availability [as measured by FMD %] - reported a statistically significant improvement in FMD following a HIIT intervention, but not resistance training ➢ Homocysteine – significant change from baseline concentrations (MD: -0.6 µmol/L, 95% CI: -1.0 to -0.1) following a HIIT intervention, but no between group differences
Sá *et al.,* 2015	No statistically significant findings	AUC oral glucose tolerance test; two-hour post-prandial blood glucose; interleukin-6; tumour necrosis-α
	Statistically significant findings	➢ Mean arterial blood pressure; a statistical reduction in the exercise group (MD: -6.8 mmHg, 95% CI: -10.6 to -3.0; 14 participants) and a significant time by group interaction, representative of a moderate effect size (*d*) was also reported (*d:* -1.0; 95% CI: -1.7 to -0.3)
Stener-Victorin *et al.*, 2009	No statistically significant findings	➢ Number of participants with acne; menstrual frequency; 5α-dihydrotestosterone; estrone; DHEA; androstenedione; 5-androstene-3β, 17β-diol, androsterone glucuronide; androstane-3α, 17β-diol-3 glucuronide; 17β-diol-17 glucuronide; insulin growth factor-1; thyroid stimulating hormone; free thyroxin 4; fibrinogen; fibrin ɒ-dimer; von Willebrand factor; factor VIII; tissue plasminogen activator and plasminogen activator inhibitor; ovarian volume; Montgomery Asberg Depression Rating Scale and the Brief Scale for Anxiety
	Statistically significant findings	➢ Estrone sulfate (E1-S) was significantly lower (P <.05) in the exercise group versus control when measured immediately post-intervention; this effect disappeared during follow-up assessment ➢ Median antral follicle counts were significantly lower (-11.7%; P = .010) from baseline to follow-up in the exercise group

Trial	Statistical significance	Outcomes
Turan *et al.*, 2015	Statistically significant findings	➢ Respiratory rate: significant within group changes in respiratory rate (-1.0±0.4 breaths/minute) following exercise, but no between group differences. ➢ Hip circumference: a statistically significant (P <.05) reduction following exercise training, and a statistical difference (P <.05) between change values in each arm
Vigorito et al., 2007	No statistically significant findings	➢ Respiratory exchange ratio; Peak HR
	Statistically significant findings	➢ AUC-insulin and AUC-glucose to AUC-insulin ratio: a statistically significant change from baseline for AUC-insulin and the ratio with AUC-glucose in the exercise group but not in the control. AUC-insulin significantly improved (*P* <.001) compared to the control group ➢ VO_2 at anaerobic threshold: within and between group statistical changes for VO_2 at anaerobic threshold (MD: 4.4 ml/kg/min; *P* <.001) ➢ Maximum workload: within and between group statistical changes (MD: 32.3 Watts; *P* <.001) ➢ Ventilatory equivalent for carbon dioxide production: only within group changes were reported (VE/VCO_2: MD: -0.6; *P* = .01) ➢ Participant leisure time physical activity (MET-hrs/wk): significantly higher (*P* <.001) following an exercise intervention
Vizza *et al*., 2016	Statistically significant findings	➢ Glycated haemoglobin (HbA1c): Resistance training statistically reduced (*P* = .037) within group HbA1c, and when compared with control (*P* = .03, *d*= 0.39). ➢ Lower, but not upper, body strength was significantly increased (*P* = .04) following a resistance training intervention; it was also significantly improved compared to a control (ES: 0.45; *P* = .03) ➢ Depression, Anxiety and Stress Scale 21 (DASS-21): the depression domain showed within group (*P* = .05) and between group (ES: 0.50; *P* = .01) reductions following resistance training

Trial	Statistical significance	Outcomes
		➢ Exercise Self Efficacy Scale: a statistically significant reduction (P = .04) of self-efficacy within the control group, but no changes in the exercise groups or differences between groups

Key: HR: heart rate; FMD: flow mediated dilation; HIIT: high-intensity interval training; MD: mean difference; DHEA: dehydroepiandrosterone; AUC: area under the curve; VO_2: volume of oxygen; VE/VCO_2: minute ventilation/carbon dioxide production; MET: metabolic equivalent of task

3.6. Effects of interventions: exercise and diet vs. control

Three included trials made the comparison of exercise and diet combined against a control/usual care group. Only one of these trials used a control group that was described as no treatment (Guzick *et al.*, 1994), whereas the other two (Hoeger *et al.*, 2004; Petrányi and Zaoura-Petrányi, 2011) compared exercise, diet and metformin (or placebo) to metformin only groups. As the pharmacological component of the intervention was present in each treatment arm, the assumption was made that any between group variations were attributable to the exercise and dietary components.

Due to insufficient data, only two outcomes (WHR and SHBG) were eligible to be included in the meta-analysis. Meta-analysis of the two trials (68 participants) reporting change from baseline to post-intervention WHR values (Guzick *et al.*, 1994; Petrányi and Zaoura-Petrányi, 2011) revealed a small statistical favourable effect of exercise and diet (MD: -0.02, 95% CI: -0.03 to -0.01; $I^2 = 0\%$) compared to control (Figure 3.10). The effect was not replicated in the comparison of post-intervention values.

Furthermore, no effect of exercise and diet combined versus control was found for the change from baseline to post-intervention circulating SHBG levels. There were insufficient data to complete analysis of post-intervention values or subgroups. A range of individual outcomes were also reported by each of these trials, which are summarised in Table 3.9.

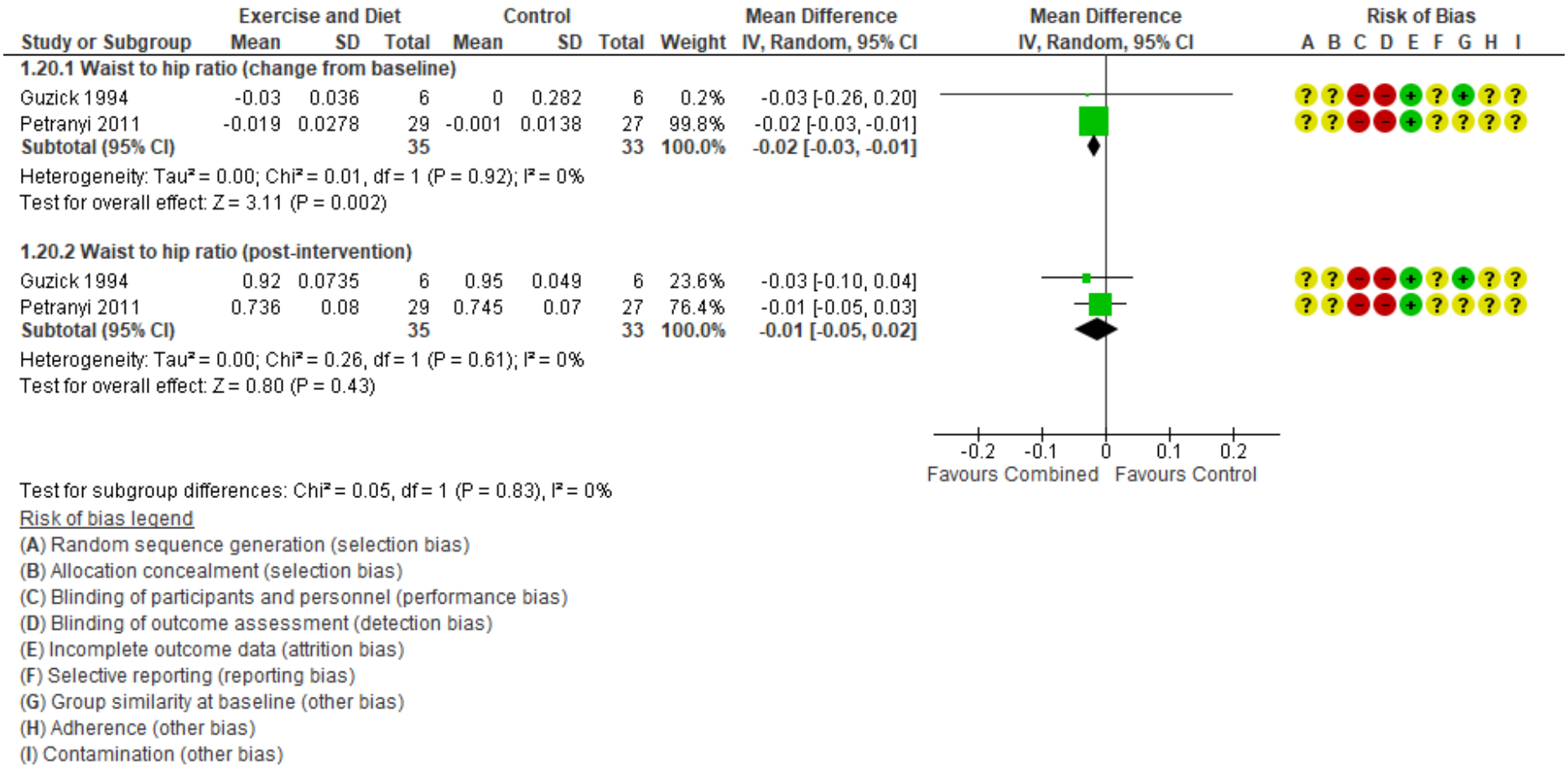


Figure 3.10. Forest plot of comparison: exercise and diet vs. control; outcome: waist to hip ratio.

Table 3.9. Combined exercise and diet versus control: summary of findings from investigative outcomes that were reported only in single trials.

Trial	Statistical significance	Outcomes
Guzick *et al*., 1994	No statistically significant findings	➢ Fasting insulin; luteinising hormone; follicle stimulating hormone.
	Statistically significant findings	➢ Body weight: statistical interaction effect (P <.0001) reflecting an improvement following combined exercise and diet intervention, but not control ➢ Free testosterone: statistical interaction effect (P =.02) following a combined exercise and diet intervention, but not control
Hoeger *et al*., 2004	No statistically significant findings	➢ Free androgen index; AUC-glucose; AUC-insulin; fasting blood glucose; ovulatory status
	Statistically significant findings	➢ Body weight: statistically significant (P <.05) within-group body weight reductions for lifestyle and placebo, but no statistical differences versus placebo alone ➢ When lifestyle was combined with Metformin, statistical differences (P <.05) compared to placebo only were reported for body weight, SHBG and FAI
Petrányi *et al*., 2011	Statistically significant findings	➢ Statistically significant (P <.001) reductions in levels of acne, FG scores and BMI following lifestyle and Metformin therapy; changes in the Metformin only arm were comparable apart for that for BMI, which was statistically higher in the combined treatment (P = .03)

Key: AUC: area under the curve; SHBG: sex hormone binding globulin; FAI: free androgen index; FG: Ferriman-Gallwey; BMI: body mass index

3.7. Effects of interventions: exercise and diet vs. diet

Three trials had intervention arms that compared the combination of exercise and diet to diet only (Bruner *et al.*, 2006; Nybacka *et al.*, 2011; Thomson *et al.*, 2008). For any assessed primary outcome (Table 3.10: FBG, FI, and HOMA-IR; all very low-quality evidence) or secondary outcome (body weight, BMI, WC, body fat, fat-free mass, testosterone, SHBG and FAI) of these studies, analyses of change from baseline and post-intervention values (Table 3.11) revealed no statistical difference between the two interventions (combined exercise and diet vs diet only). Furthermore, there were insufficient data to complete subgroup analyses within this comparison.

All these three included trials reported a range of other outcomes which were not eligible to be included in this meta-analysis and are summarised in Table 3.12.

Table 3.10. Summary of findings for primary outcomes: exercise and diet versus diet only.

Intervention: Exercise and diet; **Comparison**: Diet

Outcomes	**Anticipated absolute effects*** (95% CI)		№ of participants (№ of studies)	Certainty of the evidence (GRADE)	Comments
	Risk with Diet	**Risk with exercise and diet**			
Fasting blood glucose (change from baseline) follow up: range 16 weeks to 20 weeks	The mean fasting blood glucose (change from baseline) ranged from **-7.0 to -3.2 mg/dL**	The mean fasting blood glucose (change from baseline) in the intervention group was 2.92 mg/dL higher (0.4 lower to 6.23 higher)	78 (2 RCTs)	⨁◯◯◯ VERY LOW [a,b]	We are uncertain about the effect of exercise and diet on fasting blood glucose (change from baseline).
Fasting insulin (change from baseline) follow up: range 12 weeks to 20 weeks	The mean fasting insulin (change from baseline) ranged from **-2.9 to -18.54 µU/ml**	The mean fasting insulin (change from baseline) in the intervention group was 2.22 µU/ml higher (3.7 lower to 8.14 higher)	90 (3 RCTs)	⨁◯◯◯ VERY LOW [a,c,d]	We are uncertain about the effect of exercise and diet on fasting insulin (change from baseline).
HOMA-IR (change from baseline) follow up: range 16 weeks to 20 weeks	The mean HOMA-IR (change from baseline) ranged from **-0.74 to -0.56**	The mean HOMA-IR (change from baseline) in the intervention group was 0.01 lower (0.45 lower to 0.43 higher)	78 (2 RCTs)	⨁◯◯◯ VERY LOW [a,b]	We are uncertain about the effect of exercise and diet on HOMA-IR (change from baseline).

***The risk in the intervention group** (and its 95% CI) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI). **CI:** Confidence interval; **MD:** Mean difference

GRADE Working Group grades of evidence

- **High certainty:** We are very confident that the true effect lies close to that of the estimate of the effect
- **Moderate certainty:** We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
- **Low certainty:** Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
- **Very low certainty:** We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Explanations

a) All trials were at an unclear risk of selection bias, reporting bias, contamination, and adherence issues. All trials were at a high risk of detection bias and attrition bias. Therefore, we downgraded by one level.
b) Small number of participants, only two trials, and wide confidence intervals in the included trials. Therefore, we downgraded by two levels.
c) Substantial heterogeneity was observed. Therefore, we downgraded by one level.
d) Small number of participants and trials, wide confidence intervals, and null/negligible effect and appreciable benefit included in the confidence interval for the mean difference. Therefore, we downgraded by two levels.

Table 3.11. Effect estimates and heterogeneity for change from baseline to immediately post-intervention, and immediately post-intervention values only, for all outcomes analysed in the comparison between combined exercise and diet versus diet only.

	Change from baseline						Immediately post-intervention					
Outcome	**Trials**	***N***	**MD**	**95% CI**		**I^2 (%)**	**Trials**	***N***	**MD**	**95% CI**		**I^2 (%)**
				Lower	**Upper**					**Lower**	**Upper**	
FBG (mg/dL)	2	78	2.92	-0.40	6.23	42	2	78	2.86	-1.56	7.29	0
FI (µIU/mL)	3	90	2.22	-3.70	8.14	62	2	64	-2.72	-7.70	2.27	0
HOMA-IR	2	78	-0.01	-0.45	0.43	0	-	-	-	-	-	-
Body Weight (kg)	2	64	-0.40	-3.64	2.83	0	2	64	1.49	-8.05	11.03	0
BMI (kg/m^2)	2	38	-0.09	-1.27	1.09	0	2	38	2.56	-1.77	6.88	0
WC (cm)	2	64	-0.47	-3.95	3.01	0	2	64	-1.51	-8.69	5.67	0
Body Fat (%)	2	78	-1.05	-4.61	2.50	85	2	78	-0.93	-3.63	1.77	10
FFM (kg) ▲	2	78	0.40	-3.24	4.03	85	2	78	2.07	-1.72	5.86	0
Testosterone (nmol/L)	3	90	0.29	-0.49	1.08	78	3	90	0.08	-0.38	0.54	0
SHBG (nmol/L) ▲	3	90	2.18	-3.15	7.51	51	3	90	6.45	-5.52	18.42	61
FAI	2	64	0.11	-2.28	2.50	0	2	64	-2.88	-6.58	0.81	0

Key: Negative values favour exercise and diet combined except where stated otherwise. ▲: positive values favour exercise and diet combined over diet only. Trials: number of studies included within analysis, *N:* number or participants included within analysis. Effect estimates are reported as mean differences (MD), and 95% confidence intervals (CI), between exercise and diet combined vs diet only groups. Heterogeneity reported using I^2 statistic. FBG: fasting blood glucose; FI: fasting insulin; HOMA-IR: homeostatic model of assessment, insulin resistance; BMI: body mass index; WC: waist circumference; FFM: fat free mass; SHBG: sex hormone binding globulin; FAI: free androgen index.

Table 3.12. Combined exercise and diet versus diet: summary of findings from study outcomes that were reported only in single trials.

Trial	Statistical significance	Outcomes
Bruner *et al.*, 2006	No statistically significant findings	➢ Resting energy expenditure; LH/FSH ratio; number of ovarian follicles (left and right)
	Statistically significant findings	➢ Sum of two skinfolds (subscapular and iliac crest): statistically lower than at baseline and a group x time interaction (*P*= .002) immediately post-intervention with a greater decrease in the combined exercise and diet group compared with the diet only group
Nybacka *et al.*, 2011	No statistically significant findings	➢ No significant changes seen in any intervention arm for ratio of upper/lower body fat. No effect seen in upper body fat (kg) for diet only or diet and exercise combined; no reduction in lower body fat for the exercise only arm ➢ Exercise and diet combined did not significantly reduce IGF-I or IGFBP-1
	Statistically significant findings	➢ In the diet only arm statistical changes in free testosterone (-3.66 pg/mL, 95% CI: -6.12 to -1.20; *P* <.001), AMH (*P* <.01), IGF-1 (17.1 µg/L, 95% CI: 0.3 to 33.9; *P* <.05), and IGFBP-1 (0.32 µg/L, 95% CI: 0.01 to 0.64; *P* <.05) were reported that were not present in the combined diet and exercise arm ➢ There were statistically significant reductions in: lower body fat for the diet only arm (-1055g, 95% CI: -1787 to -322; *P* <.01) and the combined diet and exercise arm (1616g, 95% CI: -2407 to -825; *P* <.001); lean body mass only in the combined diet and exercise arm (-2.66kg, 95% CI: -4.14 to -1.18; *P* <.001); mean ovarian follicle number in both diet only (*P* <.05) and the combined arm (*P* <.05); as well as improvements to ovulatory function in both intervention arms (diet: *P* <.001; combined: *P* <.05) ➢ Mean ovarian volume was reduced only in the combined diet and exercise arm (*P* <.05)
Thomson *et al.*, 2008	No statistically significant findings	The Centre of Epidemiologic Studies Depression Scale was also used, but there were no differences in post-intervention scores compared to baseline
	Statistically significant findings	➢ Statistically significant reductions (*P* ≤.03) to fat mass and abdominal fat mass in all groups; both exercise arms were also statistically different (P ≤.03) to the diet only arm.

Trial	Statistical significance	Outcomes
		➢ Levels of endothelial function were also measured; vascular cell adhesion molecule-I (*P*= .01), plasminogen activator inhibitor-I (*P* <.001) and intra-cellular adhesion molecule-I (*P* <.001) were reduced in all treatment arms with no statistical differences between treatments ➢ PCOS-Q was used to assess quality of life, showing statistical improvements (*P* ≤.001) across all treatment arms in each domain apart from body hair scores, while no differences were found between treatment arms

Key: LH/FSH: luteinising hormone/follicle stimulating hormone; IGF-1: insulin-like growth factor-1; IGFBP-1; insulin-like growth factor binding protein-1; AMH: anti-Müllerian hormone; PCOS-Q; polycystic ovary syndrome questionnaire.

3.8. Effects of interventions: Exercise vs. diet, and exercise and diet vs exercise.

A relevant meta-analysis was not possible as only one trial (Nybacka *et al.*, 2011) compared exercise with diet, and exercise combined with diet to exercise alone. Effects from this study for the diet only and combined diet and exercise groups have been reported in Section 3.6 and in Table 3.12. In addition to these findings, the exercise only intervention reduced BMI (-0.85 kg/m^2, 95% CI: -1.69 to -0.02; P <.05), but the magnitude of these changes was smaller than those seen in the other treatment arms (*i.e.* diet only, or combined diet and exercise). Upper body fat was statistically reduced only in the exercise group (-1.57 kg, 95% CI: -2.86 to -0.28; P <.05) and also, mean follicle number exhibited the greatest improvement in the exercise only group (P <.01). There were no within-group effects reported for body fat (%), lower body fat (kg), lean body mass, free testosterone, insulin-like growth factor-1, insulin-like growth factor binding protein-1, FBG, FI, HOMA-IR, LH, FSH, testosterone, SHBG, T/SHBG ratio, AMH, or mean ovarian volume.

3.9. Discussion

3.9.1. Summary of Main Results

The current systematic review and meta-analysis provides up-to-date analyses of the existing published evidence regarding the impact of exercise interventions in the management of PCOS. The present analyses identified a marked paucity of well-designed, long-term and large studies in this field, whilst the existing evidence, albeit of low quality, suggests that incorporation of exercise interventions in the management of PCOS may result in relatively small benefits in certain PCOS patient groups and with certain interventions. Indeed, when exercise interventions were compared to a no-intervention/usual care control group, there were statistically beneficial effects observed for both change from baseline to immediately post-intervention and comparison of immediately post-intervention values, for fasting insulin, total cholesterol, LDL-C, and VO_2 max. The clinical relevance of these changes is unclear. Furthermore, there were statistically favourable changes from baseline (but not post-intervention comparison) for HOMA-IR, triglycerides, waist circumference, and body fat percentage, as well as statistically lower post-intervention values for BMI and resting heart rate. Whether such changes translate to clinically meaningful benefit for the long-term cardio-metabolic risk of women with PCOS (particularly for different PCOS phenotypes) remains to be studied in long-term prospective trials. Moreover, although the analysis comparing exercise and diet to a control incorporated few outcomes and a limited number of studies, there was a small statistical effect in favour of exercise and diet for waist-hip-ratio. In the combined exercise and diet versus diet only comparison, there was no evidence of intervention effect for any included outcome, but again there were strikingly scant data available. Indeed, only one study made the remaining comparisons (*i.e.* exercise versus diet, and exercise combined with diet versus exercise only) and, whilst there are some beneficial effects reported, it is difficult to draw meaningful conclusions from a single trial.

3.9.2. Primary Outcomes

A small beneficial change in SBP from baseline to post-intervention was observed with supervised exercise interventions versus control. Searches of existing literature suggest that this is the first systematic review to report on the effects of exercise on blood pressure in women with PCOS. A

previous study in the general population provides evidence suggesting that it is aerobic exercise interventions which induce the greatest improvements to blood pressure (both SBP and DBP) in hypertensive participants (Cornelissen and Smart, 2013), and that there are less marked effects in normotensive participants (*i.e.,* small decreases in DBP and no effect on SBP). The mean SBP (116 mmHg) and DBP (73 mmHg) values in the current systematic review indicate that most enrolled women with PCOS in the included studies were normotensive at baseline, which explain, at least partly, why a larger beneficial effect was not observed.

When surrogate markers of insulin resistance (FI and HOMA-IR) were considered, statistically beneficial changes from baseline (FI and HOMA-IR), and more favourable post-intervention values (FI) for exercise compared with control were observed. Subgroup analyses were also indicative that the greatest improvements appeared in women with PCOS and BMI $\geq$ 25.0 kg/m^2, and from interventions which were supervised, aerobic-based and shorter in duration. These findings are partially supported by two previous systematic reviews (Domecq *et al.,* 2013; Moran *et al.,* 2011); however, these reviews did not make the distinction between exercise, diet or a combination of the two, but instead indiscriminately combined these three terms under the heading 'lifestyle interventions' which was compared to a control group. The more recent of these systematic reviews (Domecq *et al.,* 2013) reported a small, but statistically significant effect on FI change from baseline (MD: -2.1 µIU/mL, 95% CI: -3.3 to -1.0; 5 trials, I^2 = 0%). The other review (Moran *et al.,* 2011) compared post-intervention FI values and reported that lifestyle interventions induced decreased FI values compared to control (MD: -2.02 µIU/mL, 95% CI: -3.28 to -0.77; 144 participants, 5 trials, I^2 = 0%). These previous findings have been expanded upon in the current meta-analysis because a greater number of trials and participants have been included and also by separating exercise only trials. What this has shown is that, based on the best available data, exercise alone has a comparable effect to that of lifestyle interventions, but the degree of benefit is likely trivial.

Although IR is not currently included within the PCOS diagnostic criteria, it is widely acknowledged to have a key role in the pathophysiology of PCOS (Iuorno, 2007). It has been reported that approximately 50-70% of women with PCOS have been diagnosed with IR and hyperinsulinaemia

(Marin *et al.*, 2003), whereas many also present evidence of glucose intolerance (Ng *et al.*, 2019). Alongside the potential metabolic implications, hyperinsulinaemia in PCOS also promotes further secretion of androgens from the ovarian theca cells, whilst also supressing hepatic SHBG secretion, which causes an increase in free androgens and further exacerbates the associated symptoms (Zhang *et al.*, 2018). Despite the integral role of IR in PCOS, there is scant information for FI reference values in the literature (Chevenne *et al.*, 1999). One historic study (Boyns *et al.*, 1969) reported FI levels ranging from 2-60 µIU/mL in a sample of 111 healthy women, and a mean value of 17.6 ± 5.7 µIU/mL in a subset of those women aged 25-34 years (n = 22). For women with PCOS, a large-scale case-control study of 1404 women with PCOS reported mean FI levels of 14.3 ± 1.6 µIU/mL, which was significantly higher than healthy controls (Zhang *et al.*, 2013). The mean baseline FI level of all intervention participants in the current systematic review was 16.21 µIU/mL, and a reduction of ~13% was reported following exercise interventions. Although statistical effects are reported in the current review, it is unclear whether these exercise-induced reductions are clinically meaningful because there is marked variability of normative FI values in PCOS.

Although FI correlates with IR, especially in normoglycemic populations (Olefsky *et al.*, 1973; Philips *et al.*, 1994) it has been shown that HOMA-IR [fasting glucose (nmol/L) * fasting insulin (µU/mL) / 22.5]; Matthews *et al.*, 1985) may be a better estimate of insulin sensitivity (Emoto *et al.*, 1999). In the current review, the mean baseline HOMA-IR values for intervention-group participants was 2.99, which dropped to 2.43 (MD: -0.57) following exercise, with no corresponding reduction in control groups. A generally adopted HOMA-IR cut-off value for the identification of IR is 2.6 (Ascaso *et al.*, 2003). This implies that exercise may have a significant effect on IR compared with usual care, which may result in less required insulin levels to maintain normoglycemia. Conversely, no statistical or clinical effect of exercise was observed for FBG; however, participants were within normal FBG at baseline.

In contrast to previous reviews (Moran *et al.*, 2011; Haqq *et al.*, 2015), an effect of exercise on lipidemic profiles was observed in the current systematic review. When compared to control, there were improvements in exercise-induced changes for circulating levels of total cholesterol, LDL-C

and triglycerides. Based on data included in this systematic review, the mean baseline values for total cholesterol (233 mg/dL) and LDL-C (142 mg/dL) are classified as borderline high or elevated in the presence of concomitant CVD risk factors (National Cholesterol Education Program, 1993). Post-intervention values for LDL-C were lower for exercise compared to control, but total cholesterol levels were comparable (approximately 229 mg/dL in both). LDL-C plays a key role in atherogenesis; the risk of coronary heart disease (CHD) progressively increases as LDL plasma levels increase (Sniderman *et al.,* 1997). Conversely, there is an inverse association between HDL-C and both atherosclerosis severity and CHD risk, with HDL-C levels ≥60 mg/dL having a potentially protective effect against CHD (National Institutes of Health Consensus Development Panel, 1993). Within this systematic review, baseline and post-intervention HDL-C values were >60 mg/dL, which may partially explain why there was no observed effect of exercise. However, in the current systematic review, total cholesterol and LDL-C levels were elevated at baseline, and there is evidence of a statistical effect following exercise, but the magnitude of the observed changes may not be clinically important (Puska, 2010; Wadhera *et al.,* 2016).

For the present analysis on the effect on circulating triglyceride levels, mean baseline concentrations were higher in the exercise group (+11 mg/dL) compared with control, although both groups were within a range that would be considered as normal (<150 mg/dL). That said, exercise statistically reduced triglyceride levels, but the comparison of post-intervention values revealed that concentrations were still lower in the control group. Triglycerides are independent predictors of CVD mortality in women (Bass *et al.,* 1993); however, in the present systematic review the magnitude of exercise-induced triglyceride reduction within the reported range is unlikely to be clinically important. Future research should aim to investigate the independent effect of exercise in women (with and without PCOS) with hypertriglyceridaemia.

3.9.3. Secondary Outcomes

A statistically significant and potentially clinically important effect was observed for VO_2 max (>3.5 ml/kg/min) when exercise was compared to control (Lakoski *et al.,* 2015) but this may be in part due

to weight loss. Subgroup analyses revealed that aerobic exercise, regardless of other all other assessed variables, improved VO_2 max in women with PCOS.

Low cardiorespiratory fitness, as measured by VO_2 max, is often associated with an increased risk of chronic disease and all-cause mortality (Blair *et al.,* 1989; Kodama *et al.,* 2009). Whilst reductions in cardiorespiratory fitness occur physiologically with age, being physically inactive may also contribute to a lower VO_2 max. The consequences of reduced cardiorespiratory fitness include impaired capability to exercise, reduced ability to perform activities of daily living, as well as lower overall quality of life (Donà *et al.,* 2017). Because of this, cardiorespiratory fitness improvement is a goal of many lifestyle interventions, although this is often overlooked in PCOS. Studies assessing patient VO_2 max in this female patient population are limited; one study in overweight (Orio *et al.,* 2006), and another in lean (Bacchi *et al.,* 2015) women with PCOS reveal markedly lower cardiorespiratory fitness than healthy controls. Only one previous systematic review of exercise/lifestyle and women with PCOS report on VO_2 max/peak (Haqq *et al.,* 2015), reporting improvements for both lifestyle (*i.e.* exercise and diet combined; MD: 5.09 ml/kg/min, 95% CI: 3.13 to 7.05, 3 trials, 137 participants) and exercise only (MD: 4.86 ml/kg/min, 95% CI: 2.83 to 6.88, 2 trials, 125 participants) interventions when compared with usual care. Although we noted a marginally smaller effect, the agreement between these results and the findings of the analysis of relative VO_2 max changes in the current systematic review, which combined data from 92 more participants than the review by Haqq *et al*. (2015), suggests that exercise has the potential to improve cardiorespiratory fitness in women with PCOS.

In addition, the current systematic review found slight reductions in WC and body fat percentage in the exercise groups, suggesting that exercise may stimulate small favourable changes to body composition in women with PCOS. As a measure of central/abdominal obesity, WC is reportedly a better independent predictor of obesity-related disorders than BMI (Lean *et al.,* 1995). This is potentially attributed, at least in part, to the pivotal role that central adiposity plays in the development of IR and T2DM, even in those with normal BMI (Gómez-Ambrosi *et al.,* 2011). However, despite statistical significance, the exercise-induced WC changes observed in this

systematic review may have questionable clinical relevance because the observed average reduction from baseline was 2.8% (95% CI: 1.31 to 4.24), which is less than the suggested 3-5% reduction considered as clinically significant (Verweij *et al.,* 2013). Findings for body fat (%) are similar; despite modest change from baseline improvements, there were no difference between those completing exercise interventions or controls.

A previous systematic review reported improvements to anthropometric outcomes (Haqq *et al.,* 2015), but the basis of these compared lifestyle modification, as opposed to exercise, with a control group. In addition to this, when comparing lifestyle interventions to control, another review (Moran *et al.,* 2011) reported statistical improvements in body mass and abdominal adiposity. In the current systematic review, exercise and diet interventions were compared to diet alone and both groups demonstrated favourable changes; however, there was no evidence of an effect in favour of either intervention for any of these outcomes.

There was no evidence of a statistical effect of exercise on the androgenic profile of included participants when compared with control. Similarly, where such analyses were possible, there was no evidence of effect favouring either diet and exercise combined or diet only. This was further supported by subgroup analyses where the evidence of relevant effects was minimal. To investigate these findings, the baseline values of women with PCOS included in this current review were considered; typically, hormonal concentrations were below recommended cut-offs for diagnosing hyperandrogenism [testosterone >2.5 nmol/L and SHBG <30 nmol/L (Vanky *et al.,* 2004)], which suggests that the included participants were not markedly hyperandrogenic. In contrast to the present study, Moran *et al.* (2011) did observe a testosterone reduction following lifestyle interventions, but no effects upon FAI (100 x total testosterone/SHBG), which is often considered a more valid marker of hyperandrogenism (ESHRE/ASRM, 2004). Previous studies have investigated the effects of exercise on androgens in the general population, Indeed, a review of exercise-induced changes on the androgenic profile of premenopausal healthy women reporting that exercise invoked an acute increase to circulating androgens (Enea *et al.,* 2011), but the longer-term effects are less clear. Another meta-analysis investigated the chronic androgenic responses which were attributable to

exercise and reported a statistical reduction in concentrations of bioavailable testosterone (MD: -0.18 pg/mL, 95% CI: -0.29 to -0.07; 1369 participants, 9 trials, $I^2 = 0\%$), as well as increased SHBG (MD: 3.93 nmol/L, 95% CI: 0.98 to 6.87; 1643 participants, 14 trials, $I^2 = 75\%$) following participation in an exercise regime in healthy women (Ennour-Idrissi *et al.*, 2015). Collectively, these data suggest that exercise interventions may play a role in the regulation of androgenic profiles. However, the optimal dose is unclear and, potentially, there could be an unquantifiable variation to responses in women with menstrual disruption (Tworoger *et al.*, 2007).

Finally, although there appears to be an increasing recognition of the deleterious effects of PCOS on HRQoL and other psychosocial components, only three of the included trials measured these outcomes in the comparison of exercise versus control. Although there was no evidence of effect in any of the PCOS-Q domains, scores in four domains of the SF-36, namely physical functioning, general health, social functioning and mental health, were all statistically improved following exercise compared to control. The meta-analysis revealed improvements of ≥10% in these four domains for exercise, supporting the notion that exercise in these patients may improve their perception of physical and mental wellbeing.

3.9.4. Overall completeness and applicability of evidence

A comprehensive and systematic search of relevant electronic databases, and the reference lists from included publications and relevant reviews was completed. From this search,16 RCTs, one quasi-RCT, and one randomised crossover trial were identified. The applied review process located and combined data from more trials, made a greater number of comparisons and included a wider range of outcomes when compared to previous systematic reviews (Harrison *et al.*, 2011; Domecq *et al.*, 2013; Moran *et al.*, 2011; Haqq *et al.*, 2015). At the time of writing, this was the first time data from 10 of the trials included in this systematic review have been meta-analysed (Almenning, *et al.*, 2015; Konopka *et al.*, 2015; Nasrekani *et al.*, 2016; Petrányi and Zaoura-Petrányi, 2011; Roessler *et al.*, 2013; Sá *et al.*, 2015; Saremi *et al.*, 2013; Saremi and Yaghoubi, 2016; Turan *et al.*, 2015; Vizza *et*

al., 2016), which suggests that this may be the most comprehensive and up-to-date systematic review on the topic of exercise in the management of women with PCOS.

Despite this, it must be noted that there are certain limitations to the studies included in this systematic review, which reflect the overall low quality of the relevant existing data. As such, it is likely that many of the included trials did not have sufficient statistical power to detect meaningful differences between study groups. In fact, only seven of the included trials stated the methods employed to calculate their target sample size. Due to small participant numbers (*e.g.*, median: exercise $n = 11$; control $n = 12$) in these studies, it is unlikely that sufficient statistical power was achieved to make findings applicable to the general population, or to ensure that false positive/negative results were not reported. Therefore, it is important that future trials are sufficiently powered to detect changes in (at least) their primary outcomes.

Moreover, PCOS is a heterogeneous condition that is categorised by a range of phenotypes; the respective phenotypic subgroups may present varying levels of hyperandrogenism, menstrual disruption and polycystic ovarian morphology (Kyrou *et al.*, 2014). Because of this, it can be assumed that each phenotype will respond differently to exercise and/or dietary interventions. Of the included trials, most did not target a specific PCOS phenotype, and in addition, the pre-defined review protocol specified that studies were eligible for inclusion as long as they defined their participants based on any of the existing PCOS definitions and/or diagnostic criteria. It would be beneficial for future studies to specify PCOS subgroups/phenotypes of their participants and investigate the phenotypic exercise-induced effects accordingly. Another concern surrounds the representativeness of the cohorts included in the studies which were included in this systematic review; it is unclear whether the ethnicity, socio-economic or educational status of included participants is a true representation of typical patients with PCOS, or whether these variables have influenced the observed (or unobserved) effects.

Finally, all included trials reported baseline and immediately post-intervention data, but only one trial (Stener-Victorin *et al.*, 2009) completed follow-up beyond the end of the intervention. Therefore, based on the existing data, it is difficult to assess the lasting, long-term effects of exercise

(or exercise and diet) for women with PCOS. Hence, future research should attempt to determine whether behaviours relating to PA are changed (*i.e.,* are patients able to maintain PA levels beyond study end) in this patient population, and whether the observed physical and physiological effects remain beyond the short-term.

3.9.5. Quality of the evidence

As outlined in the methods (Section 2.1.3), due to the general nature of exercise interventions, all included trials were deemed to have a high risk of performance bias, but the quality of the evidence was not downgraded based on this alone. All included trials, but one (Vigorito *et al.*, 2007), was judged to have a high risk of detection bias due to lack of blinding outcome assessors. Although the logistics of assessor blinding can be more difficult and additional resources may be required, reasonable steps could have been taken to reduce detection bias risk in each of these trials. Both selection and reporting bias were inadequately reported in >50% of trials, resulting in judgements of unclear risk being made. Furthermore, nearly 45% (n = 8) of the included trials were judged to be at a high risk of attrition bias. There were six trials with an unclear or high risk of imbalances between study groups at baseline, whilst, in most instances, contamination was also unreported resulting in an unclear judgement. Again, few studies reported data on intervention adherence (33%, n = 6); however, from the trials that did report these data, adherence rates were generally good (median: 90%). Similarly, in the 10 trials that did report on attrition, the median value was 19.5%, with five of these trials reporting drop-out values that were under the 20% attrition threshold outlined in the protocol (Section 2.1.3).

Statistical effects were observed in 13 of the main analyses, but in three of those evidence of at least substantial heterogeneity (I^2 ≥50%) was reported. However, this was largely explained by investigation of subgroups and/or the removal of trials with the most extreme values. For all primary outcomes, the quality of evidence was rated as very low to low due to a combination of unclear or high risk for randomisation or allocation procedures, lack of blinding, unclear or improper handling of missing data, high attrition rates, unclear risk of selective reporting bias, contamination, low

adherence, or evidence of considerable heterogeneity. The quality of evidence for all outcomes was downgraded due to imprecision resulting from a small number of participants and either wide confidence intervals surrounding the effect estimate, or the inclusion of a null or negligible effect, as well as an appreciable benefit, within the confidence interval for the mean difference.

3.9.6. Limitations and potential biases in the review process

In addition to the limitations of the included studies mentioned in Section 3.9.4, there are potentially further limitations to the current systematic review. As such, despite a thorough and comprehensive search of relevant databases, it is possible that searches may have missed trials that were eligible for inclusion. Furthermore, no additional studies were identified from the reference lists of the included publications; whilst this may support the comprehensiveness of completed searches, it may also represent a potential methodological flaw. Also, no language restriction was placed upon the searches resulting in the return of several foreign language papers; three trials were written in Persian (Farsi) (Nasrekani *et al.*, 2016; Saremi *et al.*, 2013; Saremi and Yaghoubi, 2016), and another was in Hungarian (Petrányi and Zaoura-Petrányi, 2011). To assess these trials, translation services and software were utilised; whilst interpretation of results tables was reasonably straightforward, evaluating the quality of these studies was a considerable challenge. Consequently, in most instances, when assessment for risk of bias in these trials was completed, judgements of 'unclear risk' had to be made.

Finally, only full publications were eligible for inclusion in the review and this may have contributed to publication bias. Whilst the inclusion of grey literature may have influenced the findings of this review, its inclusion may also have increased the risk of associated bias. However, due to a lack of eligible trials, publication bias analysis was not performed.

3.9.7. Future Directions

Based upon the findings of this review, it is evident that there is a dearth of trials comparing exercise and diet in combination with other comparators (*e.g.*, diet only, exercise only, and usual care control). Considering that lifestyle changes (*i.e*. both diet and exercise) are recommended in the management of PCOS, there is a scarcity of studies assessing the effectiveness of these interventions and the available data are not sufficient to draw definitive conclusions/recommendations for use by clinical practitioners. Therefore, future trials should aim to make comprehensive comparisons involving interventions that incorporate both exercise and diet.

Furthermore, the majority of studies included in the current systematic review have small sample sizes, and even where studies have reported power calculations, it is likely that they are still insufficiently powered to detect meaningful changes in all reported outcomes. Therefore, it is important that future studies are robustly designed and sufficiently powered to more reliably inform future clinical practice, and support relevant evidence-based guidelines and recommendations. Considering the high prevalence of PCOS in reproductive-aged women, large RCTs studying the effectiveness of lifestyle interventions in this young patient population are still clearly needed.

There was also a striking lack of follow-up testing, beyond that which was done immediately post-intervention, to assess the longer-term effects of such interventions. Without follow-up reassessments, it is impossible to determine whether any intervention-induced improvements are maintained, and to ascertain whether the studied intervention has resulted in sustained changes to the lifestyle behaviours of participants, an aspect which is vital for the long-term management of these patients.

The final consideration for future works surrounds the internal validity within each study. The majority of included studies in this review had an unclear or high risk of bias across multiple risk of bias domains. The risk of bias could be addressed in the following ways; to address selection bias, participants must be randomised appropriately (*i.e.,* computer generated randomisation, allocation concealed from researchers, opaque envelopes, *etc.*) and importantly, these methods should be explicitly described within publications. Due to the nature of exercise interventions, it is difficult to

reduce performance bias but this should not necessarily reduce the quality of the evidence. In pharmaceutical trials, performance bias can be avoided through the use of placebos, but creating sham exercise intervention arms is challenging (El-Kotob and Giangregorio, 2018). One potential solution could be through the use of multiple comparisons; this review highlights a paucity of trials that incorporate diet and exercise in their intervention arms. Future studies could compare lifestyle intervention A (*i.e.* diet and education) to lifestyle intervention B (*i.e.* diet and exercise), effectively masking participants to which is the true intervention.

Detection bias was also high across studies, within the current systematic review. Detection bias can quite easily be negated by blinding those who will assess outcomes, to the allocation of the participant, which logistically can be achieved quite simply. The methods adopted for this process should also be clearly reported within methodologies. Attrition from study groups is another way in which bias can be introduced to RCTs, thus reducing statistical power and affecting generalisability of findings (Fewtrell *et al.*, 2008). Attrition can be managed in many ways; regular communication with all participants, reduced burden upon participants (*e.g.*, shorter questionnaires, reduced follow-up, *etc.*), provision of incentives for continued engagement, and behavioural strategies (Brueton *et al.*, 2011) could all be used to reduce attrition bias in future studies. Another bias that can be reduced in a straightforward manner is reporting bias; by pre-registering a trial, explicitly stating which outcomes will be analysed, and the methods used for analysis, reporting bias can be nullified. Pre-registration of trials has also been shown to substantially reduce publication bias (Warren, 2018). Finally, studies should also record and report adherence to an intervention and any potential contamination in comparator groups (El-Kotob and Giangregorio, 2018). Providing these data can provide important information about whether the intervention group actually received what they were prescribed and whether, inadvertently or not, the control group received a dose of the intervention.

If researchers are able to address these issues, there will undoubtedly be a reduced risk of bias within their work and ultimately, improved confidence in their findings.

3.10. Conclusions

When data were combined in meta-analyses, only FI, total cholesterol, LDL-C and VO_2 max had statistically favourable changes for both change from baseline and comparison of post-intervention values. In addition to this, HOMA-IR, triglycerides and waist circumference were statistically improved from baseline after exercise, but were not different to post-intervention control values. Finally, RHR and BMI following exercise were statistically lower than post-intervention values of control groups. Compared with control, exercise also improved multiple domains of the SF-36; the physical functioning, general health, social functioning, and mental health domains all improved following exercise.

Subgroup analyses revealed that the greatest favourable changes when exercise was compared to control were seen in participants with overweight (FI, HOMA-IR, triglycerides, VO_2 max and waist circumference) or obesity (BMI, body mass and waist circumference); post-intervention value analyses also showed beneficial effects in those who were overweight (LDL-C, VO_2 max, RHR, BMI, waist circumference and body fat percentage). Aerobic exercise interventions improved FI, HOMA-IR, total cholesterol, triglycerides, VO_2 max, BMI, waist circumference and body fat percentage. Conversely, resistance training lowered HDL-C concentrations and increased BMI, but reduced waist circumference, whilst post-intervention improvements in HDL-C were also evident following resistance training interventions. Compared to control, supervised exercise interventions improved outcomes more than those that were unsupervised. Shorter duration interventions performed better than longer interventions; improved change from baseline FI, HOMA-IR, total cholesterol, LDL-C, triglycerides, VO_2 max and waist circumference was found in shorter duration trials, compared with only improved VO_2 max in those >12 weeks. Based on limited available data, there were no notable differences between the effects of exercise and diet combined, and diet alone. Due to lack of available trials, it was not possible to compare the effectiveness of exercise versus diet or exercise and diet combined versus diet.

Although the evidence presented within this systematic review is largely drawn from RCTs, a cautious approach should be adopted when interpreting the present findings. Many of the outcomes

presented statistically significant, yet modest, effects and generally had wide confidence intervals around the effect estimate (indicating greater uncertainty). Furthermore, the observed statistical effects in many of the analyses were sensitive to the addition/removal of individual trials regardless of their weighting within the analysis. Using the GRADE approach, the quality of evidence was rated as very low or low for all primary outcomes. It is recommended that future trials should be rigorously designed and sufficiently powered so that they are more generalizable to the wider PCOS population. In order to be more closely aligned with current treatment recommendations, future studies should ideally include a dietary component alongside exercise interventions.

Chapter 4: Results

Study 2.

A comparison of self-reported energy expenditure, health-related quality of life, and attitudes towards physical activity in women with, and without polycystic ovary syndrome

4.1. Background/rationale

Polycystic ovary syndrome (PCOS) is a common endocrine disorder that affects up to 21% of reproductive-aged women (Yildiz *et al.*, 2012). Diagnosis is usually made when women are hyperandrogenic with the presence of oligomenorrhea and/or polycystic ovaries and this often manifests in a range of undesirable symptoms such as acne, hirsutism and infertility. Alongside these defining characteristics, many women with PCOS experience metabolic complications; some estimated 40-60% are either overweight or obese (Moran *et al.*, 2009), and when compared to women without PCOS, up to 70% have increased insulin resistance (IR) (Hutchison *et al.*, 2011), independent of obesity (Teede, Hutchison and Zoungas, 2007; Dunaif *et al.*, 1989).

In addition to these hyperandrogenic and metabolic symptoms, PCOS is also recognised as a chronic disorder that adversely affects patient health-related quality of life (HRQoL; Jedel *et al.*, 2010). HRQoL is a widely reported outcome which is deemed an important consideration in the treatment and management of chronic disease (Dokras *et al.*, 2018). This is because it relates to the patient-reported physical, social and emotional effects of an underlying condition and the treatments that are associated with it (Colwell *et al.*, 1998). The assessment of HRQoL is usually completed via generic self-report questionnaires, such as the Short Form 36-item (SF-36) Health Survey, or disease specific questionnaires, such as the PCOS-Questionnaire (PCOS-Q).

Previous studies have consistently reported lower HRQoL in women with PCOS when compared with either normative data from the general population (Jones *et al.*, 2010; Ching, Burke and Stuckey, 2007), or control groups (Li *et al.*, 2011; Barnard *et al.*, 2007; Coffey, Bano and Mason, 2006; Hahn *et al.*, 2005). Furthermore, women with PCOS are more likely to experience greater levels of psychological morbidity (*i.e.* anxiety and/or depression) than women without PCOS (Deeks *et al.*, 2011). The psychological components of HRQoL for women with PCOS are also reportedly poorer than in women with other physical conditions, such as asthma, epilepsy, coronary heart disease or arthritis (Coffey, Bano and Mason, 2006).

Previous studies have compared physical activity (PA) levels of women with PCOS against controls, with the majority of such findings reporting no statistical differences in energy expenditure (MET-mins/wk) between these groups (Mario *et al.*, 2012; Rodino *et al.*, 2016), despite poorer physical and mental health in women with PCOS. It can therefore be assumed that PA, or lack thereof, is not implicated in the health differences between women with PCOS and healthy women. However, there is less evidence about the effectiveness of increasing PA to manage PCOS symptoms. Indeed, the physical benefits of increasing PA levels (*e.g.* improvements to fitness, body composition or metabolic health) have been widely reported across a range of populations. However, it has also been stated that increased PA may improve HRQoL; following exercise interventions, beneficial effects have been reported in both cancer patients receiving treatment and in cancer survivors (Mishra *et al.*, 2012), in patients with chronic respiratory conditions (Eichenberger, *et al*., 2013) and in patients with rheumatoid arthritis or osteoarthritis (Kelley, *et al.*, 2015).

Despite the high global prevalence of PCOS, less is known about the effectiveness of PA at improving HRQoL within this patient population. The author's recent systematic review (Chapter 3) reported statistical improvements in a range of outcomes (*i.e.* fasting insulin, lipids, and cardiorespiratory fitness), including components of HRQoL. Moreover, whilst no benefit was observed in any domain of the PCOS-Q, statistical improvements were also found in multiple domains of the SF-36. When exercise was compared to usual care, physical functioning [mean difference (MD): 11.81, 95% confidence interval (CI): 2.36 to 21.25; I^2 = 74%], general health (MD: 10.05, 95% CI: 3.89 to 16.20; I^2 = 0%), social functioning (MD: 11.75, 95% CI: 2.56 to 20.95; I^2 = 6%) and mental health (MD: 11.70, 95% CI: 1.27 to 22.13; I^2 = 47%) domains were all statistically improved from baseline. Another recent systematic review (Lim, *et al.*, 2019) assessed the effectiveness of lifestyle interventions (*i.e.* exercise and diet) upon HRQoL (as measured by the PCOS-Q) in women with PCOS. This systematic review reported statistical improvements for the Emotions (MD: 0.77, 95% CI: 0.30 to 1.23; I^2 = 92%) and Infertility (MD: 0.68, 95% CI: 0.21 to 1.14; I^2 = 87%) domains of the PCOS-Q. Whilst these findings appear promising, there are some limitations, namely small sample sizes (n = 84 and n = 95, respectively), evidence of considerable

heterogeneity and also the lack of a current consensus as to what constitutes a clinically meaningful difference for these metrics within this patient population (Teede, *et al.*, 2018).

Although there is limited evidence supporting the effectiveness of exercise interventions, there appears to be a paucity of research on the role of habitual PA in the management of PCOS. In contrast to structured exercise interventions, habitual activity accounts for activities of daily living (*e.g.* walking for transport, shopping or household chores) and, when undertaken in combination with aerobic exercise, may potentiate the cardio-metabolic health benefits of exercise alone (Swift *et al.*, 2012). Independently, higher habitual PA levels have been linked with improved health outcomes in a range of patient populations (Dwyer *et al.*, 2007; Jennersjö *et al.*, 2012; Manjoo *et al.*, 2012). More specifically, in women with type 2 diabetes, a 1000 steps per day increment has been associated with clinically important reductions in systolic (-2.6 mmHg, 95% CI: -4.1 to -1.1) and diastolic (-1.4 mmHg, 95% CI: -2.2 to -0.6) blood pressure. Furthermore, another study reported that women who demonstrated moderate levels of ambulatory activity (≥7500 steps/day) had ~50% lower levels of depression (assessed via the computerised Composite International Diagnostic Interview) when compared to their sedentary (<5000 steps/day) counterparts (McKercher *et al.*, 2008); these associations were not evident when males were analysed.

In women with PCOS, studies that have used objective measurements of habitual PA are sparse. One such cross-sectional study used pedometers to measure habitual PA, comparing health outcomes of those who were active (≥7500 steps/day) against those who were inactive (Mario *et al.*, 2017), and reported lower BMI, waist circumference, serum androgens and an improved lipid profile in the former more active group. A more recent study also reported a modest reduction in BMI (-0.20 kg/m^2, 95% CI: -0.38 to -0.01) alongside clinically important differences in inflammatory biomarkers (interleukin-6: -0.81 ng/L, 95% CI: -1.37 to -0.25; and C-reactive protein: -0.68 mg/L, 95% CI: -1.30 to -0.06) for each additional 1000 steps per day (Webb *et al.*, 2018).

In other studies, participants with PCOS have self-reported their PA levels, often using the International PA Questionnaire (IPAQ). One such study (Lamb *et al.*, 2011) categorised participants

as either meeting (active), or not meeting (inactive) the US Department of Health and Human Services (DHHS) PA guidelines (*i.e.* 150 mins per/week of moderate-intensity PA, or 75 mins per/wk of vigorous-intensity PA, or an equivalent combination of both). This study reported that ~59% of women with PCOS in this study sample (n = 88) met the DHHS PA guidelines and that, when compared to active women, the inactive group had a higher BMI, waist circumference and increased weight fluctuations. Furthermore, inactive participants also had higher fasting blood glucose (94.0 ± 21.3 mg/dL versus 87.8 ± 12.1 mg/dL) and lower sex hormone binding globulin (SHBG) concentrations (40.4 ± 5.9 nmol/L versus 68.4 ± 7.3 nmol/L) than their physically active counterparts. One additional finding was that inactive participants had higher depression scores on the Beck Depression Inventory-Fast Screen (BDI-FS) questionnaire.

A more recent study (Greenwood *et al.*, 2016) sampled 326 women with PCOS and found that 56% met the Department of Health and Human Services (DHHS) PA guidelines, with 83% of these women meeting the guidelines primarily through vigorous-intensity exercise (defined as: 'activities that take hard physical effort and make you breathe much harder than normal') as opposed to moderate intensity. In this study, women classified as vigorous exercisers had a lower median BMI, smaller waist circumference and greater concentrations of high-density lipoprotein cholesterol (HDL-C) when compared to those who were inactive or moderately active. Moreover, despite no statistical between-group differences in fasting blood glucose and insulin concentrations, the vigorous group demonstrated an improved metabolic profile [*i.e.* significant decrease in oral glucose tolerance test (OGTT) and homeostasis model assessment-insulin resistance (HOMA-IR) values], and fewer diagnoses of metabolic syndrome when compared to those who were inactive.

In summary, women with PCOS reportedly have poorer HRQoL than their healthy counterparts regardless of PA levels. However, whilst the evidence of exercise's potential to improve HRQoL is limited, there appears to be some evidence that if women with PCOS increase their activities of daily living, either independently or alongside exercise, HRQoL can be improved.

4.2. Objectives and hypotheses

Based on the aforementioned background, the hypotheses explored in the present study can be summarized in the following points:

1. Women with PCOS have a poorer quality of life than women without PCOS, irrespective of PA levels.
2. The higher the PA level of women with PCOS, the less severe their symptoms and higher their HRQoL.
3. Greater perceived benefits, and fewer perceived barriers to PA leads to increased self-efficacy for exercise, which in turn improves PA levels.
4. A diagnosis of PCOS decreases participant self-esteem, which also reduces HRQoL.

Accordingly, the objectives of the current study were to identify whether there are significant differences in HRQoL (including physical and psychological components) between women with PCOS and a healthy control group, and to explore whether higher levels of PA facilitate improved PCOS-related symptoms (as measured by PCOS-Q). Furthermore, this study aimed to estimate any potential simultaneous impact of not only a PCOS diagnosis, but also of other predictive factors (*i.e.*, PA and its determinants, BMI, self-esteem), upon the mental and physical domains of HRQoL.

4.3. Methods

For complete details of methodology, please refer to Chapter 2, Sections 2.2.1 to 2.2.6.

4.4. Results

4.4.1. Study cohort - Participants characteristics

Following the recruitment period, 194 participants accessed the online surveys and agreed to participate (Figure 4.1). Based upon the exclusion criteria, 40 participants were deemed ineligible for participation and their involvement in the study ceased. Reasons for exclusion were: asthma (*n* = 19), depression (*n* = 4), endometriosis (*n* =3), inflammatory arthritis (*n* = 3), hypothyroidism (*n* = 2), osteoporosis, osteoarthritis, irritable bowel syndrome, Sjorgen's syndrome, chronic migraines, bipolar disorder, type 2 diabetes mellitus, hypermobility syndrome, Ehlers-Danlos syndrome, multiple sclerosis, Raynaud's phenomenon, spinal cord injury and pregnancy (all *n* = 1). Three participants cited two conditions.

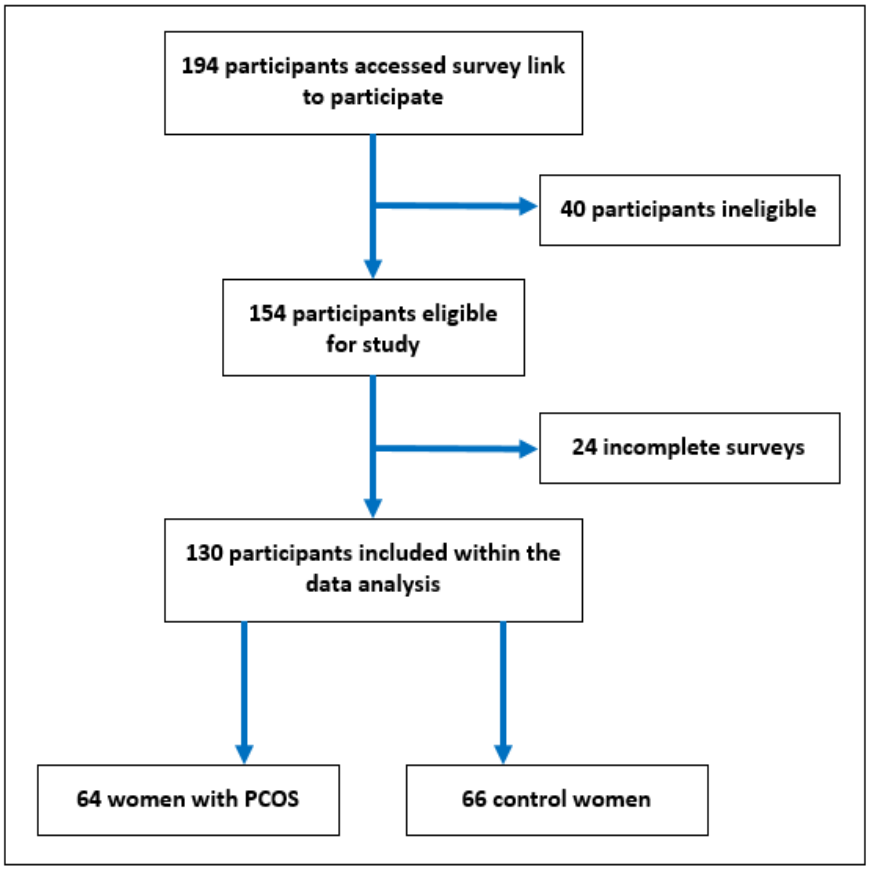


Figure 4.1. Participant flow diagram

Of the remaining 154 participants, 24 failed to complete any data beyond the demographic questionnaire and, thus, were deemed to have provided insufficient data to warrant inclusion in the analysis. Descriptive statistics for women with PCOS (n = 64) and the control group (n = 66) are presented in Table 4.1.

Table 4.1. Demographics of participants.

Variable	Total	Women with PCOS (n = 64)	Control (n = 66)
Age Range (years)	19-45	21-45	19-45
Ethnicity			
White	108 (83.1%)	57 (89.1%)	51 (77.3%)
Asian or Asian British	13 (10%)	2 (3.1%)	11 (16.6%)
Black or Black British	3 (2.3%)	-	3 (4.5%)
Chinese	1 (0.8%)	1 (1.6%)	-
Gypsy or Traveller	1 (0.8%)	1 (1.6%)	-
Other Mixed background	3 (2.3%)	3 (4.7%)	-
Declined to specify	1 (0.8%)	-	1 (1.5%)
Marital Status			
Single	52 (40%)	23 (35.9%)	29 (43.9%)
Married	52 (40%)	22 (34.4%)	30 (45.5%)
Divorced	6 (4.6%)	5 (7.8%)	1 (1.5%)
Widowed	1 (0.8%)	1 (1.6%)	-
Civil-partnership	3 (2.3%)	1 (1.6%)	2 (3.0%)
Other	16 (12.3%)	12 (18.8%)	4 (6.1%)
Children			
Yes	38 (29.2%)	14 (21.9%)	24 (36.4%)
No	92 (70.8%)	50 (78.1%)	42 (63.6%)
Occupation			
Full-time employed	63 (48.5%)	34 (53.1%)	29 (43.9%)
Part-time-employed	16 (12.3%)	10 (15.6%)	6 (9.1%)
Student	35 (26.9%)	10 (15.6%)	25 (37.9%)
House person	5 (3.8%)	5 (7.8%)	-
Unemployed	2 (1.5%)	1 (1.6%)	1 (1.5%)
Other	9 (6.9%)	4 (6.3%)	5 (7.6%)
Education			
Secondary	6 (4.6%)	6 (9.4%)	-
College	28 (21.5%)	15 (23.4%)	13 (19.7%)
Undergraduate	50 (38.5%)	26 (40.6%)	24 (36.4%)
Postgraduate	32 (24.6%)	14 (21.9%)	18 (27.3%)
Doctorate	14 (10.8%)	3 (4.7%)	11 (16.7%)
Income			
≤ £39,999	77 (59.2%)	43 (67.3%)	34 (51.5%)
£40,000 - £79,999	42 (32.3%)	16 (25%)	26 (39.4%)
≥ £80,000	11 (8.4%)	5 (7.8%)	6 (9.1%)

All percentage data rounded to one decimal place.

The majority of participants were of White ethnicity (83.1%) and were either single (40%) or married (40%), in full-time employment (48.5%), educated to degree level (38.5%), had a household income ≤ £39,999 (59.2%), and did not have children (70.8%). There were some demographical differences between the groups; ~63% of women with children came from the control group. Women in the control group also tended to be educated to a higher level and have a greater household income than their counterparts. A larger number of women in the control group also self-reported that they were currently a student; this trend may be associated with the location that the recruitment took place (*i.e.*, Aston University).

Table 4.2. Comparison of self-reported variables between women with PCOS and control group.

	Total (*n* = 130)	**PCOS** (*n* = 64)	**Control** (*n* = 66)	**Missing** PCOS / CON	**Median or mean diff**	**95% CI for MD**		**Cohen's *d***	***P* value**
						Lower	**Upper**		
Age (years)	31 (11.25)	30 (9.75)	32 (13.5)	0 / 0	-2.00	-	-	-0.192	.297
Height (cm)	165.1 (10.00)	165.1 (8.06)	164.5 (9.25)	0 / 0	0.60	-	-	0.107	.349
Weight (kg)	65.16 (38.27)	91.3 (45.31)	63.0 (18.3)	0 / 0	28.30	-	-	1.121	**<.001**
BMI (kg/m^2)	24.51 (13.29)	32.91 (16.6)	23.28 (5.25)	0 / 0	9.63	-	-	1.129	**<.001**
WC (cm)	81.28 (31.60)	101.60 (40.62)	74.00 (16.66)	11 / 8	27.60	-	-	1.315	**<.001**
HRQoL									
SF-12 Physical	23.00 (4.25)	20.0 (7.0)	23.0 (3.0)	6 / 1	-3.00	-	-	-0.999	**<.001**
SF-12 Mental▲	20.01 (4.88)	16.71 (4.66)	22.29 (3.63)	6 / 1	-5.58	-7.07	-4.10	-1.346	**<.001**
SF-12 Total	43.00 (10.50)	36.0 (12.0)	46.0 (7.50)	6 / 1	-10.00	-	-	-1.327	**<.001**
IPAQ									
MET-min/wk	3478.5 (3257.9)	3447.0 (3648.0)	3010.5 (3295.5)	5 / 5	436.50	-	-	0.158	.280
Sitting-mins/wk	2370.0 (1515.0)	2460.0 (1665.0)	2340.0 (1830.0)	1 / 0	120.00	-	-	0.060	.603
SES▲	16.86 (5.68)	14.0 (5.21)	19.34 (5.20)	5 / 1	-5.34	-7.17	-3.47	-1.022	**<.001**
EBBS									
Benefits	88.50 (18.00)	84.0 (14.75)	93.5 (17.0)	4 / 0	-9.50	-	-	-0.595	**.001**
Barriers	39.00 (8.25)	38.0 (9.25)	41.5 (8.25)	4 / 0	-3.50	-	-	-0.482	**.003**
Total▲	128.85 (25.00)	122.78 (18.92)	133.47 (20.50)	4 / 0	-10.69	-16.59	-4.78	-0.639	**.001**
SEE●	42.40 (17.77)	40.44 (21.35)	43.64 (15.95)	2 / 0	-3.20	-9.83	3.43	-0.171	.341

Key: Unless otherwise indicated, data presented as median (IQR) and Mann-Witney U tests performed; ▲ Parametric data reported as mean (SD) and Student's t-test performed; ● Parametric data reported as mean (SD) and Welch's test performed; CON: Control group; CI: confidence interval; MD: Mean difference; BMI: Body Mass Index; WC: waist circumference; HRQoL: Health-related Quality of Life; IPAQ: International Physical Activity Questionnaire; SES: self-esteem scale; EBBS: Exercise Benefits/Barriers Scale; SEE: self-efficacy for exercise; *P* value: significant values in bold font indicated by $P \leq .05$

There were no statistically significant differences between the age and height of the two groups but, women with PCOS had higher body weight, BMI and waist circumference than the control group (Table 4.2). In addition to this, women with PCOS reported lower scores in the physical and mental domains of the SF-12, and overall poorer HRQoL as assessed by the total SF-12 score. They also had significantly lower self-esteem than control women did. In addition to this, women with PCOS also perceived fewer benefits and greater barriers to exercise.

When individual perceived benefits of exercise were identified, the top three with the highest mean (± SD) scores (out of 4) in sequence were: '*exercising increases my level of physical fitness*' (3.45 ± 0.55), '*exercising improves functioning of my cardiovascular system*' (3.38 ± 0.56), and '*exercise gives me a sense of personal accomplishment*' (3.33 ± 0.65). In contrast, the three perceived benefits with the lowest scores were: '*exercising increases my acceptance by others*' (2.14 ± 0.78), '*exercising lets me have contact with friends and persons I enjoy*' (2.23 ± 0.90), and '*exercising is a good way for me to meet new people*' (2.29 ± 0.85). When each group's benefits were assessed separately, the top three perceived benefits for the control group were identical to the whole cohort. Women with PCOS perceived '*exercise improves my flexibility*' (3.52 ± 0.75) as their second benefit, just behind '*exercising increases my level of physical fitness*' and ahead of '*exercising improves functioning of my cardiovascular system*'. The least cited perceived benefits for each group were different to the whole cohort. In women with PCOS, '*exercising lets me have contact with friends and persons I enjoy*' scored lowest, followed by '*exercise is good entertainment for me*' and '*exercising is a good way for me to meet new people*'. The least scored benefits in the control were the same as the whole cohort, but with '*exercising lets me have contact with friends and persons I enjoy*' and '*exercising is a good way for me to meet new people*' reversed.

When the perceived barriers were scored, the ranking for the highest scores were virtually identical for the whole cohort and each population, respectively. These are: 1) '*exercise tires me*'; 2) '*exercise is hard work for me*', and; 3) '*I am fatigued by exercise*'. The mean score for the barrier '*exercising takes too much of my time*' was also equal 3rd in the control group. Similarly, the lowest scoring barrier across both groups and the whole cohort was '*I think people in exercise clothes look funny*'.

This was followed by '*places for me to exercise are too far away*' and '*there are too few places for me to exercise*' in the PCOS group and whole cohort. The control group barriers with the 2nd and 3rd lowest scores were '*I am too embarrassed to exercise*' and '*my spouse (or significant other) does not encourage exercising*', respectively.

In contrast to EBBS, there were no statistical differences in self-efficacy for exercise, and when data from the IPAQ was analysed, there were also no between-group statistical differences in either MET-min/wk or sitting time.

Women with PCOS were asked to stipulate to which phenotypic subgroup they belonged. The majority (51.6%) self-reported as having excess androgens, menstrual dysfunction and polycystic ovaries as identified by ultrasound (Figure 4.2). A further 11 (17.1%) of the women with PCOS in this study were unsure of the characteristics of which their diagnosis had been made.

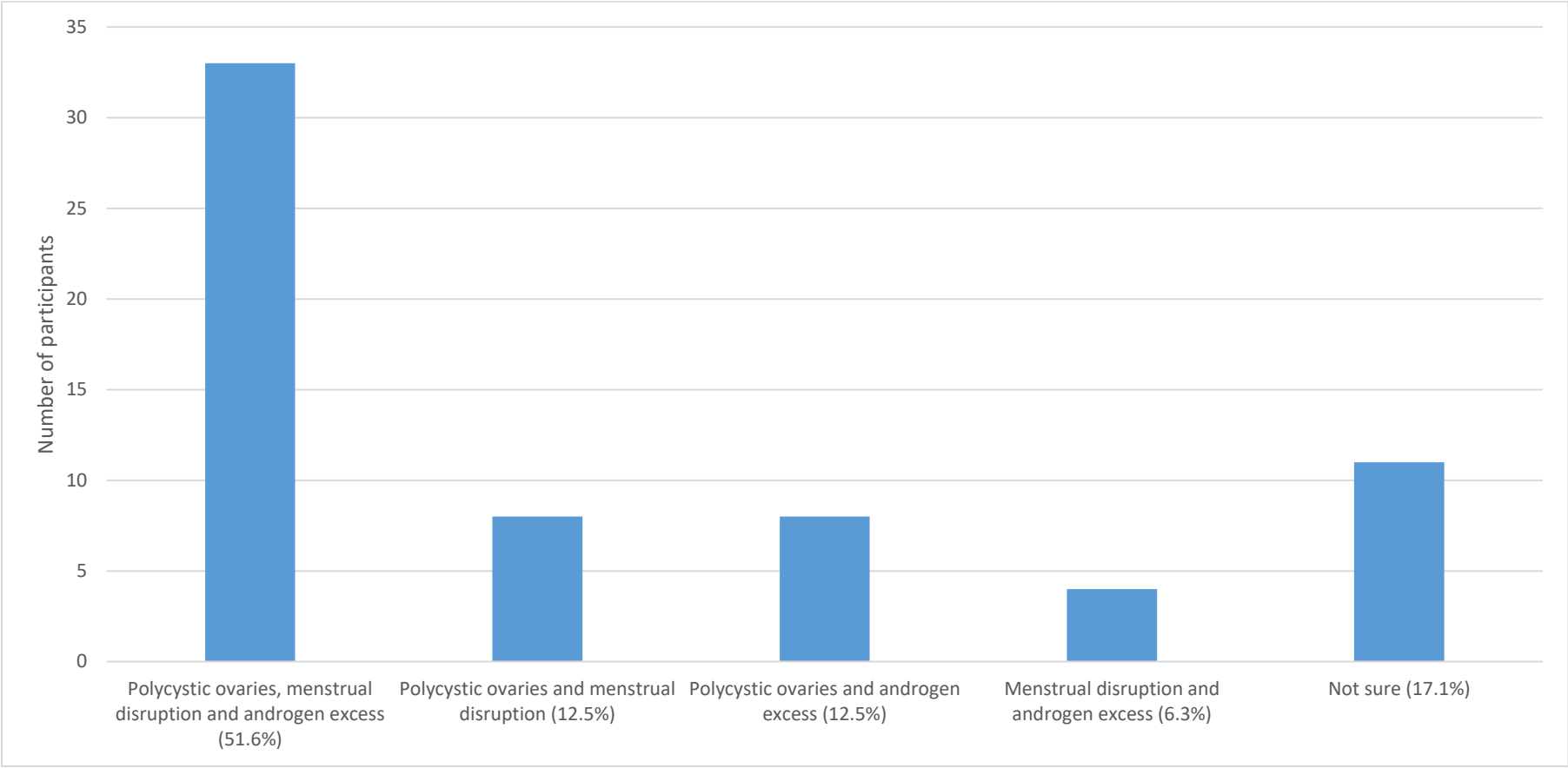


Figure 4.2. Self-reported phenotypic subgroups of women with PCOS.

Table 4.3. PCOS-Q domain specific scores from women with PCOS.

PCOS-Q Domain	Median Score	Interquartile Range
Emotions	3.05	1.57
Body Hair	2.60	3.30
Weight	1.90	2.75
Infertility	2.50	3.81
Menstrual Problems	3.00	2.25

Key: Emotions: Emotion domain scores from the PCOS-Q; Body Hair: Body hair domain scores from the PCOS-Q; Weight: Weight domain scores from the PCOS-Q; Infertility: Infertility domain scores from the PCOS-Q; Menstrual Problems: Menstrual Problems domain scores from the PCOS-Q. A lower score indicates a greater severity of problem.

Results from the PCOS-Q are presented in Table 4.3; it is evident that concerns about body weight and infertility are the most prevalent in this sample. When domain scores were compared, there were statistical differences between five domains (Table 4.4); the Weight domain was significantly lower than the other four domains and Infertility domain scores were significantly lower than the Emotions domain.

There were no statistically significant differences in PCOS-Q domain scores between the self-reported phenotypic subgroups.

Table 4.4. Pairwise comparison (Durbin-Conover) of PCOS-Q domains.

Pairwise Comparison		Median Diff	Statistic	*P* value
Emotions	Body Hair	-0.45	1.389	.166
Emotions	Weight	-1.15	4.291	**<.001**
Emotions	Infertility	-0.55	2.037	**.043**
Emotions	Menstrual Problems	-0.05	0.772	.441
Body Hair	Weight	-0.70	2.902	**.004**
Body Hair	Infertility	-0.10	0.648	.517
Body Hair	Menstrual Problems	0.40	0.617	.538
Weight	Infertility	0.60	2.253	**.025**
Weight	Menstrual Problems	1.10	3.519	**<.001**
Infertility	Menstrual Problems	0.50	1.266	.207

Key: Emotions: Emotion domain scores from the PCOS-Q; Body Hair: Body hair domain scores from the PCOS-Q; Weight: Weight domain scores from the PCOS-Q; Infertility: Infertility domain scores from the PCOS-Q; Menstrual Problems: Menstrual Problems domain scores from the PCOS-Q; *P* value: significant values in bold font indicated by $P \leq .05$

4.4.2. Correlations

Waist circumference was strongly (P <.001) correlated with BMI (τ_b = .652). Whilst waist circumference is an important clinical measure of central obesity, it was felt that because it had a number of missing values in both groups (11 and 8 in the women with PCOS and control group, respectively), and a strong correlation with BMI, that BMI was the most suitable weight-related variable for subsequent analysis. Similarly, the Benefit (τ_b = .819) and Barrier (τ_b = .611) scores were strongly associated with the total score from the EBBS; so too were the Physical (τ_b = .737) and Mental (τ_b = .844) domain scores with the total score from the SF-12. Accordingly, only the total scores from the EBBS and SF-12 were used in subsequent analyses.

BMI had a statistically significant negative correlation with the SF-12 total score (τ_b = -.297), RSE Scale (τ_b = -.271), and EBBS total score (τ_b = -.188). The implication here is that the higher the BMI of a participant, then the lower the score in these metrics. Conversely, there was a smaller, but statistically significant positive correlation between BMI and MET-mins/wk (τ_b = .139). Furthermore, there was a significant positive correlation between SEE Scale scores and EBBS (τ_b = .340) and MET-mins/wk (τ_b = .212). Moreover, those who reported higher SEE had lower weekly sitting (τ_b = -.167). The EBBS total scores were positively correlated with RSE Scale scores (τ_b = .234), HRQoL (τ_b = .339) and with MET-mins/wk (τ_b = .175). Higher self-esteem had a significant association with improved HRQoL (τ_b = .463) and a higher volume of MET-mins/wk was negatively associated with sitting time (τ_b = -.161). These results are summarized in Table 4.5.

Further Kendall's τ_b were completed only for women with PCOS and included their responses to the PCOS-Q domains. There were statistically significant negative correlations between BMI and the Weight (τ_b = -.638), Emotions (τ_b = -.415), Body Hair (τ_b = -.213), and Infertility (τ_b = -.193) domains, whilst there was no statistical correlation between the Menstrual domain and BMI. The Emotions domain was positively correlated with SEE (τ_b = .199), EBBS (τ_b = .338), RSE Scale (τ_b = .436), and SF-12 (τ_b = .506). Although to a lesser extent, the Body Hair (τ_b = .202), Weight (τ_b = .333) and Menstrual (τ_b = .450) domains were positively correlated with SF-12 Total Scores. Three PCOS-Q domains were positively correlated with RSES scores, namely the Weight (τ_b = .319), Infertility (τ_b

= .226) and Menstrual (τ_b = .241) domains. The Body Hair domain also had a statistically significant positive association with the EBBS (τ_b = .204). Of the PCOS-Q domains, only the Emotions domain was associated with Weekly Sitting Time, with a negative correlation (τ_b = -.209) between these two variables. Many of the PCOS-Q domain scores also demonstrated correlation with other domain scores; the Emotions domain showed statistical association with all other domains (Table 4.6), with the strongest association being with the Weight domain (τ_b = .525). The Weight domain was also correlated with both the Infertility (τ_b = .288) and Menstrual (τ_b = .274) domains. The Menstrual and Infertility domains also showed significant correlation (τ_b = .300).

Aside from the PCOS-Q domain scores, a number of other outcomes correlated with each other. Similarly, to the whole cohort analysis, negative correlations remained for BMI with RSE Scale (τ_b = -.253), and BMI with SF-12 scores (τ_b = -.296). SEE was still correlated with EBBS (τ_b = .483), MET-min/wk (τ_b = .373) and Weekly Sitting Time (τ_b = -.217). SEE also showed a moderate correlation with RSES (τ_b = .207), which was not present when the whole cohort was included in the analysis. EBBS remained correlated with RSE Scale (τ_b = .291), MET-min/wk (τ_b = .278) and SF-12 (τ_b = .407). SF-12 total scores also remained correlated with RSE Scale (τ_b = .393). Finally, a moderate negative association remained between MET-min/wk and Weekly Sitting Time (τ_b = -.246).

Table 4.5. Correlation matrix (Kendall's τ_b) for self-reported variables for all study participants.

		Age	Weight	BMI	WC	SEE	Benefit	Barrier	EBBS	SES	SF12 Phys	SF12 Mental	SF12 Total	MET-min/wk	Sitting
Age	τ														
	P														
Weight	τ	.062													
	P	.305													
BMI	τ	.039	.804												
	P	.519	<.001												
WC	τ	.040	.685	.652											
	P	.543	<.001	<.001											
SEE	τ	.031	-.036	-.078	-.069										
	P	.609	.546	.196	.294										
Benefit	τ	.148	-.125	-.169	-.218	.305									
	P	.017	.041	.006	<.001	<.001									
Barrier	τ	.053	-.097	-.141	-.176	.319	.411								
	P	.400	.115	.022	.008	<.001	<.001								
EBBS	τ	.131	-.140	-.188	-.244	.340	.819	.611							
	P	.034	.022	.002	<.001	<.001	<.001	<.001							
SES	τ	.094	-.271	-.211	-.326	.095	.234	.229	.270						
	P	.137	<.001	<.001	<.001	.130	<.001	<.001	<.001						
SF12 Phys	τ	.084	-.288	-.305	-.342	.063	.338	.236	.347	.299					
	P	.196	<.001	<.001	<.001	.331	<.001	<.001	<.001	<.001					
SF12 Mental	τ	.119	-.288	-.252	-.352	.112	.267	.308	.313	.535	.543				
	P	.062	<.001	<.001	<.001	.077	<.001	<.001	<.001	<.001	<.001				
SF12 Total	τ	.103	-.297	-.280	-.365	.083	.303	.305	.339	.463	.737	.844			
	P	.102	<.001	<.001	<.001	.184	<.001	<.001	<.001	<.001	<.001	<.001			
MET-min/wk	τ	.060	.163	.139	.137	.212	.155	.152	.175	.058	.057	.015	.037		
	P	.338	.008	.025	.041	<.001	.014	.018	.006	.369	.395	.815	.573		
Sitting	τ	-.095	.006	-.019	.028	-.167	.009	-.051	-.015	-.042	.000	-.039	-.020	-.161	
	P	.118	.924	.749	.668	.006	.879	.406	.800	.498	.996	.534	.748	.010	

Key: BMI: Body mass index; WC; waist circumference; SEE: self-efficacy for exercise; Benefit: Exercise Benefits Barriers Scale perceived benefits score; Barrier: Exercise Benefits Barriers Scale perceived barriers score; EBBS: Exercise Benefits Barriers Scale total score; SES: self-esteem scale; SF12 Phys: Physical domain scores from the SF-12; SF Mental: Mental domain scores from the SF-12; SF12 Total: Total scores from the SF-12; MET-min/wk: Metabolic equivalent of task minutes per week as measured by the IPAQ-LF; Sitting: minutes per week of self-reported sitting as measured by the IPAQ-LF; τ: Kendall's Tau B; *P*: significance value.

Table 4.6. Correlation matrix (Kendall's τ_b) for self-reported variables for women with PCOS.

		Age	BMI	SEE	EBBS	SES	SF12 Total	Emotions	Body Hair	Weight	Infertility	Menstrual	MET-min/wk	Sitting
Age	τ													
	P													
BMI	τ	-.032												
	P	.740												
SEE	τ	.172	.027											
	P	.079	.776											
EBBS	τ	.181	-.136	.483										
	P	.065	.157	<.001										
SES	τ	.133	-.253	.207	.291									
	P	.182	.010	.035	.003									
SF12 Total	τ	.171	-.296	.123	.407	.393								
	P	.083	.002	.206	<.001	<.001								
Emotions	τ	.181	-.415	.199	.338	.436	.506							
	P	.066	<.001	.041	.001	<.001	<.001							
Body Hair	τ	.089	-.213	.091	.204	.157	.202	.295						
	P	.370	.028	.354	.037	.114	.041	.003						
Weight	τ	.081	-.638	.012	.162	.319	.333	.525	.311					
	P	.420	<.001	.905	.103	.002	.001	<.001	.002					
Infertility	τ	.172	-.193	.006	.167	.226	.179	.370	.051	.288				
	P	.086	.050	.949	.092	.025	.072	<.001	.611	.005				
Menstrual	τ	.155	-.179	-.026	.158	.241	.450	.429	.080	.274	.300			
	P	.118	.066	.794	.108	.016	<.001	<.001	.418	.006	.003			
MET-min/wk	τ	.093	.037	.404	.278	.094	.050	.157	.086	.098	.020	-.044		
	P	.338	.699	<.001	.004	.338	.607	.104	.376	.320	.837	.652		
Sitting	τ	-.136	-.007	-.269	-.113	-.155	-.180	-.209	-.072	-.164	.006	-.185	-.246	
	P	.166	.943	.002	.242	.115	.064	.031	.462	.098	.949	.059	.011	

Key: BMI: Body mass index; SEE: self-efficacy for exercise; EBBS: Exercise Benefits Barriers Scale total score; SES: self-esteem scale; SF12 Total: Total scores from the SF-12; Emotions: Emotion domain scores from the PCOS-Q; Body Hair: Body hair domain scores from the PCOS-Q; Weight: Weight domain scores from the PCOS-Q; Infertility: Infertility domain scores from the PCOS-Q; Menstrual: Menstrual Problems domain scores from the PCOS-Q; MET-min/wk: Metabolic equivalent of task minutes per week as measured by the IPAQ-LF; Sitting: minutes per week of self-reported sitting as measured by the IPAQ-LF; τ: Kendall's Tau B; *P*: significance value.

4.4.3 Path Analysis

The causal path model of this study is presented in Figure 4.3. Both a diagnosis of PCOS and BMI have direct effects on HRQoL (standardised β = .230, and β = .234, respectively) and self-esteem (β = .364 and β = -.213, respectively). Furthermore, PCOS also has the largest indirect (β = .177) and total effect (β =.407) upon HRQoL, although it is closely followed by BMI (indirect β = -.116; total β = -.351, respectively). BMI also had a direct effect on EBBS (β = -.180). Self-esteem (β = .326) and EBBS (β = .254) also demonstrate a direct effect upon HRQoL and although they have no significant indirect effects (via weekly PA as a mediator), both demonstrate a total effect upon HRQoL in this model (β = .324 and β = .265, respectively). EBBS also had a significant total effect on weekly PA (β = .376). Self-efficacy for exercise and weekly PA had no effect (either direct or indirect) on any other variable in the model (Table 4.9).

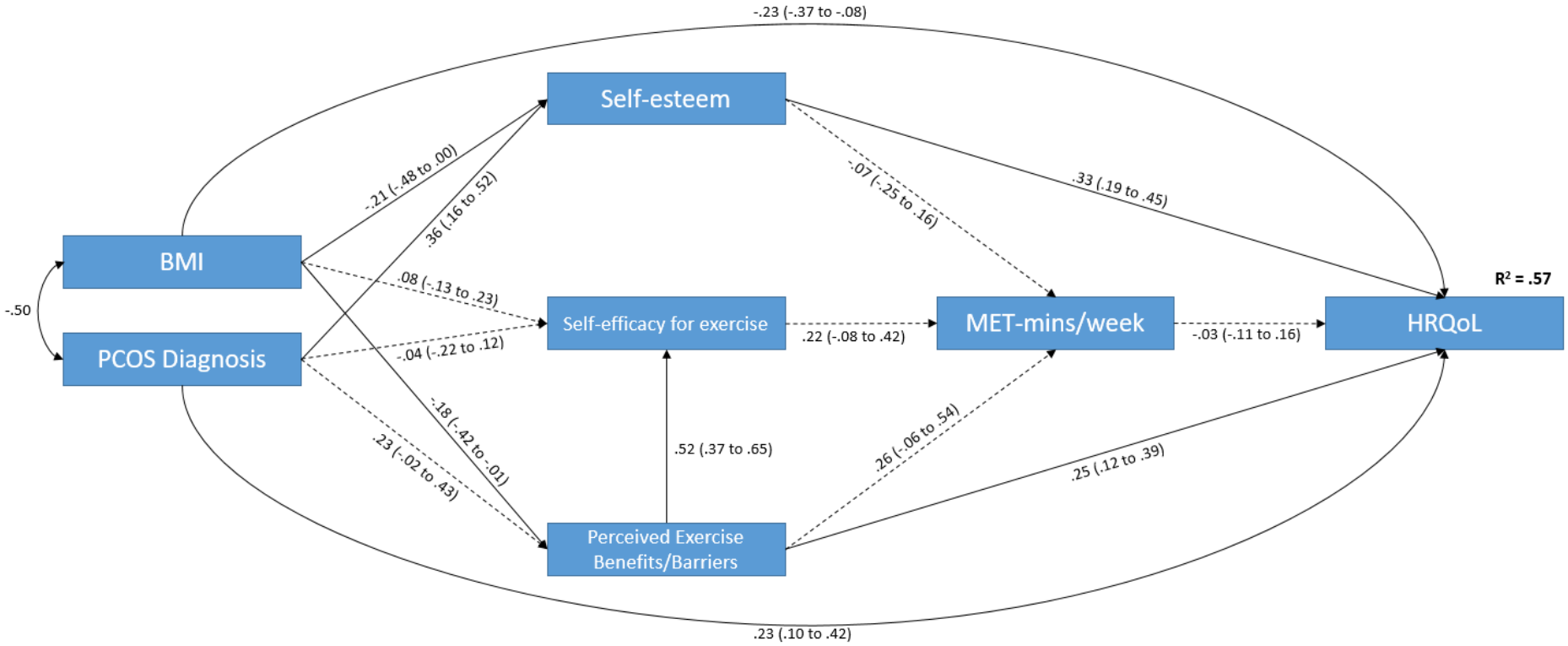


Figure 4.3. Path analysis model based on the findings of this study. Bi-directional arrows indicative of correlation between exogenous variables. Single directional arrows indicate significant standardised path coefficients and bias corrected 95% confidence intervals for direct effects. A dashed arrow indicates non-significant direct effects. All values rounded to two decimal places.

Table 4.9. Direct, indirect and total effects of variables in the HRQoL causal model.

	Standardised Coefficient	95% CI for β		2-tailed Significance
		Lower	Upper	
Direct Effects				
PCOS → HRQoL	.230	.099	.420	**.004**
PCOS → Self-esteem†	.364	.163	.517	**.021**
PCOS → SEE†	-.039	-.220	.116	.517
PCOS → EBBS†	.225	-.022	.426	.068
BMI → HRQoL	-.234	-.368	-.076	**.007**
BMI → Self-esteem†	-.213	-.478	.000	**.050**
BMI → SEE†	.082	-.134	.233	.460
BMI → EBBS†	-.180	-.415	-.009	**.038**
Self-esteem → HRQoL	.326	.185	.445	**.018**
Self-esteem → MET-min/wk†	-.069	-.252	.162	.612
SEE → MET-min/wk†	.224	-.078	.419	.118
EBBS → MET-min/wk	.260	-.055	.542	.119
EBBS → HRQoL	.254	.124	.391	**.010**
EBBS → SEE†	.520	.368	.654	**.006**
MET-min/wk → HRQoL†	.029	-.113	.164	.740
Indirect Effects				
PCOS → HRQoL	.177	.072	.283	**.018**
PCOS → MET-min/wk†	.051	-.057	.177	.305
BMI → HRQoL	-.116	-.266	-.018	**.023**
BMI → MET-min/wk†	-.035	-.154	.059	.519
Self-esteem → HRQoL	-.002	-.035	.008	.626
SEE → HRQoL†	-.006	-.026	.038	.411
EBBS → MET-min/wk	.117	-.006	.285	.066
EBBS → HRQoL	.011	-.038	.093	.684
Total Effects				
PCOS → HRQoL	.407	.255	.571	**.011**
BMI → HRQoL	-.351	-.507	-.194	**.005**
Self-esteem → HRQoL	.324	.196	.446	**.012**
EBBS → HRQoL	.265	.127	.407	**.012**
EBBS → MET-min/wk	.376	.130	.586	**.015**

Key: PCOS: Diagnosed with PCOS; HRQoL: Health-related Quality of Life as measured by total SF-12 score; SEE: self-efficacy for exercise; EBBS: Exercise Benefits/Barriers Scale; MET-min/wk: Metabolic equivalent of task minutes per week as measured by the IPAQ-LF; 95% CI: Bias corrected 95% confidence interval; → denotes causal direction of path; † data reported are equal to Total Effects for path.

4.5. Discussion

4.5.1. Key Findings within the study cohort and the PCOS study group

The present findings show that despite no statistical differences in self-reported PA or sitting time, women with PCOS had greater body weight, BMI and waist circumference than women without PCOS. Moreover, women with PCOS also demonstrated lower physical, mental and total domain scores for HRQoL than their non-PCOS counterparts, whilst also reporting lower self-esteem. When determinants of PA were considered, women with PCOS perceived fewer benefits and more barriers to exercise than women without PCOS, but there were no statistical between group differences for self-efficacy for exercise. In addition, for the whole cohort higher PA was, unsurprisingly, associated with greater self-efficacy for exercise and higher EBBS scores. In this context, what was less expected was the positive association between body mass and anthropometric outcomes (*i.e.,* body weight, BMI and waist circumference) and MET-mins performed each week. However, PA-levels were not associated with SF-12 results or self-esteem. In contrast, BMI had a statistical negative association with HRQoL, self-esteem and total EBBS scores (*i.e.,* more perceived barriers and fewer perceived benefits). Increased self-esteem was positively associated with HRQoL and this effect was greater in the mental health domain. When determinants of PA were considered, EBBS scores were associated with self-efficacy for exercise, HRQoL and self-esteem. Apart from its association with EBBS, self-efficacy for exercise was only associated with one other variable, exhibiting a negative association with sitting time.

Overall, the path analysis in this study indicated that both PCOS diagnosis and participant's BMI had statistical direct, indirect and total effects upon HRQoL, with the former having the largest total effect. Interestingly, self-esteem levels had the largest direct effect upon HRQoL, and also constituted an important mediator between PCOS/BMI and HRQoL.

In the analyses only for the group of participants with PCOS, there were no statistical differences between the subgroups for each self-reported PCOS phenotype. It was perhaps surprising to note that ~17% of these women were unsure regarding the PCOS phenotype which characterized their PCOS diagnosis. Regarding the PCOS-Q domains, the weight domain was scored statistically lower than

all other domains, indicating that this represents the largest concern for women with PCOS in this study cohort. Of note, the Weight, Emotions, Body Hair and Infertility PCOS-Q domains were all negatively associated with BMI. Furthermore, HRQoL (total SF-12 score) was correlated with all PCOS-Q domains apart from Infertility. Interestingly, self-esteem was correlated with the Weight, Infertility and Menstrual domains. The Emotions domain was also positively correlated with self-efficacy for exercise, EBBS and self-esteem, but showed a negative association with sitting time, whilst EBBS scores correlated with the body hair domain. In addition to PCOS-Q domain scores, the associations between PA-levels and both self-efficacy for exercise and EBBS scores remained, and were in fact stronger for the former, in the PCOS study group, whilst a negative association between PA and sitting time also remained. Similarly, BMI still exhibited a negative association with both self-esteem and HRQoL, and self-efficacy for exercise with EBBS score and sitting time. The results within the PCOS study group varied from the whole cohort analysis by demonstrating a moderate correlation between self-esteem and self-efficacy for exercise.

4.5.2. Between group differences

Regardless of PCOS phenotype, it is clear that the symptomatic manifestations of PCOS may impact upon the HRQoL of a woman with PCOS. Indeed, it has been widely reported that women with PCOS demonstrate reduced HRQoL compared to healthy control groups or normative population data (Barnard *et al.*, 2007; Ching, Burke and Stuckey, 2007; Coffey, Bano and Mason, 2006; Hahn *et al.*, 2005; Jones *et al.*, 2010; Li *et al.*, 2011). The findings reported here tend to agree with these previous studies, with statistically significant differences noted for the Physical, Mental and Total scores from the SF-12. Despite statistical differences, the degree of clinical importance associated with this is less straightforward to determine. To the author's knowledge, this is the first study that has used the SF-12 to assess HRQoL in women with PCOS, so there are no normative data with which to compare. However, Costa *et al.* (2018) reported change in the scores for each domain of the SF-36 before and after an aerobic exercise intervention. Whilst it is not possible to compare these scores directly, the effect size (*d*) of the most relevant domains can be compared. In the present study,

the effect size for between group differences for the SF-12 Physical Domain (d = -0.999) is comparable to the SF-36 Physical Functioning (d = 1.2, 95% CI: 0.5 to 1.9) and Role Physical (d = 1.0, 95% CI: 0.1 to 1.8) change from baseline effect sizes reported in Costa *et al.* (2018). Using magnitude-based inference (Hopkins *et al.,* 2009) they state that the magnitude of effect sizes for Role Physical 'Very Likely', and Physical Functioning 'Almost Certainly' represent clinically significant changes. Similarly, they report a 'Likely' chance that the Role Emotional (d = 0.8, 95% CI: 0.0 to 1.6) and Mental Health (d = 1.0, 95% CI: 0.1 to 1.8) domain scores are clinically important; the magnitude of the effect size they report is smaller than that reported in the present study (d = -1.346), suggesting that the observed differences in perceived mental health between women with PCOS and controls may have clinical relevance.

However, there is some contrasting evidence presented in a study of non-PCOS participants (those who had received anterior cervical discectomy and fusion) that did use the SF-12 to evaluate HRQoL (Parker *et al.,* 2013). Using the minimum detectable change (MDC) method, they do quantify minimal clinically important differences for domain scores of the SF-12; MDC defines clinically important differences as the smallest change which can be observed above the measurement error with a given level of confidence (*e.g.,* 95% CI). Therefore, the minimal clinically important difference is equal to the upper value of the 95% CI (Copay *et al.,* 2007). Parker *et al.* (2013) identify the threshold for clinically important change as ±8.1 points for the SF-12 Physical domain and ± 4.7 points for the mental domain scores. Although caution should be adopted because of the variation in different populations, the implication in the current study is that women with PCOS have statistically and potentially clinically lower levels of perceived mental health than women without PCOS. Furthermore, whilst statistically lower physical wellbeing scores were observed in women with PCOS, the corresponding degree of clinical importance is unclear. For greater confidence in these findings, future work should focus upon identifying the HRQoL thresholds for clinically important changes in women with PCOS.

The present findings also document that the SF-12 mental health domain was the most severely impacted domain by a diagnosis of PCOS. Notably, there was a statistically significant difference

between mental and physical domain scores (mean difference = 2.67, 95% CI: 1.74 to 3.60; P <.001) in women with PCOS that was not present in the control group (mean difference = 0.60, 95% CI: -0.25 to 1.45). This suggests that, according to the SF-12, PCOS has a greater psychological than physical effect. Indeed, these findings agree with those reported by Bazarganipour *et al.* (2013), although the SF-36 domain scores were utilised in this study. This underlines the importance of incorporating assessments of psychological wellbeing, and the appropriate treatment strategies during the management of PCOS. When differences between phenotypic subgroups were evaluated, there were no statistical differences for either SF-12 domain.

Aside from the SF-12, there were also additional differences between women with PCOS and the control group, with women with PCOS reporting higher body mass, BMI and waist circumference than the control group. Across the UK general population, obesity levels have increased from 15% in 1993, to 27% in 2015 (Moody *et al.*, 2016). Specifically among UK women, a 31% obesity prevalence has been reported, with the highest rates in the 55-64 years age group (~34%), and considerably lower rates (~22%) in younger age groups (*i.e.*, 16-44) (Moody *et al.*, 2016). The association between obesity and PCOS has been widely reported, with older data indicating that the obesity prevalence in UK women with PCOS was at 35-38% (Kiddy *et al.*, 1990; Balen *et al.*, 1995). Given the increasing obesity prevalence rates in general population over the past few decades, it is a reasonable assumption that a similar growth may have been observed in women with PCOS. Indeed, in the current study ~63% of women with PCOS had a BMI ≥30 kg/m^2 (compared to ~17% in the control group). This finding is further supported by a North American review (Yildiz *et al.*, 2008) which found that the temporal trends of obesity prevalence in women with PCOS appear to reflect the increases in obesity prevalence in the general American population (obesity prevalence in women with PCOS increased from 51% in 1987 to 74% in 2002).

BMI may provide a standardized index to evaluate an individual's weight status, but it does not distinguish between fat, muscle or bone mass, whilst it also makes no inference about body composition, or the distribution of adipose tissue. Although the current study did not assess body composition, data on self-reported waist circumference were captured as an index of central

adiposity. As expected, median waist circumference was statistically higher in women with PCOS. Of note, based on the waist circumference cut-off points outlined in clinical guidelines (NICE, 2014) to identify women at a high risk of developing health problems (waist circumference ≥80 cm, and further increased risk for waist circumference 88 cm), most of the control study group self-reported as being under this threshold (mean: 77.2 cm; median: 74 cm), but most study participants with PCOS far exceeded these waist circumference cut-off points (mean: 101 cm; median: 101.6 cm). This indicates that there are both statistical and clinically meaningful differences regarding waist circumference - and so central adiposity - between the two study groups. Interestingly, it has also been reported that the distribution of adipose tissue in PCOS tends to be more central (Talbott *et al.*, 1995; Taponen *et al.*, 2003), thus further promoting metabolic disturbances (Pasquali, 2006) and increasing the risk of metabolic syndrome and type 2 diabetes (Kousta and Franks, 2007). Whilst instances of metabolic disturbances are beyond the scope of the current study, the degree of difference in waist circumference certainly tends to agree with previous findings, suggesting that these women with PCOS are at an increased risk of cardio-metabolic disease.

There were no statistical differences in the amount of self-reported PA (MET-mins/wk) or sitting time between women with PCOS and the control group. This result partially agrees with previous findings. Moran and colleagues (2013) analysed data from a large cohort study and found that there were no differences in total PA levels between women with PCOS and controls, whilst this tends to agree with the findings of others too (Douglas *et al.*, 2006; Wright *et al.*, 2004; Alvarez-Blasco *et al.*, 2011; Cutler, Pride and Cheung, 2018). Similarly, a more recent study (Lin *et al.*, 2019) also reported no statistical differences of PA levels between women with PCOS and a control group. In fact, in this study there were no differences ($P \geq 0.14$) in the duration, type or intensity of PA regardless of the measurement methods used (*i.e.*, accelerometry or self-report). This implies that PA-levels may not be a key contributing factor to the widely reported increased body weight and metabolic complications associated with PCOS. Indeed, both Moran *et al.* (2013) and Lin *et al.* (2019) reported less favourable values in women with PCOS for body weight and BMI, despite no differences in PA. Considering this, it is possible that these disparities may be due to differences in

dietary intake. Although a dietary analysis was beyond the scope of the present study, it has been investigated in numerous previous studies, with Moran *et al.* (2013) reporting higher daily energy intake (but better dietary quality) in women with PCOS (P = .02). Contrastingly, Lin *et al.* (2019) cite no difference in daily kcal consumption (P = .64) or key macronutrients between these populations, which also agrees with previous studies (Douglas *et al.*, 2006; Wright *et al.*, 2004; Alvarez-Blasco *et al.*, 2011; Cutler, Pride and Cheung, 2018). In fact, a systematic search of the relevant literature identified only two additional studies in which energy intake was reportedly higher in women with PCOS, with Zhang *et al.* (2015) reporting that Southwest Chinese women with PCOS (n = 169) had a higher energy (KJ) and fat intake, but consumed fewer carbohydrates than their age-matched non-PCOS counterparts (n = 338), whilst Ahmadi *et al.* (2013) have also reported similar findings with women with PCOS consuming more calories (P = .001) and fat (P = .019) than women without PCOS.

As stated earlier, the present study observed no differences in weekly sitting time between women with PCOS and women without. To the best of the author's knowledge, only one study has reported findings that contradict this result. Moran *et al.* (2013) report increased sitting time in women with PCOS compared to a control group (6.3 ± 2.8 *vs* 5.8 ± 2.9 hr/day; P = .008). Conversely, there is a significant body of literature that agrees with the present findings and reports no statistical differences in sitting time (Ahmadi *et al.*, 2013; Wright *et al.*, 2004; Alvarez-Blasco *et al.*, 2011; Lin *et al.*, 2019). Variation in these factors (PA-levels, diet and sitting time) which are established contributors to energy balance may well have a role in weight and anthropometric differences between women with and women without PCOS. However, given the contrasting findings of the evidence to date (including the present study) it is difficult to draw definitive conclusions.

The current study also revealed that self-esteem, as measured by the Rosenberg Self-esteem Scale, was statistically lower in women with PCOS than in the control group. This is in contrast to a previous study (Annagür, Tazegül, & Akbaba, 2014) which reported no statistical difference in self-esteem between women with PCOS and a control group. This may be attributed, at least partly, to the group characteristics in this study, since the mean BMI of both study groups was <24 kg/m^2. In

contrast, the women with PCOS in the present study were mostly overweight or obese. Indeed, another study (Acmaz *et al.*, 2013) reported significantly lower (P <.001) self-esteem across subsets of women with PCOS (*i.e.*, subgroups with hirsutism, infertility and obesity) compared to healthy controls, but it was those with obesity (BMI ≥ 30 kg/m^2) and PCOS who were far more likely to have the lowest self-esteem. This also tends to reflect what has been widely reported in the general population, with Strauss (2000) reporting that Caucasian females with obesity demonstrated lower levels of self-esteem than their non-obese counterparts. Another longitudinal study (n ~2400) followed girls from adolescence to 22 years and found that BMI was an important predictor of self-esteem (Biro *et al.* 2006), whilst an Australian study also found that overweight/obesity preceded poor self-esteem in a mixed-gender cohort of >1100 participants (Hesketh, Wake and Waters, 2004).

Despite the apparent relationship between excess body weight and self-esteem, there is a dearth of literature on this topic in women with PCOS. Many women with PCOS reportedly have poorer self-esteem compared to their healthy counterparts, but it is unclear whether the sequelae of PCOS is responsible for poorer self-esteem in these women, and also whether the excess weight gain can be attributed to an effect that low self-esteem and depression may have upon diet (*e.g.*, emotional eating) and a reluctance to exercise (Elsheik and Murphy, 2008). Alternatively, it is also likely that the genetic predisposition to obesity in PCOS, and the subsequent weight gain, are contributing factors to poorer self-esteem and that the clinical manifestations of PCOS, which are seemingly worsened by obesity, further exacerbate this problem. A key question to address then surrounds the impact of weight loss upon self-esteem and whether a concomitant improvement in physical manifestations of PCOS is also a driving factor in improved self-esteem.

A recent study in a predominantly female population (Borbón-Castro *et al.*, 2019) has reported the effectiveness of exercise, compared to control, at improving self-esteem (post-intervention mean difference: P = .005). Within PCOS cohorts, previous studies have also evaluated the effect of exercise/lifestyle interventions on self-esteem, with reported findings suggesting a similar effect. Indeed, a study in women with PCOS who completed a 6-month lifestyle intervention (incorporating aerobic exercise and gradual dietary changes) reported improved self-esteem (Clark *et al.*, 1995).

What is less clear is whether these observed improvements were attributable to the undertaking of an exercise intervention or the resultant weight loss (mean change from baseline: -6.3 ± 4.2 kg; P <.001). Two additional studies also report similar findings (Clark *et al.,* 1998; Galletly *et al.,* 1996). As such, a larger cohort of women with obesity and fertility problems (including women with PCOS) demonstrated statistical improvements in self-esteem following a 6-month exercise intervention (Clark *et al.,* 1998). Concomitantly, the women completing this intervention also reduced their BMI (-3.7 ± 1.6 kg/m^2; P <.001). Galletly *et al.* (1996) also documented improvements in self-esteem alongside weight loss in a cohort of infertile women following an exercise and dietary advice group intervention. Given the scarcity of literature exploring the role of exercise upon self-esteem in PCOS, and the uncertainty surrounding the role of adiposity, future studies should further explore the relationship between self-esteem and body weight.

Despite there being no statistical differences in self-reported PA between the women with PCOS and the control group, it is evident that women with PCOS perceive both fewer benefits (d = -0.595; P = .001), and a greater number of barriers (d = -0.482; P = .003) to participation in exercise. Contradictory to the findings of the current study, but rather unsurprisingly, it has been reported in non-PCOS individuals that those who perceive greater benefits and fewer barriers to exercise participation exercise more regularly (Grubbs and Carter, 2002) and tend to be more physically active in comparison to those who perceive less benefit and more barriers (Bonheur and Young, 1991; Jones and Nies, 1996). Indeed, it has been previously reported that perceived barriers are the most powerful predictor of a health behaviour (Janz and Becker, 1984). Although there is mixed evidence as to whether women with PCOS are less active than their non-PCOS counterparts, few studies have investigated the relationship between barrier/benefit perception and PCOS. One such study compared women with PCOS to controls and found many similarities between groups (Banting *et al.,* 2014); both groups cited ‘lack of time’, ‘fatigue’ and ‘not confident can maintain’ as the top three barriers to participation in exercise. However, following Chi-squared goodness of fit analyses, they found that a statistically significant greater proportion of women with PCOS cited a lack of confidence to maintain (χ^2 = 3.97; P = .046), fear of injury (χ^2 = 4.08; P = .043) and physical limitations (χ^2 =

11.92; $P = .001$) when compared to women without PCOS. A more recent study (Thomson, Buckley and Brinkworth, 2016) reported that the most common perceived barriers to PA were those related to physical exertion (*i.e.*, exercise is tiring, hard work and fatiguing) and these findings exactly match the top cited barriers identified in this study. However, Thomson and colleagues (2016) further stated that exposure to a lifestyle intervention may improve these perceptions; total EBBS and barriers only scores improved 10-weeks into a lifestyle intervention (either diet only, diet and aerobic exercise, or diet combined with resistance and aerobic exercise), with no further statistical improvements over the final 10-weeks. Commencement of a new exercise regime is often perceived as fatiguing, particularly to those leading a previously sedentary lifestyle, and this may explain why the perceived barriers were reduced in the initial 10-weeks ($P \leq .001$), but not during the subsequent 10-20 week period ($P = 1.0$).

When perceived benefits were considered, slight differences between the top three cited benefits were reported between women with PCOS and controls; all three in PCOS focussed on improving fitness (*i.e.*, exercise improves physical fitness, flexibility and cardiovascular fitness), whereas the control group did not mention flexibility, but cited that '*exercise gives me a sense of personal accomplishment*' instead of flexibility. Despite using different methods, Banting *et al.* (2014) found no differences between women with PCOS and controls for the top five exercise motivators (1: weight control; 2: health improvements; 3: increased energy; 4: stress reduction; and 5: health maintenance), or the number of motivating factors reported by each cohort (PCOS: 8.51 ± 3.27; Control: 8.27 ± 2.68). In fact, it was only '*controlling a medical condition*' that was more commonly cited as an exercise motivator by women with PCOS (40% vs 19%, $P = .006$) which is unsurprising as women with PCOS were compared to healthy controls. When Thomson and colleagues (2016) measured changes to the perceived benefits in response to a lifestyle intervention, they report findings that contrasted those of perceived barriers. Whereas perceived barriers were reduced in the first 10-weeks of an intervention, there was no corresponding improvement in perceived benefits ($P = .90$). However, there were statistical improvements from weeks 10-20 ($P = .003$) for perceived benefits; it is suggested that there may be a phase, during the early stages of a lifestyle regime where the

individual is becoming accustomed to the changes, and a period of adaptation before the benefits become apparent to the individual (Thomson *et al.*, 2016). The implication of this is that in order to improve PA levels in PCOS, an initial focus should be placed upon removing barriers to participation; PA and exercise do not influence the existence of barriers, but those who are physically active appear to be more capable of overcoming them (Van Zanten *et al.*, 2015). As the intervention progresses, there should be a change of focus to targeting the perceived benefits in order to promote long-term exercise adherence (Pentecost and Taket, 2011).

Whilst poor self-efficacy itself is not a barrier to behaviours, it reportedly influences the chosen activity (*i.e.*, PA), the effort expended within those activities and the degree of persistence when faced with adverse events or barriers (Bandura, 1977; Bandura, 1986). Accordingly, self-efficacy for exercise has been shown to be a significant predictor of both exercise adoption and exercise regime adherence (Sallis *et al.*, 1986). Furthermore, there is a reciprocal relationship between self-efficacy and the activity. For example, increased self-efficacy may contribute to the uptake of an exercise programme, but a negative experience may serve to lower self-efficacy, which in turn may cause early attrition and resistance to future participation (McAuley and Jacobson, 1991). The current study found no statistically significant differences in SEE between PCOS and control groups, but also no difference in self-reported PA or sedentary activities. Although this does suggest that there are no differences between the two groups in this study, it does not mean that SEE does not contribute to PA behaviours. Whilst not direct measures of SEE, Banting *et al.* (2014) report that women with PCOS cited a lack of confidence to maintain an exercise programme, a greater fear of injury, and their own physical limitations as barriers to participation, more frequently than their healthy counterparts. SEE was not measured in the study by Banting *et al.* (2014), but it could be argued that these traits are strong examples of low confidence in their own ability to perform PA.

4.5.3. PCOS-Q domains

Based on the data on the five PCOS-Q domains, the largest HRQoL concern of the women with PCOS in this study relates to weight, with the weight domain score being significantly lower than all

other domain scores. This is further supported by a strong negative correlation between Weight domain scores and self-reported BMI, indicating that the higher the BMI the greater weight-related concerns (*i.e.,* lower scores) were for this cohort. This finding has also been documented in previous studies (McCook, Reame and Thatcher, 2004; Coffey, Bano and Mason, 2006). However, other studies have shown various domain scores as having the greatest impact upon HRQoL. For example, baseline data from a randomised controlled trial (RCT) reported that either the Weight or Infertility domains had the lowest mean score depending upon treatment arm (Dokras *et al.,* 2016), whilst another study found the Infertility domain score to be the lowest (followed by weight) in two treatment arms (Vizza *et al.,* 2016). Another RCT reported the lowest baseline data for the Body Hair domain in two out of three treatment arms (Stener-Victorin *et al.,* 2013), with the Weight domain being the lowest in the third study arm.

Although studies utilising the PCOS-Q are limited, a number of studies in non-PCOS populations that have used the SF-36 to assess HRQoL in patients with obesity. In this context, individuals with obesity undertaking a weight loss programme in the United States scored significantly lower across all domains of the SF-36 when compared to normative population data (Fontaine, Chesky and Barofsky, 1996). A similar study in Australia also reported that subjects with obesity scored significantly lower on the SF-36 Physical and Emotional domain scores when compared to age-matched normative data (Anandacoomarasamy *et al.,* 2009), whilst they also reported significantly lower Assessment of Quality of Life (AQoL) scores and significantly higher fatigue (as measured by the Multidimensional Assessment of Fatigue). Of note, it has also been previously reported that HRQoL is improved following a small to moderate weight loss (via non-surgical methods) (Rippe *et al.,* 1998).

A landmark study in women with PCOS (Balen *et al.,* 1995) reported that an increased BMI was associated with increased rates of infertility and menstrual disruption, as well as with elevated serum testosterone levels and hirsutism prevalence. Furthermore, whilst women with PCOS are at a greater risk of metabolic abnormalities (Azziz *et al.,* 2016), it has also been reported that obesity further

exacerbates metabolic complications (Sam, 2007). Although these symptoms are not direct measures of HRQoL, it is reasonable to suggest that these may also have a detrimental effect on both the physical and emotional components of HRQoL. The reduced self-esteem of women with PCOS in the current study, alongside significantly elevated BMI, also suggests that there may be a causal relationship between these two outcomes.

4.5.4. Correlations

The current study revealed no statistical relationship between total SF-12 scores and PA for women with PCOS, or the combined cohort. Similarly, there were also no significant relationships between the Mental and Physical domains from the SF-12 and PA performed. Furthermore, when the domain scores from the PCOS-Q were considered, there was no evidence of a relationship between PA and any of the domains; the author's recent systematic review (Kite *et al.*, 2019) also reported no statistical benefit, following exercise interventions, upon any domain of the PCOS-Q. The current study also found no relationship between HRQoL and weekly sitting time. Although no studies have assessed the role habitual PA has upon the HRQoL of women with PCOS directly, Banting *et al.* (2014) reported on PA status for some relevant outcomes, which tend to contradict the findings of the current study. Indeed, Banting and colleagues (2014) have found that both anxiety and depression (as measured by the Hospital Anxiety and Depression Scale) were significantly lower in physically active women with PCOS compared to inactive women with PCOS. Similarly, Lamb *et al.* (2011) found that inactive women with PCOS presented with mild depression, whereas their active counterparts did not (BDI: 5.5 ± 4.4 vs 3.6 ± 3.6; $P = .005$). This study also reported data on numerous cardio-metabolic outcomes (BMI, waist circumference, weight fluctuations, and fasting glucose) which all statistically favoured those who were active.

It is not possible from the current results to state that women with PCOS (or in fact healthy women) have reduced symptoms and improved quality of life if they are more physically active. To date, there is some evidence that exercise interventions can contribute to perceived improvements in HRQoL; however, the quality of this evidence is limited and the clinical importance of these changes

uncertain (Kite *et al.,* 2019). What is more widely known is that increasing PA levels in the general population lowers the risk of morbidity. For example, Kyu *et al.* (2016) meta-analysed data from 174 studies (149.2 million total person years of follow-up) and reported that higher levels of PA are associated with lower risk for five common chronic diseases, including type 2 diabetes, ischaemic heart disease and stroke. Although reporting incidence of disease is not a direct measure of HRQoL, it is previously reported that those with symptoms deemed as risk factors for chronic disease (Fontaine, Chesky and Barofsky, 1996; Anandacoomarasamy *et al.,* 2009), or those living with chronic disease (Safita *et al.,* 2016; Westin *et al.,* 1998) have markedly lower HRQoL than their healthy counterparts. The present study also indicates that women with PCOS have poorer HRQoL than controls, but the relationship with PA is less clear. This represents an under-researched area, with existing studies utilizing mostly self-report measures of PA and no direct measure of HRQoL. Future work should therefore employ objective measures of PA alongside well-established and validated measures of HRQoL.

When correlation analyses were completed using all study participants, there were some statistically significant positive associations between the self-reported PA (MET-mins/wk) and Weight (τ = .163), BMI (τ = .139), waist circumference (τ = .137), SEE (τ = .212) and EBBS (τ = .175), albeit with a small magnitude (Cohen, 1988). The anthropometric findings are somewhat surprising, since, typically, it would be expected that those performing more PA would have lower body weight and waist circumference (Hu *et al.,* 2004; Mario *et al.,* 2016), and, if the PA type completed was mainly aerobic in nature, also lower BMI (Kite *et al.,* 2019). When the study groups are split, associations between PA with weight and BMI only remain in the control group (τ = .262 and τ = .178, respectively) and the effect is slightly increased. Waist circumference ceases to have statistical association in either group. These findings make the data difficult to interpret and presents a potential limitation of this study, since it is possible that waist circumference could have been incorrectly measured and also that participants may have over reported their weight or height. It is also likely that participants have overstated their PA levels. Indeed, Prince *et al.* (2008) compared self-report measures with objective measures of PA in large meta-analyses (173 studies), showing that, whilst

there were no clear trends in the over- or under-reporting of PA, females self-reported higher levels of PA compared to accelerometers (mean percent difference = 138%). Of note, these meta-analyses also revealed that self-report measures that categorise PA by level of exertion (such as the IPAQ-LF) led to the largest total over-report, largely due to over-reporting in the higher intensity categories (*i.e.*, vigorous PA). Another study (Ferrari, Friedenreich and Matthews, 2007) also found that self-report data from females was less accurate when compared to accelerometry, but, in addition, the greater an individual's BMI, then the less accurate their responses were. This may partially explain the surprising effects of increased weight and BMI being associated with increased PA that were observed in the current study. Furthermore, the self-reported MET-mins/wk in the current study appear to be considerably higher than both the current WHO (2010) PA guidelines (600 MET-mins/wk) and the corresponding female normative data from the UK (Love-Koh and Taylor, 2018). Thus, it is likely that in the current study a response bias, potentially due to social desirability, has influenced the degree of over-reporting, particularly in those women with increased BMI (Prince *et al.*, 2008).

Further highlighting the importance of SEE in the uptake and maintenance of exercise/PA, there was a small ($\tau < .3$), but statistically significant, positive correlation between SEE and MET-mins/wk ($\tau = .212$; $P < .001$), and a negative correlation between SEE and weekly sitting time ($\tau = -.167$; $P = .006$) when all participants were included in the analysis. Furthermore, the magnitude of each relationship was greater in only the PCOS group (MET-mins/wk $\tau = .404$; $P < .001$, and sitting time $\tau = -.269$; $P = .002$) than in the whole group. Neither of these associations were present when the control group was analysed independently. Although participation in PA can be attributed to a number of environmental, physical, social and psychological factors, an individual's belief (*e.g.*, self-efficacy for exercise) in his/hers capabilities to perform physical activity is consistently cited as a predictor of exercise adherence and compliance (McAuley and Blissmer, 2000). Notably, previous research indicates that it may be only in the early stages of an exercise programme (*i.e.*, the adoption phase) where self-efficacy levels successfully predict the behaviour (Oman and King, 1998). Another study reported that SEE was an effective predictor of long-term adherence to home-based exercise

programmes, suggesting that self-efficacy could be targeted in interventions, alongside barrier removal, to improve adherence (King *et al.,* 1995). The contrasting viewpoint is that engagement in PA itself promotes improved SEE; acute bouts of exercise typically improve efficacy, but prolonged exposure to PA is where the greatest improvements are seen (McAuley, Lox and Duncan, 1993). The implication here is that, without knowing how participants in the current study were engaging in PA, and at what stage they were at in their PA lives, it is difficult to determine the effect of improved self-efficacy on PA levels. Equally, it is difficult to hypothesise whether the causal relationship is reversed, and that it is engagement in PA which is promoting self-efficacy levels, especially since there were no statistical differences between either outcome in the PCOS or control groups.

A statistical correlation was also observed between MET-mins/wk and EBBS when all participants were included in the analysis, with the magnitude being smallest ($\tau = 0.175$; $P = .006$) with small increases when only women with PCOS ($\tau = 0.278$) or the control ($\tau = 0.290$) were considered. When Benefits and Barriers were considered separately, there were statistically significant correlations of small magnitude for PCOS, control and all participants for each domain. This indicates that more PA is performed (and so self-reported for the last seven days) if participants perceive greater benefits and fewer barriers to exercise. Theoretical models, such as the Health Belief Model (Rosenstock *et al*., 1988), purport that an individual simultaneously evaluates the perceived exercise benefits and barriers associated in engaging with PA before making what they consider to be an informed decision whether or not to engage. This theory appears to be supported by the findings of the current study. Whilst the current study reports that women with PCOS perceive statistically fewer benefits and greater barriers to PA, the benefits and barriers with the highest mean score are similar across groups, with the top three barriers relating to fatigue and the top three benefits mainly related to improving fitness. These findings tend to be consistent with previous findings both in women with PCOS (Thomson, Buckley and Brinkworth, 2016) and in healthy participants (Chung-Yan Chan, 2014). As such, Chung-Yan Chan (2014) reports that both perceived benefits and perceived barriers to exercise were independent predictors of exercise behaviour, and that self-efficacy for exercise is a predictor of physical activity in healthy participants. Whilst the current study revealed a statistical association

between SEE and MET-mins/wk, there was a statistical correlation of greater magnitude between SEE and EBBS (τ = .340; P <.001) and this relationship was further enhanced when only women with PCOS (τ = .483) were included in the analysis. Indeed, SEE is widely reported to improve PA behaviours (Higgins *et al.,* 2014; Schwarzer, *et al.,* 2008) and the observed relationship between SEE and EBBS may play a pivotal role in contributing to PA. Bandura (1997) states that if an individual interprets physiological and psychological feedback positively, self-efficacy can be enhanced. If this is true, it can be assumed that the inverse may also be true; therefore, the three highest scored barriers which all relate to exercise being hard work, fatiguing and tiring, all likely contribute to decreased SEE and hence reduce PA levels.

Finally, a notable correlation between self-esteem and the total and domain scores from the SF-12 was also noted in the present study. In the whole study cohort, the magnitude of association between the total SF-12 and self-esteem (τ = .463) suggests that self-esteem may have an important role in HRQoL. Indeed, it is possible that that the causal effect may be reversed (*i.e.*, improved HRQoL promotes higher self-esteem), but this effect can be further explored. When domain scores from the SF-12 are correlated with self-esteem, a large (τ = .535) magnitude is evident with the Mental domain which is far less pronounced in the Physical domain (τ = .299). Self-esteem has previously been identified as a major factor that impacts upon HRQoL, with Bazarganipour *et al.* (2014) stating that self-esteem, alongside body image and sexual function, significantly impaired HRQoL. However, using structured equation modelling, it was suggested that it was a greater severity of PCOS symptoms that promoted decreased self-esteem leading to poorer HRQoL. Although physical measurements in the current study were not taken, when PCOS-Q domain scores were correlated with self-esteem, it was found in all domains that the more severe participants perceived their symptoms (*i.e.,* lower domain scores), the lower their self-esteem. This was observed with increased body weight and BMI also having statistically negative associations. Whilst increased body weight/BMI may be a driver of self-esteem, it is unlikely that it explains all of the variance. Miller and Downey (1999) meta-analysed data on this topic, revealing only a modest relationship between the two (d = -0.36, 95% CI: -0.33 to -0.40), although this did increase as individuals aged. Due to the

nature of PCOS symptoms, it is likely that the severity of individual's physical manifestations directly contributes to decreased self-esteem which in turn indirectly reduces HRQoL.

4.5.5. Path Analysis

To the best of the author's knowledge, this is the first study that has attempted to quantify the role of PA on HRQoL within PCOS, whilst simultaneously examining the relationship between exogenous variables, determinants of PA and their direct/indirect effects upon each other within a causal model. The use of path analysis in the study of chronic disease (including PCOS) is rare. Path analysis is a multivariate technique that is both exploratory and confirmatory, allowing relationship testing between multiple constructs within a theoretical framework. As such, it has been cited as a suitable statistical approach for examining associations within complex, heterogeneous diseases, such as PCOS (Bazarganipour *et al.,* 2013).

Here, the path analysis model showed little evidence that PA had any influence upon HRQoL. Indeed, no single outcome had any direct effect upon the amount of self-reported PA, and similarly PA did not have a direct effect on SF-12 scores. A statistical total effect of perceived exercise benefits/barriers upon MET-mins/week is reported, but this is likely due to the strength of the direct effect (standardised β = .52, 95% CI: .37 to .65) of EBBS upon self-efficacy for exercise (as a mediating variable). This relationship was the only instance where an outcome was deemed to have any effect upon PA levels. There may be no legitimate relationship between PA and HRQoL, but having previously noted some of the limitations with self-reported data, particularly that associated with PA, it is difficult to definitively state this. In this context, it is of note that the MET-mins/wk data generated from participant's responses to the IPAQ in this study is certainly higher than UK population normative data. For example, using participant data from the Health Survey England (2014), Love-Koh and Taylor (2018) report gender specific MET-mins/wk data for different age groups (*i.e.,* 16-24, 25-34, 35-44, *etc*.). Utilising MET calculations derived from the same source as the IPAQ (Ainsworth *et al*., 2011), the normative data they report for the most active female group (16-24 years) is considerably lower than that observed in the present study (mean: 2854 versus 3953

MET-mins/wk). It is of course possible that the participants within the current study are indeed very active. However, there are also a number of well documented limitations with self-reported PA data which may have influenced the results in the present study. Recalling PA is a highly complex cognitive task (Baranowski, 1988) and the instruments available may also vary in their cognitive demands (Sallis and Saelens, 2000). Whilst the wording of the IPAQ-LF is reasonably unambiguous, it is possible that the respondents have misread the question or failed to fully understand what they were being asked, leading to errors in the answer provided. Also, responses may have been influenced by social-desirability bias leading to the over reporting of weekly PA (Warnecke *et al.,* 1997). The validity of PA levels may have been improved if the IPAQ-LF had been administered by a member of the research team either as a face-to-face (Van Dyck *et al.,* 2015) or telephone (Hallal *et al.,* 2010) interview, but this was beyond the scope of the present study.

When direct and indirect effects were summed, the causal model identified a diagnosis of PCOS as having the largest total effect (standardised $\beta = .407$, 95% CI: .255 to .571) on participant's HRQoL. When only direct effects were considered, the effect of PCOS upon HRQoL was comparable to the effect of BMI. A diagnosis of PCOS had a greater direct effect on participant self-esteem than did BMI, but both were statistically significant. In-turn, self-esteem had the largest direct effect upon HRQoL. In fact, self-esteem emerged as a key mediator between the exogenous variables and HRQoL (Total effect: standardised $\beta = .324$, 95% CI: .196 to .446). It is therefore likely that women with PCOS have a two-fold effect upon their self-esteem; that is, managing a chronic disease and its associated symptoms promoting lower self-esteem (Reitzes and Mutran, 2006), but also increased BMI (a frequent component of PCOS) contributing to poor self-esteem (Tiggemann, 2005). Another key consideration is the bi-directional effects of self-esteem, where chronic disease reduces an individual's self-esteem, but the reverse is also true. Previous studies have identified that low self-esteem is associated with physical dysregulation (in the context of stress) and physical health complications (Liu *et al.,* 2014; Cott *et al.,* 1999).

In this study, women with PCOS had statistically lower self-esteem than women without. These findings agree with a large community-based cohort study that also reported lower self-esteem in

women with PCOS compared to healthy controls (Tay *et al.*, 2019). This cohort study also found, that as BMI increased, self-esteem was further reduced and levels of psychological distress (as measured by the Kessler Psychological Distress Scale) were increased. Within the general population, it has been previously reported that individuals with a higher body weight/BMI (Tiggemann, 2005) or self-perception of being overweight (Kim and Kim, 2001) are more prone to impaired self-esteem. Furthermore, this effect is greater in females than males (Becerra *et al.*, 2015). In the current study, when the whole cohort was included in analyses, this notion is supported, with the noted statistically significant negative correlation (τ = -.211; P <.001) indicating that as BMI increases, self-esteem decreases. However, when the data are split into case (PCOS) or control, the result changes. The relationship between self-esteem and BMI is still evident in women with PCOS (τ = -.250; *P* = .006), but disappears in the control group (*P* = .222). A potential reason for this is that the median BMI in the latter group is indicative that most of these women are of a weight/BMI within the healthy range.

However, there are well-documented limitations with BMI, particularly regarding characterizing body fat distribution (Nuttall, 2015) which is an important variable in the assessment of metabolic health and mortality risk. As a measure of central adiposity, participants in the current study were asked to provide their waist circumference measurement, which again had statistical and clinical differences between groups. Whilst waist circumference was not included in the path analysis due a number of missing values in both groups, there was a high degree of correlation between the two variables (τ = .652). Self-esteem was negatively correlated with waist circumference in whole group analyses (τ = -.326) and in PCOS (τ = -.394; *P* <.001), whist this effect again disappeared when only the healthy women were included. Therefore, it is difficult to conclude from the present study whether it is BMI or central adiposity alone that are influencing self-esteem, or whether it is the sequelae associated with PCOS. A previous study (Bazarganipour *et al.*, 2013) investigated self-esteem in 300 women with PCOS, showing that there was no relationship with BMI, although poorer self-esteem was reported for women who were hirsute and/or infertile. This tends to suggest that it is the other manifestations of PCOS driving the impairment of self-esteem in these women. However,

delineating these effects was beyond the scope of this study and so to add clarity to this point, future research should utilise age- and BMI-matched control groups to further investigate the role of BMI upon self-esteem and other outcomes in women with PCOS.

Similarly, future work may also wish to manipulate the causal flow of the model. Previously, Joseph *et al.* (2014) utilised path analysis in a healthy student population to assess the role of PA on HRQoL. This analysis incorporated PA as the exogenous variable, and found that individuals with higher total PA had increased self-esteem ($\beta = .10$; $P < .001$), which in turn promoted improved QoL ($\beta = 0.30$; $P < .001$). Regardless of the theoretical model's structure, it appears that self-esteem has an important role in the perception of HRQoL. Therefore, interventions designed to improve HRQoL in women with PCOS (and potentially in other populations) should also focus at improving self-esteem and may wish to incorporate components utilising PA interventions (Ekeland, Heian and Hagen, 2005).

4.6. Limitations

Despite the present study achieving the recruitment target, there are some potential limitations, mostly associated with the manner in which study participants were recruited. Firstly, it could be argued that the case and control groups were recruited from different populations; or more precisely, that a proportion of the control group was recruited through advertisements within the local university. Communications were sent via internal university systems calling out for volunteers with or without PCOS. The control group has a higher number of participants who were students, and those who have a doctorate level qualification at the time of survey. Whilst not explicitly stated, this data tends to suggest that they may have come from the university population. The degree to which this may have influenced the results or introduced bias into the study is difficult to assess. When the recruitment methods are measured against the National Institute of Health (NIH) Quality Assessment tool (NIH, 2014), then it could be argued that participants were drawn from the same population. The eligibility criteria stipulated that women had to be UK-based, either with or without PCOS and recruitment for both case and control was conducted at the same time. Indeed, this could be indicative that women were drawn from the same population. However, if a significant proportion of one group

were drawn from a Birmingham-based university, the outcomes of interest may have been influenced by environmental factors associated with university attendance or education status.

It was also stipulated that participants had to have PCOS or be healthy; that is free from any other condition that may affect their ability to perform PA. Whilst the objective behind this decision was an attempt to isolate the impact of a PCOS diagnosis, rather than any other comorbidity, this decision meant that 40 study respondents were deemed as ineligible from participating, with the majority of these being women with PCOS. It is widely stated that women with PCOS are more susceptible to a range of physical (*e.g.,* CVD, diabetes and insulin resistance, metabolic syndrome, endometrial cancer) and psychological (*e.g.,* anxiety and depression) conditions (Gilbert *et al.,* 2018). By excluding these women, some key aspects of living with PCOS, and important data about the prevalence of comorbidities may have been overlooked.

The fact that study data was also entirely self-reported is a methodological concern. Given the time constraints and the pilot nature of this study, this approach was necessary, but it may have led to inaccuracies in the captured data. The issues around the self-report of physical activity data are discussed earlier in the chapter, with the main concern being the over reporting of PA and under reporting of sedentary behaviours. In fact, a large systematic review has reported that there was a 44% mean difference (favouring over-reporting) when self-report was compared to objective measures of PA, whilst the mean percentage difference increased to 138% when only females were analysed, (Prince *et al.,* 2008). Similarly, another large meta-analysis reports that sedentary behaviours are under-reported by ~105 minutes per day when compared to objective measures. Single-item questionnaires (such as the IPAQ-LF used in this study) further increased the level of disagreement with sedentary time being under-reported by ~160 minutes per day (Prince *et al.,* 2020). The implication for the current study is likely that participants have either consciously or unconsciously over-reported PA or under-reported sedentary time, which may be a contributing factor to these outcomes not contributing to variance in HRQoL.

There may also be inconsistencies with the reporting of a PCOS diagnosis in this study, since formal diagnoses from a medical professional were not obtained for the participants. Whilst the majority of

participants undoubtedly will have received a formal diagnosis, it is not certain that all would have been through this process. The majority of women with PCOS indicated the phenotype with which they had been diagnosed, but ~17% were unable to provide this information. This may be due to a poor diagnosis experience or lack of information provision about the condition (Sills *et al.*, 2001; Gibson-Helm *et al.*, 2016). However, it is also possible that participants have self-diagnosed, without medical intervention, based upon their own experiences and research.

Participants were also asked to provide their own anthropometric measurements. Previous research has shown that people tend to over-report their height and under-report their weight (Gorber *et al.* 2006). In fact, one study in individuals with overweight or obesity has shown under reporting of weight by up to 6 kg (Nawaz *et al.*, 2001). In the current study, participants were asked to also self-report their waist circumference measurement, but ~15% of respondents did not provide that information. Waist circumference is an important (and practical) measure of central obesity that is often self-reported in population-level studies of obesity. However, it has previously been reported that self-reported waist circumference measurements are under-reported by participants, and that the degree of inaccuracy tends to be greater in women and in those with a higher BMI (Bigaard *et al.*, 2005), much like the women with PCOS in this study. The survey used in the current study provided some basic instructions about how to measure waist circumference, but in retrospect these instructions may have been insufficient, meaning that the size of error may have been greater, and uncertainty about how to measure waist circumference may also have contributed to the high rate of the relevant missing data.

4.7. Conclusions

The current study has highlighted less favourable values for women with PCOS, across most included outcomes, suggesting that women with PCOS have worse physical and mental health than their age-matched counterparts without PCOS. In fact, women with PCOS in this study had higher body weight (median difference: 28.3 kg) and BMI (median difference: 9.63 kg/m^2). Whilst there was no direct measure of body fat, the magnitude of difference reported for waist circumference (+27.60 cm) in

women with PCOS demonstrates that central adiposity is also statistically and clinically worse, increasing their risk of type 2 diabetes, cardiovascular disease and all-cause mortality. In addition, women with PCOS had poorer self-esteem and HRQoL, with the size of these effects suggesting that these between group differences may be clinically important and should be a target for improvement in future treatment strategies. There were no between group differences observed for weekly PA levels or sitting time, and nor were there any differences in perceived self-efficacy for exercise. Despite this, observations revealed that women with PCOS perceived fewer benefits and greater barriers to PA participation. What is pertinent is that despite no observed differences in the current study, and the fact that PA and sedentary time are integral variables in the energy balance equation, there are still large differences in weight between cohorts.

Another key study finding was that a PCOS diagnosis has a more severe impact upon perceived mental health than it does upon physical health. Whilst both HRQoL domains were impaired in women with PCOS, the mental health domain was more affected. Contrary to our initial hypothesis, reduced HRQoL was evident irrespective of PA levels and sitting time. There was no evidence in the current study to suggest that those women who performed more PA and spent less time sitting had improved HRQoL, and this was true in both PCOS and control groups. Whilst there are the aforementioned limitations with self-report data, particularly self-reporting of PA and sedentary time, previous studies indicate that there may be some precedent for further research on the role of PA (and sitting time) on HRQoL in women with PCOS. Future studies may wish to utilise measurement methods with greater validity; objective measurement of these metrics alongside HRQoL outcomes should be a research priority, with the inclusion of biochemical analyses where possible.

Whereas PA had no effect upon HRQoL in either study group, self-esteem emerged as an important outcome in individuals' perceived HRQoL. For the whole cohort, there were highly significant positive correlations between self-esteem and the mental, physical and total domain scores of the SF-12. However, when data were separated, the magnitude of this effect increased in women with PCOS; any relationship between self-esteem and the physical domain ceased in the control group. The importance of self-esteem was further augmented during the path analysis, where self-esteem

demonstrated the single largest direct effect upon HRQoL, more so than a diagnosis of PCOS or BMI. Self-esteem is also a key mediator between PCOS and BMI, with both outcomes having a direct effect upon self-esteem, which in turn influences HRQoL. What is less clear is how self-esteem can be improved in these women. BMI may be influenced by lifestyle and pharmaceutical interventions, which may in turn improve self-esteem, but PCOS cannot be prevented and there is no cure. Perhaps increasing PA can improve self-esteem (Spence, McGannon and Poon, 2005), but there is little evidence to support this notion in the current study, since the correlation coefficients between PA and self-esteem were close to zero regardless of the group being analysed. However, the role of PA in the promotion of self-esteem also warrants further investigation. The MET-mins/wk reported in the current study far exceed population norms causing a null relationship. Thus, it should be reiterated that objective measures of PA in women with PCOS should be a research priority. The importance of self-esteem in the promotion of HRQoL is paramount and the impaired self-esteem of women with PCOS is alarming. Investigation of this phenomena should be a priority for medical professionals and researchers alike.

Despite the lack of evidence supporting the notion here, there is a consensus that participation in exercise and PA is beneficial for the participant's mental and physical health. Prior evidence (although not particularly strong) suggests that exercise interventions have the potential to improve a range of outcomes in PCOS (Kite *et al.*, 2019). Therefore, it is important that women with PCOS are achieving at least the current PA recommendations. In this context, the current study revealed that women with PCOS perceive more barriers and fewer benefits to PA participation than the women in the control group. Despite this, the top cited barriers and benefits between groups are similar with an emphasis upon the physical components as opposed to mental; it is largely the degree to which these are perceived that differs. What is also apparent in the present data is that it is those women who perceive fewer barriers and greater benefits that have greater self-efficacy for exercise, and, in-turn, higher levels of PA. Correlation analyses revealed that this effect is greater in women with PCOS than in the control. In fact, the control group had no statistical relationship between self-efficacy for exercise and PA performed. Despite this, the path analysis model revealed exercise

benefits and barriers as the only variable that had any influence upon PA, and this was achieved through self-efficacy as a mediator. The significance of this is that education may have an important role in improving both individual perceptions of PA and confidence to perform exercise. Furthermore, support should be offered to facilitate barrier removal and to design programmes of exercise that are suitable for the needs and abilities of these women. However, the methods in which this could be delivered are uncertain and certainly warrant further investigation.

In summary, the present study highlights that poorer HRQoL is reported by women with PCOS compared to those without PCOS, and highlights self-esteem as a key factor in the promotion of health in this female population. Whilst previous studies suggest that PA has a key role at improving health in a range of populations, the link here is not apparent. This demonstrates a need for future studies, preferably using objective measures to better understand the impact of PA on the health of these women.

Chapter 5: Results
Study 3.
Clustering of cardiometabolic risk factors and their association with physical activity and sedentary time in women with and without polycystic ovary syndrome

5.1. Background/rationale

Polycystic ovary syndrome (PCOS) is a common endocrine disorder, affecting up to 21% of reproductive-aged women (Lizneva *et al.* 2016). PCOS is characterised by hyperandrogenism and ovulatory disruption, and typically manifests in a range of undesirable symptoms such as acne, hirsutism and infertility (Barber *et al.,* 2006). Women with PCOS are also at an increased risk of cardiometabolic disturbances, such as obesity (particularly central obesity with increased waist circumference), and elevated circulating cholesterol levels (Lim *et al.,* 2012; Stepto *et al.,* 2013; Cussons, Stuckey and Watts, 2007). Typically, management of PCOS focuses on lifestyle changes (Legro *et al.,* 2013), incorporating increased physical activity (PA), aiming to alleviate symptoms, and lower the associated risk of type 2 diabetes and cardiovascular disease (CVD).

Although increasing PA is a well-established method for improving CVD risk factors (Ekelund *et al.,* 2015; Samitz *et al.,* 2011), the effect of reducing sedentary behaviours has become a more recent focus. The proportion of sedentary time an individual engages in has been strongly linked to cardiometabolic risk, independent of PA (Tremblay *et al.,* 2010). If a diagnosis of PCOS increases cardiometabolic risk, then those who are not physically active and also spend more time sitting are further increasing the risk of disease.

As such, better understanding of the characteristics of women with PCOS (and their counterparts without PCOS), including risk factors that are modifiable (*e.g.*, PA levels), may lead to more effective and targeted interventions for these risk factors, improving health-related outcomes in these women. Furthermore, this may allow earlier identification of individuals at risk of metabolic syndrome as a co-morbidity, and thus enable intervention that may reduce risk, decrease morbidity, improve patient quality of life, and lower subsequent health care cost.

Consistent with the results of previous studies (Barnard *et al.*, 2007; Ching, Burke and Stuckey, 2007; Coffey, Bano and Mason, 2006; Hahn *et al.*, 2005; Jones *et al.*, 2010; Li *et al.*, 2011), the previous results chapter identified that women with PCOS have poorer health related quality of life (HRQoL) than women without PCOS. It was also revealed that women with PCOS have lower self-esteem than healthy controls and this was a key mediator in the promotion of HRQoL (Chapter 4). There is

contrasting evidence about the relationship between PCOS and self-esteem, but one theory relates to the prevalence of overweightness and obesity in PCOS, and how this influences self-esteem. Indeed, the previous chapter identifies that women with PCOS were statistically and clinically more overweight than their healthy counterparts, placing them at an increased risk of cardiovascular disease.

Despite the aforementioned indicators of poorer health in women with PCOS, there appears to be no difference in the volume of PA performed (Douglas *et al.*, 2006; Wright *et al.*, 2004; Alvarez-Blasco *et al.*, 2011; Cutler, Pride and Cheung, 2018) when compared to women without PCOS. Similarly, there are no observed differences for time engaged in sedentary behaviour when compared to a control group (Ahmadi *et al.*, 2013; Wright *et al.*, 2004; Alvarez-Blasco *et al.*, 2011; Lin *et al.*, 2019). Although contrasting the findings of a previous population-based study (Legro *et al.*,2013), this may be indicative that these outcomes have no role in the maintenance and promotion of health in women with PCOS.

Previous studies have investigated the effectiveness of exercise interventions for improvement in a range of physiological and psychological outcomes in these women. When the results of these studies are meta-analysed (as in Chapter 3), some beneficial effects of exercise are observed but confidence in these findings are uncertain. Effect estimates generally have wide 95% confidence intervals (CIs) that often include negligible change and furthermore, included studies often have poorly reported methodologies and results. One such area that is poorly reported is the degree of compliance to prescribed exercise programmes; indeed, of the studies included in Chapter 3, 61% (n = 11) did not report any information on adherence. Without appropriate reporting, it is difficult to evaluate intervention integrity (Higgins and Deeks, 2008); for instance, negligible effects may be observed due to non-participation, and conversely, if participants receive a higher dose than that prescribed, results may be inflated. Whilst objective measures of PA may be preferential, self-report methods remain the most practical means for population-based measurement (Prince *et al.*, 2019). One such example of this is the UK Biobank project; 502,617 UK-based adults were recruited into the project

between 2006-2010, with the vast majority completing self-report measures of PA and sedentary behaviour (Guo *et al.*, 2020).

The beneficial health-effects of PA are widely reported in previous literature; whilst these effects may be ambiguous in women with PCOS, it is fair to say that information about the health-effect of excessive sitting/sedentary behaviours as part of daily living in these women is even less clear. Analysis of data from the UK Biobank will provide self-report activity data for nearly all participants and where available, objective data for a subset too.

5.2. Objectives and Hypotheses

The objective of the present study is to address some of the research gaps identified in Chapter 4. The UK Biobank provides objective data on a range of physical/physiological outcomes (*e.g.*, weight, glucose, testosterone, *etc*.) that are often associated with a greater severity of symptoms in PCOS. These objective measures allow a direct comparison between women with PCOS and their age-, and age + BMI-matched counterparts, providing an indication of the degree to which a PCOS diagnosis increases cardiometabolic risk. Therefore, cluster analysis was used to identify whether cardiometabolic risk factors cluster together in women with PCOS, and if they do, assess whether the clustering pattern is different compared to those without PCOS.

Furthermore, measures of PA and sedentary behaviour also permit an evaluation of their importance as independent risk factors in the development of cardiometabolic conditions. Using objective and self-report measures, the impact of low PA levels and high sedentary time can be measured and the existence of differences between women with PCOS, and those without, assessed. In addition, the characteristics of participants included in the study can be classified according to their PA and sedentary behaviour profiles. Those classifications are: high PA + low sedentary behaviour (low cardiometabolic risk, reference group); high PA + high sedentary behaviour, and low PA + low sedentary behaviour (intermediate risk); or low PA + high sedentary behaviour (high cardiometabolic

risk), and the associations between groups and prevalence of PCOS, CVD, type 2 diabetes, and metabolic syndrome.

The exploratory hypotheses for the current study are therefore:

1. Women with PCOS have less favourable physical and psychological health than their counterparts without PCOS. Outcomes will differ between women with PCOS and the age + BMI matched group but the greatest differences will occur between women with PCOS and the age-matched only cohort.
2. Women with PCOS will generally have poorer metabolic health and will demonstrate a greater clustering of risk factors. The age + BMI-matched cohort will share many similar traits, but this will differ considerably from the age-matched group.
3. There will be no between group statistical differences for PA and sedentary behaviour, but, those who are more physically active with the lowest time spent in sedentary behaviour will have the lowest cardiometabolic risk regardless of study cohort.

5.3. Methods

For complete details of methodology, please refer to Chapter 2, Sections 2.3.1 to 2.3.11.

5.4. Results

5.4.1. Descriptive Characteristics

The descriptive characteristics of the three groups are presented in Table 5.1. Assumption checks (distribution, skewness and kurtosis) were completed for each outcome and are reported in Appendix 7.17. No single outcome met the assumptions for all criteria across each group. In the PCOS group, SBP, DBP, total cholesterol, LDL-C and IGF-1 did meet the assumptions of normality. However, none of these outcomes met the same requirements in the other groups; indeed, stature (age-matched and BMI + age-matched) and body fat % (age-matched) were the only other outcomes to meet these assumptions. Therefore, non-parametric statistics were used for further analyses. Kruskal-Wallis tests were therefore completed and between group differences were observed for all outcomes apart from stature, total cholesterol, LDL-C, glucose and oestradiol (Table 5.1).

Table 5.1. Descriptive characteristics (median and IQR) of study participants and between group comparisons (Kruskal-Wallis) including outcomes relating to sedentary behaviour and physical activity.

Outcome	PCOS	Age-matched	Age + BMI-matched	Between-group comparison (χ^2)	*p*-value	ε^2
Stature (cm)	164.0 (159-167)	164 (160-168)	164 (160-168)	3.603	.165	0.00228
Body weight (kg)	80.0 (66.8-97.8)	68.5 (61.1-80.6)	81.5 (68.5-97.3)	139.010	**< .001**	0.08793
BMI (kg/m^2)	30.1 (25.5-36.8)	25.4 (22.6-29.7)	30.1 (25.5-36.8)	152.935	**< .001**	0.09673
Waist circumference (cm)	92.0 (80.8-106)	81.0 (73.0-91.0)	92.0 (79.0-105.0)	136.841	**< .001**	0.08650
Hip circumference (cm)	108 (100-120)	101 (95.0-109)	109 (101-121)	138.988	**< .001**	0.08786
WHR	0.835 (0.780-0.890)	0.80 (0.76-0.85)	0.82 (0.77-0.88)	48.442	**< .001**	0.03062
WHtR	0.570 (0.490-0.650)	0.49 (0.45-0.56)	0.56 (0.48-0.64)	138.832	**< .001**	0.08776
Body fat (%)	40.8 (34.3-46.0)	35.4 (29.5-40.6)	41.0 (34.2-46.0)	120.922	**< .001**	0.07817
Systolic BP (mmHg)	129 (119-140)	126 (116-137)	128 (118-140)	7.864	**.020**	0.00527
Diastolic BP (mmHg)	83.0 (74.0-89.0)	79.0 (72-86)	82.0 (73-88)	19.159	**< .001**	0.01283
Total cholesterol (mmol/L)	5.35 (4.70-6.05)	5.41 (4.76-6.05)	5.42 (4.75-6.17)	0.502	.778	3.39e-4
HDL-C (mmol/L)	1.30 (1.11-1.56)	1.49 (1.28-1.80)	1.38 (1.17-1.66)	59.311	**< .001**	0.04433
LDL-C (mmol/L)	3.37 (2.87-3.91)	3.28 (2.78-3.81)	3.30 (2.88-3.96)	2.980	.225	0.00202
Triglycerides (mmol/L)	1.48 (0.96-2.14)	1.09 (0.81-1.64)	1.29 (0.93-1.81)	37.323	**< .001**	0.02523
HbA1c (mmol/mol)	35.2 (32.0-38.2)	33.5 (31.1-35.9)	33.9 (31.4-36.6)	27.836	**< .001**	0.01897
Glucose (mmol/L)	4.81 (4.49-5.24)	4.80 (4.50-5.13)	4.84 (4.50-5.21)	0.341	.843	2.55e-4
IGF-1 (nmol/L)	22.2 (16.9-26.3)	23.1 (19.7-27.1)	21.8 (17.9-26.4)	16.338	**< .001**	0.01108
Testosterone (nmol/L)	1.19 (0.91-1.62)	1.14 (0.82-1.49)	1.18 (0.84-1.56)	6.532	.038	0.00491
SHBG (nmol/L)	43.0 (28.9-68.2)	57.7 (39.8-80.3)	49.6 (33.8-72.3)	32.383	**< .001**	0.02463
Oestradiol (pmol/L)	357 (255-619)	443 (304-637)	421 (300-660)	5.248	.072	0.00621
CRP (mg/L)	2.00 (0.74-4.81)	1.18 (0.50-2.45)	1.93 (0.79-5.24)	53.979	**< .001**	0.03650
MET-mins/wk	1386 (516-2880)	1646 (733-3230)	1493 (688-3282)	4.32	.115	0.00311
Summed PA (mins/wk)	80 (40-150)	90 (50-155)	90 (45-160)	3.08	.214	0.00222
MET-mins/wk VPA	0 (0-480)	240 (0-960)	240 (0-800)	16.18	**< .001**	0.01165
MET-mins/wk MPA	360 (40-900)	360 (120-960)	360 (75-840)	2.18	.336	0.00157
Screen time (hrs/day)	3.9 (2.9-5.0)	3.0 (2.0-4.0)	3.0 (2.0-5.0)	11.63	**.003**	0.00729
Sedentary time (hrs/day)	4.9 (3.0-6.0)	4.0 (3.0-5.8)	4.0 (3.0-5.9)	7.64	**.022**	0.00479

Key: BMI: body mass index; WHR: waist-to-hip-ratio; WHtR: waist-height-ratio; BP: blood pressure; HDL-C: high-density lipoprotein cholesterol; LDL-C: low-density lipoprotein cholesterol; HbA1c: glycated haemoglobin; IGF-1: Insulin-like growth factor-1; SHBG: sex hormone-binding globulin; CRP: C-reactive protein; SD: standard

deviation; IQR: interquartile range; %: percentage of data missing form sample; *P*: significance level from Kruskal-Wallis test of difference; χ^2: chi-square statistic; df: degrees of freedom (equals 2 for all outcomes); *P*: significance level; ε^2: epsilon square effect size; reported outcomes are based upon fasted data; MET: metabolic equivalent of task; mins/wk: minutes per week; hrs/day: hours per day; PA: physical activity; Screen time: summed values of computer usage and TV viewing; Sedentary time: screen time and driving time summed.

Table 5.2. Dwass-Steel-Critchlow-Fligner pairwise comparisons of between group characteristics.

Outcome	Group One	Group Two	Median diff (95% CIs)	W	*P*
Weight (kg)	PCOS	Age	11.50 (8.40, 16.20)	-11.601	< .001
	PCOS	BMI + Age	-1.50 (-5.70, 2.90)	0.789	.843
	Age	BMI + Age	-13.00 (-17.00, -10.70)	15.682	< .001
BMI (kg/m^2)	PCOS	Age	4.70 (3.19, 6.13)	-13.0269	< .001
	PCOS	BMI + Age	0.00 (-1.78, 1.35)	0.0212	1.000
	Age	BMI + Age	-4.70 (-6.03, -3.84)	15.9619	< .001
Waist circumference (cm)	PCOS	Age	11.0 (8.0, 12.5)	-12.92	<.001
	PCOS	BMI + Age	0.00 (-4.00, 4.00)	-1.18	.684
	Age	BMI + Age	-11.00 (-14.00, -9.00)	14.69	<.001
Hip circumference (cm)	PCOS	Age	7.00 (5.00, 11.00)	-11.15	< .001
	PCOS	BMI + Age	-1.00 (-5.00, 1.00)	1.47	.553
	Age	BMI + Age	-8.00 (-12.00, -7.00)	15.89	< .001
WHR	PCOS	Age	0.04 (0.02, 0.05)	-8.99	< .001
	PCOS	BMI + Age	0.02 (-0.01, 0.03)	-3.08	.075
	Age	BMI + Age	-0.02 (-0.04, -0.01)	7.30	< .001
WHtR	PCOS	Age	0.08 (0.05, 0.09)	-13.44	< .001
	PCOS	BMI + Age	0.01 (-0.02, 0.03)	-1.62	.488
	Age	BMI + Age	-0.07 (-0.08, -0.05)	14.47	< .001
Body fat (%)	PCOS	Age	5.40 (3.80, 7.40)	-11.467	< .001
	PCOS	BMI + Age	-0.20 (-2.20, 1.30)	0.328	.971
	Age	BMI + Age	-5.60 (-7.60, -4.70)	14.255	< .001
Systolic BP (mmHg)	PCOS	Age	3.00 (1.00, 6.99)	-3.294	.052
	PCOS	BMI + Age	1.00 (-4.00, 3.00)	-0.570	.915
	Age	BMI + Age	-2.00 (-7.00, -1.00)	3.361	.046
Diastolic BP (mmHg)	PCOS	Age	4.00 (1.00, 5.00)	-5.22	< .001
	PCOS	BMI + Age	1.00 (-2.00, 3.00)	-1.09	.720
	Age	BMI + Age	-3.00 (-5.00, -1.00)	5.16	< .001
HDL-C (mmol/L)	PCOS	Age	-0.19 (-0.2, -0.12)	10.31	< .001
	PCOS	BMI + Age	-0.08 (-0.13, -0.03)	4.53	.004
	Age	BMI + Age	0.09 (0.01, 0.22)	-7.31	< .001

Outcome	Group One	Group Two	Median diff (95% CIs)	W	*P*
Triglycerides (mmol/L)	PCOS	Age	0.39 (0.17, 0.47)	-7.95	< .001
	PCOS	BMI + Age	0.19 (0.04, 0.25)	-3.74	.022
	Age	BMI + Age	-0.20 (-0.33, -0.14)	5.97	< .001
HbA1c	PCOS	Age	1.70 (0.50, 2.10)	-7.48	< .001
	PCOS	BMI + Age	1.30 (0.50, 2.10)	-4.93	.001
	Age	BMI + Age	-0.40 (-0.70, 0.70)	3.11	.072
IGF-1 (nmol/L)	PCOS	Age	-0.90 (-3.78, -1.10)	4.286	.007
	PCOS	BMI + Age	0.40 (-2.29, 0.45)	0.351	.967
	Age	BMI + Age	1.3 (0.47, 2.63)	-5.202	< .001
Testosterone (nmol/L)	PCOS	Age	0.05 (-0.05, 0.17)	-3.57	.300
	PCOS	BMI + Age	0.01 (-0.06, 0.18)	-2.15	.281
	Age	BMI + Age	-0.04 (-0.09, 0.09)	1.84	.394
SHBG (nmol/L)	PCOS	Age	-14.70 (-22.74, -10.92)	7.62	< .001
	PCOS	BMI + Age	-6.60 (-12.03, -0.92)	3.80	.020
	Age	BMI + Age	8.10 (5.26, 15.58)	-5.19	< .001
CRP (mg/L)	PCOS	Age	0.82 (0.34, 1.05)	-8.099	< .001
	PCOS	BMI + Age	0.07 (-0.29, 0.38)	-0.235	.985
	Age	Age and BMI	-0.75 (-0.83, -0.29)	9.277	< .001
MET-mins/wk VPA	PCOS	Age	-240 (-336, 36)	5.59	<.001
	PCOS	BMI + Age	-240 (-244, 34.0)	4.47	.005
	Age	BMI + Age	0 (-31.1, 193)	-1.81	.406
Screen time (hrs/day)	PCOS	Age	0.90 (0.32, 1.30)	-4.87	.002
	PCOS	BMI + Age	0.00 (-0.09, 0.61)	-3.05	.078
	Age	BMI + Age	0.00 (-0.60, 0.09)	2.07	.307
Sedentary time (hrs/day)	PCOS	Age	0.90 (0.20, 1.35)	-3.881	.017
	PCOS	BMI + Age	0.00 (-0.07, 0.65)	-3.058	.078
	Age	Age and BMI	0.00 (-0.48, 0.05)	0.885	.806

Key: BMI: body mass index; WHR: waist-hip-ratio; WSR: waist-stature-ratio; BP: blood pressure; HDL-C: high-density lipoprotein cholesterol; LDL-C: low-density lipoprotein cholesterol; HbA1c: glycated haemoglobin; IGF-1: Insulin-like growth factor-1; SHBG: sex hormone-binding globulin; CRP: C-reactive protein; W: test statistic; *P*: significance level from pairwise comparison; PCOS: women with PCOS; Age: age-matched women; BMI + Age: BMI and age-matched women; 95% CIs: 95% confidence intervals (lower, upper)

When post-hoc DSCF pairwise comparisons were made to identify where the differences lie, it was revealed that the majority of differences were found when the PCOS cohort were compared to the age-matched control group (Table 5.2). Of the 15 statistical differences that were identified, seven related to body composition: weight, BMI, waist circumference, hip circumference, waist-to-hip-ratio, waist-stature ratio, and body fat percentage. Given that these groups were not BMI-matched, this is to be expected. Indeed, for these seven outcomes, there were no significant differences between the PCOS and the age + BMI-matched group, and furthermore, these differences were still apparent when the BMI + age-matched was compared to the age-matched only.

Aside from body weight/composition associated outcomes, eight additional outcomes differed between the PCOS and age-matched cohort: DBP, HDL-C, triglycerides, HbA1c, IGF-1, testosterone, SHBG and CRP, with all these exhibiting less favourable values in the PCOS group. The age + BMI-matched group compared to the age-matched group exhibited similar differences; all outcomes were similar with the exception of no differences between HbA1c, and testosterone, whilst there was also a significant difference in SBP, with the age + BMI-matched group having higher values. There were also four outcomes that were significantly different when PCOS was compared to the age + BMI-matched group; HDL-C, triglycerides, HbA1c and SHBG all had less favourable values in the PCOS cohort.

5.4.2. Physical Activity and Sedentary Behaviour

For these analyses, outcomes related to either self-reported PA levels or sedentary behaviours (*i.e.*, screen time, passive transport), with the least missing data were identified. None of these outcomes met the assumptions of a normal distribution (Appendix 7.17), so non-parametric analyses were completed.

There were no statistically significant between group differences for summed MET-mins/week, summed weekly minutes of activity or MET-mins/week completed through moderate intensity PA (Table 5.1). The Kruskal-Wallis test identified a statistical difference (P <.001; $\varepsilon^2 = 0.01165$) in the weekly amount of MET-mins/wk achieved through vigorous intensity PA. DSCF pairwise

comparisons (Table 5.2) identified that these differences lay between the women with PCOS and the age-matched cohort (W = 5.59; *P* <.001), and between the women with PCOS and the BMI + age-matched cohort (W = 4.47; *P* = .005). There were no differences when the two non-PCOS cohorts were compared. Despite statistical significance, it should be noted that, where differences exist, the volume of VPA is relatively small, so too is the effect size, and the 95% CIs include negligible differences.

The screen time outcome, derived from the summation of TV viewing time and computer usage (hrs/day), and the sedentary time outcome, derived from screen time summed with driving time, both reported statistical significance (*P* = .003 and *P* = .022) when Kruskal-Wallis analyses was completed. For both outcomes, post-hoc DSCF tests (Table 5.2) revealed that between groups differences lay between the PCOS group and the age-matched group only (screen time: W = -4.87; *P* = .002, and sedentary behaviours: W = -3.881; *P* = .017). As with VPA, caution should be adopted in translating these findings; when comparing women with PCOS to the age-matched control, the actual difference represents only a small difference in minutes per day. The lower CIs represent < 20 minutes per day for screen time and 12 minutes per day for summed sedentary behaviours.

When activity levels were assessed via IPAQ categorical score, a statistical effect was found (χ^2 = 10.7; *P* = .031) indicating that there is a significant difference between the expected and the observed frequencies in one or more categories (Table 5.3). The frequencies reported suggest that a greater percentage of women with PCOS are in the lowest level activity group when compared to the other groups. Furthermore, a smaller percentage of women with PCOS are categorised as achieving the highest activity levels.

Table 5.3. Frequencies of IPAQ activity categorisation.

	Study Group		
IPAQ Activity Group	**PCOS** *n* = 253	**Age-matched** *n* = 525	**BMI + Age -matched** *n* = 612
Low	77 (30.4)	109 (20.7)	134 (21.9)
Moderate	98 (38.7)	215 (40.9)	247 (40.4)
High	78 (30.8)	201 (38.2)	231 (37.7)

Key: IPAQ: International Physical Activity Questionnaire; PCOS: polycystic ovary syndrome; age-matched: control group age-matched only to case; BMI: body mass index; BMI + age-matched: control group both BMI + age-matched to case; *n*: total participants in analysis; Data are presented as number of participants in each IPAQ activity group (percentage of study group).

Similarly, using the IPAQ results to assess whether participants met national PA recommendations (Appendix 7.18), a chi-squared test revealed a statistical effect (χ^2 = 6.06; *P* = .048). Despite statistical significance, confidence can be reduced in these findings due to significance being so close to the threshold (*P* <.05) for statistical significance (likely false positive) and from looking at the walking data. When walking was not included, fewer women (~9%) with PCOS met PA recommendations than the age-matched group; compared to the age-matched group, BMI + age-matched women also reported 5.5% less that were meeting PA recommendations. When walking was included, there were no statistical differences.

5.4.3. Health ratings and satisfaction scores

All UK Biobank participants were asked to self-rate their perceived health. A chi-squared analysis revealed a statistically significant effect (χ^2 = 80.1; *P* <.001) indicating a difference from the expected frequencies (Table 5.4). Compared to the age-matched group and the BMI + age-matched group, a lower percentage of women with PCOS rated their health as either "Excellent" or "Good". Furthermore, a higher percentage also self-rated their health as "Poor". There are differences between the age- and BMI + age-matched groups, but scores tend to be less favourable in the PCOS cohort. The relatively high proportion of participants who have rated their health as "Excellent" or "Good" seems to be at odds with the high prevalence of comorbidity (as identified by ICD-10) in the study participants.

Table 5.4. Frequencies for Health Rating.

	Study Group		
	PCOS	**Age-matched**	**Age and BMI-matched**
	n = 319	*n* = 637	*n* = 638
Excellent	32 (10.0)	103 (16.2)	95 (14.9)
Good	133 (41.7)	382 (60.0)	327 (51.3)
Fair	95 (29.8)	122 (19.2)	162 (25.4)
Poor	58 (18.2)	29 (4.6)	50 (7.8)
Do not know	1 (0.3)	1 (0.2)	3 (1.3)
Prefer not to answer	0 (0)	0 (0)	1 (0.2)

Key: IPAQ: International Physical Activity Questionnaire; PCOS: polycystic ovary syndrome; age-matched: control group age-matched only to case; BMI: body mass index; BMI + age-matched: control group both BMI + age-matched to case; *n*: total participants in analysis; Data are presented as number of participants in each IPAQ activity group (percentage of study group).

A smaller number of participants (PCOS: *n* = 134; age-matched: *n* = 217; age + BMI-matched: *n* = 223) responded to the questions about their satisfaction with a range of outcomes related to psychosocial components of their life. When their health satisfaction was assessed (Figure 5.1A), chi-squared analysis found a statistical effect ($\chi^2 = 44.6$; $P < .001$). A higher proportion of women with PCOS stated that they were either "extremely unhappy", or "very unhappy" with their health compared to the other cohorts. Conversely, a lower percentage stated that they were "Very Happy" with their health; the age-matched group had the highest proportion in this category (29%).

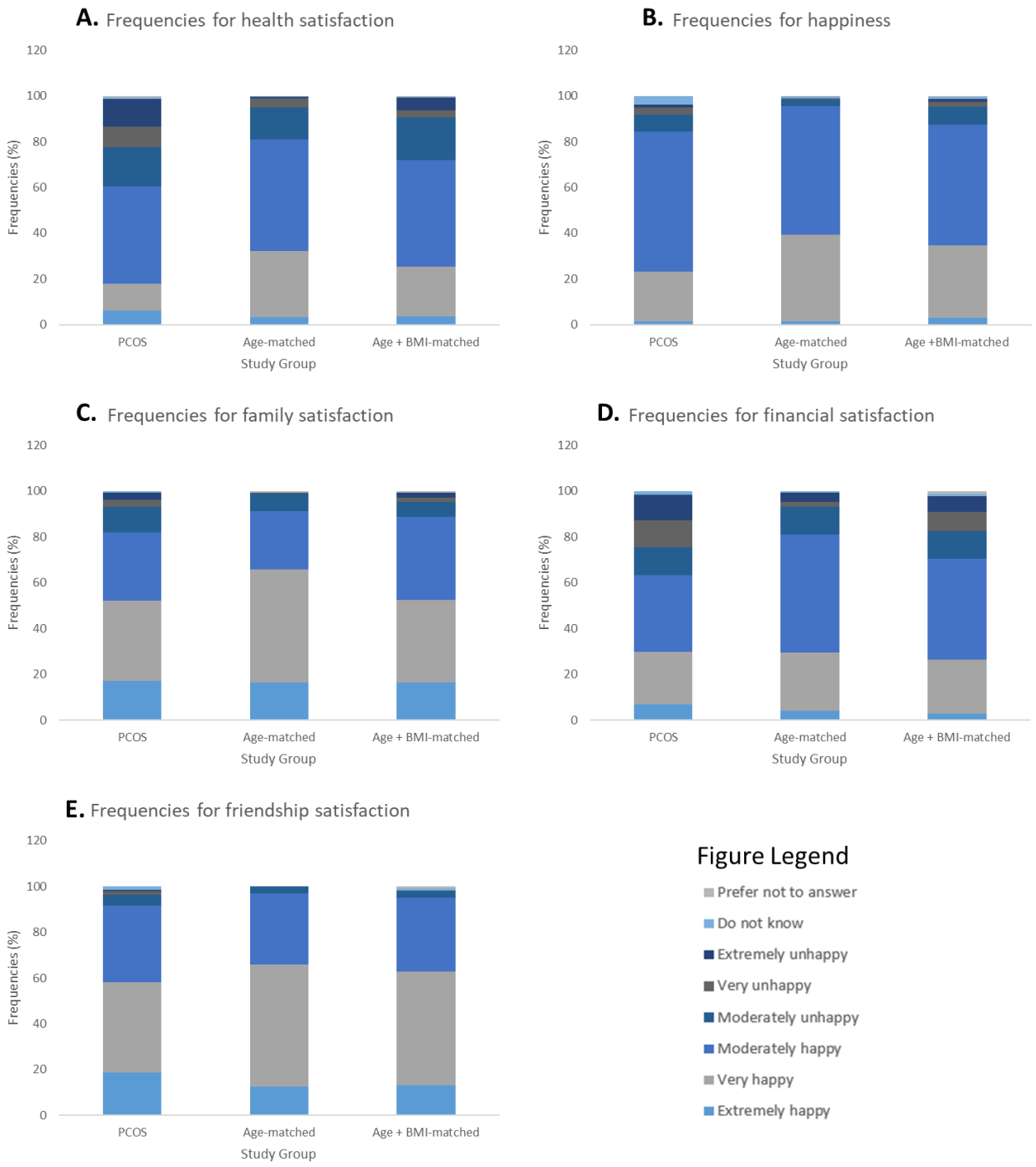


Figure 5.1. A. Frequencies for health satisfaction; **B.** Frequencies for happiness; **C.** Frequencies for family satisfaction; **D.** Frequencies for financial satisfaction; **E.** Frequencies for friendship satisfaction.

Similarly, when participants were asked to rate their happiness (Figure 5.1B), it was the age-matched cohort who were the happiest, with ~38% having stated that they were "Very Happy" compared to

31.4% and 21.6% in the Age + BMI-matched and the PCOS group, respectively. A higher proportion of women with PCOS were “Moderately Happy” compared to their non-PCOS counterparts. Women with PCOS and the Age + BMI group had similar levels of moderate unhappiness, but a greater number with PCOS were either “Very Unhappy”, or “Extremely Unhappy”.

Analysis of frequencies for Family Satisfaction (Figure 5.1C) revealed a statistically significant effect ($\chi^2 = 27.0$; $P = .019$). A greater proportion of women with PCOS were either “Moderately Unhappy”, “Very Unhappy” or “Extremely Unhappy” with their familial situation. In contrast, the largest proportion of those who were “Very Happy” (49.3%) came from the Age-matched group.

When responses on Financial Satisfaction were analysed (Figure 5.1D), a significant effect was observed ($\chi^2 = 33.6$; $P = .002$); a greater number of women with PCOS reported greater dissatisfaction with their financial situation than the number of women in the other groups. In contrast, the large majority of the Age-matched group (81%) responded positively to this question. Analysis of Friendship Satisfaction (Figure 5.1E) revealed no statistical significance between groups ($P = .069$).

Table 5.5. Frequencies and χ^2 tests of frequencies for comorbidities (interpreted from ICD10).

	PCOS (*n* = 318)		**Age** (*n* = 638)		**Age + BMI** (*n* = 638)		χ^2	***P***
	Frequency	**%**	**Frequency**	**%**	**Frequency**	**%**		
Type 2 diabetes	49	15.4	17	2.7	30	4.7	63.90	**<.001**
Type 1 diabetes	7	2.2	4	0.6	6	0.9	5.11	.078
Hypertension	97	30.4	61	9.6	91	14.3	71.60	**<.001**
Hypercholesterolaemia	35	11.0	18	2.8	29	4.5	29.70	**<.001**
Hyperlipidaemia	2	0.6	4	0.6	7	1.1	1.05	.592
Heart Disease	10	3.1	8	1.3	11	1.7	4.27	.118
Obesity	59	18.5	27	4.2	58	9.1	52.7	**<.001**
Liver Disease	2	0.6	1	0.2	0	0	4.51	.105
Kidney Disease	10	3.1	4	0.6	7	1.1	10.70	**.005**
Chronic Lung Disease	57	17.9	58	9.1	69	10.8	16.6	**<.001**
Anxiety/Depression	15	4.7	13	2.0	19	3.0	5.28	.071
Benign Cancer	40	12.5	32	5.0	32	5.0	23.70	**<.001**
Malignant Cancer	20	6.2	37	5.8	28	4.4	1.96	.376

Key: *n*: number from each study group; %: percentage with or without comorbidity from each cohort; Yes: positive diagnosis for comorbidity; No: no diagnosis for comorbidity; χ^2: chi-square statistic; *P*: significance value from χ^2; PCOS: women with PCOS group; Age: age-matched control group; Age + BMI: age and BMI-matched control group.

When tests of frequencies were analysed for comorbidities (Table 5.5), statistically significant chi-squared tests were returned for type 2 diabetes, hypertension, hypercholesterolaemia, obesity, kidney disease, chronic lung disease, and benign cancers. Without exception, prevalence was higher for women with PCOS than the other two cohorts. In fact, prevalence of type 2 diabetes was ~6 times greater in women with PCOS than in the age-matched group, and ~3 times that of the BMI + age-matched group. Similar trends were observed for hypertension (approximately twice more prevalent than the BMI + age-matched group), hypercholesterolemia (approximately 2.5 times as prevalent than the BMI + age-matched group) and kidney disease (approximately 3 times more prevalent than the BMI + age-matched group). The difference between women with PCOS and the age-matched only group was even more pronounced.

When the total number of comorbidities, and only those linked to metabolic health were summed (Appendix 7.19), an ANOVA Kruskal-Wallis test revealed differences for both total comorbidities ($\chi^2 = 109$; $P < .001$) and metabolic morbidity ($\chi^2 = 116$; $P < .001$). DSCF analyses revealed that women with PCOS had significantly more comorbidities (PCOS vs age-matched: W = -14.26; $P < .001$; PCOS vs age + BMI-matched: W = -11.12; $P < .001$) and metabolic complications (PCOS vs age-matched: W = -14.83; $P < .001$; PCOS vs age + BMI-matched: W = -10.45; $P < .001$) than those without. The age- and BMI + age-matched groups were more similar although there all comparisons were statistically significant.

5.4.4. Cluster Analysis

As described in the methods section, outliers 1.5*IQR above the third quartile or below the first quartile were replaced with the 95th or 5th centile, respectively, for each outcome in each group. The vast majority of outliers were transformed from above the third quartile with only seven cases trimmed below the first quartile. A description of the cut-off points and replacement outlier values (*i.e.*, 5th and 95th centile) are presented in Appendix 7.13.

Many outcomes demonstrated highly significant correlations (Table 5.21), but the magnitude of the test was low in many instances (*i.e.*, $r \pm < 0.30$). Assessment of magnitude allows removal of outcomes where they are essentially reporting the same phenomenon. For example, LDL-C and total cholesterol ($r = .946$), SBP and DBP ($r = .742$), BMI with waist circumference ($r = .917$), and BMI with body fat ($r = .888$) are all unsurprisingly highly positively correlated. Because the strength of correlation is so great between these pairs, one from each pair can be removed. Although BMI is a measure of body weight widely used to define obesity status, it was felt that waist circumference and body fat reveal more about the metabolic status of the participant regarding central obesity and adiposity, respectively, and thus BMI was removed from additional analyses.

Table 5.6. Pearson correlation coefficients using complete pairs between any two variables.

	HDL-C	SHBG	IGF-1	Test	CRP	WC	BF%	HbA1c	DBP	TG	SBP	BMI	LDL-C	TC
HDL-C														
SHBG	.400***													
IGF-1	.063	-.101***												
Test	-.100**	-.108***	.036											
CRP	-.247***	-.142***	-.328***	.139***										
WC	-.486***	-.421***	-.300***	.145***	.519***									
BF%	-.413***	-.396***	-.260***	.132***	.461***	.861***								
HbA1c	-.174***	-.256***	-.176***	.071*	.247***	.344***	.257***							
DBP	-.163***	-.247***	-.085**	.100***	.262***	.439***	.465***	.138***						
TG	-.408***	-.263***	-.169***	.016	.198***	.432***	.364***	.255***	.245***					
SBP	-.058*	-.191***	-.110***	.077**	.191***	.320***	.336***	.213***	.742***	.200***				
BMI	-.476***	-.422***	-.303***	.142***	.595***	.917***	.888***	.361***	.467***	.414***	.330***			
LDL-C	-.006	-.151***	-.040	.018	.062*	.181***	.219***	.100***	.197***	.334***	.175***	.169***		
TC	.266***	-.041	-.042	-.017	.009	.074*	.119***	.076**	.160***	.331***	.171***	.055*	.946***	

Key: HDL-C: high-density lipoprotein cholesterol; SHBG: sex hormone binding globulin; IGF-1: insulin-like growth factor-1; Test: testosterone; CRP: C-reactive protein; WC: waist circumference; BF%: body fat percentage; HbA1c: glycated haemoglobin; DBP: diastolic blood pressure; TG: triglycerides; SBP: systolic blood pressure; Group: cohort; LDL-C: low-density lipoprotein cholesterol; TC: total cholesterol. *: $P \leq .05$; **: $P \leq .01$; ***: $P \leq .001$.

Correlations ≥ 0 to .199 | .200 to .399 | .400 to .599 | .600 to .799 | ≥ .800 | < 0 to -0199 | -.200 to -.399 | -.400 to -.599 | -.600 to -.799 | ≤ .800

When missing data was analysed, SHBG (17.5%) had the largest amount of data missing, followed by, Total testosterone (16.6%), and HDL-C (16.1%). All other outcomes had ≤ 8% of data missing (Appendices 7.14 and 7.15). In fact, of these data, no single variable had zero missing data, 36.1% (*n* = 575) cases had some missing data. This equates to ~9% of values missing.

Missing variable analysis revealed 39 unique patterns of missing data (Appendix 7.16). The most common pattern (*n* = 1020 cases) is no missing data for the included variables (Appendices 7.4 and 7.5). The next most common missing patterns are SHBG and HDL-C as a pair (*n* = 112 cases); testosterone only (*n* = 107 cases); cholesterol, CRP, triglycerides, LDL-C, IGF-1, HDL-C, testosterone and SHBG all missing (*n* = 64 cases); HbA1c only (*n* = 64); DBP and SBP are also frequently missing pattern together (*n* = 62 cases).

Although cholesterol and LDL-C had a similar amount of missing data, they exhibit a near perfect correlation (*r* = .946) with each other; it is not necessary to include both outcomes on further analyses. Total cholesterol was removed as it was more positively correlated with HDL-C (*r* = .266).

Similarly, waist circumference and body fat have a strong positive correlation (*r* = .861) and are similar measures. Because waist circumference has fewer missing data and better describes the distribution of adiposity in participants, body fat was removed from subsequent analyses.

Removing these outcomes reduced missing data patterns to 31 and marginally increased the number of complete cases to 1043. Little's MCAR test was re-run on the amended data set (χ^2 = 274.289, df = 227; *P* = .017) indicating data missing not at random. Accordingly, data imputation was carried out on the following outcomes: waist circumference, DBP, SBP, CRP, HbA1c, HDL-C, IGF-1, LDL-C, SHBG, testosterone and triglycerides.

Four participants had no data for any of the included outcomes so it was not possible to impute missing data. These participants were removed from subsequent analysis. The descriptive statistics (mean ± SD) for each outcome before and after multiple imputation are presented in Appendix 7.21. Once missing values had been imputed, data was standardised before K-means cluster analysis was

completed (Appendix 7.22). Change in cluster centres and between cluster ANOVAs are presented in Appendices 7.22 and 7.23, respectively.

Table 5.7. Descriptive statistics of outcomes included in Cluster 1 and Cluster 2.

	Mean ± SD	95% Confidence Interval	Median (IQR)	Skewness	Kurtosis	Shapiro-Wilk *P*
Cluster 1						
Waist circumference (cm)	80.26 ± 10.38	79.61 to 80.9	79 (15)	.624	.251	**<.001**
Diastolic BP (mmHg)	76.85 ± 8.86	76.30 50 77.40	77 (12)	.148	.105	**.047**
Systolic BP (mmHg)	124.03 ± 14.09	123.16 to 124.90	123 (18)	.590	.500	**<.001**
C-reactive protein (mg/L)	1.58 ± 2.03	1.45 to 1.70	0.95 (1.51)	3.433	15.73	**<.001**
HbA1c	32.77 ± 3.34	32.57 to 32.98	32.8 (4.3)	.313	1.844	**<.001**
HDL-C (mmol/L)	1.59 ± 0.32	1.57 to 1.61	1.57 (0.44)	.288	-.377	**<.001**
IGF-1 (nmol/L)	24.13 ± 5.56	23.79 to 24.27	23.87 (7.49)	.253	-.113	**<.001**
LDL-C (mmol/L)	3.21 ± 0.69	3.17 to 3.25	3.16 (0.91)	.316	.285	**<.001**
SHBG (nmol/L)	70.51 ± 29.92	68.66 to 72.36	66.40 (36.75)	.856	.884	**<.001**
Testosterone (nmol/L)	1.18 ± 0.47	1.15 to 1.21	1.13 (0.59)	.718	.552	**<.001**
Triglycerides (mmol/L)	1.11 ± 0.46	1.08 to 1.14	1.01 (0.58)	1.254	2.080	**<.001**
Cluster 2						
Waist circumference (cm)	104.83 ± 13.44	103.73 to 105.92	103 (16)	.473	.050	**<.001**
Diastolic BP (mmHg)	88.07 ± 8.74	87.36 to 88.78	87 (11)	.094	-.264	**.003**
Systolic BP (mmHg)	138.56 ± 14.59	137.38 to 139.75	137 (21)	.287	-.149	**.001**
C-reactive protein (mg/L)	6.25 ± 5.07	5.83 to 6.66	4.75 (6.72)	1.252	1.369	**<.001**
HbA1c	37.28 ± 5.17	36.86 to 37.70	36.4 (5.7)	1.168	2.051	**<.001**
HDL-C (mmol/L)	1.23 ± 0.25	1.21 to 1.25	1.21 (0.31)	.633	.638	**<.001**
IGF-1 (nmol/L)	20.13 ± 5.96	19.65 to 20.62	19.98 (8.43)	.214	-.265	**.019**
LDL-C (mmol/L)	3.65 ± 0.75	3.58 to 3.71	3.60 (1.09)	.018	-.240	.176
SHBG (nmol/L)	39.15 ± 21.53	37.40 to 40.90	34.54 (23.46)	1.749	6.319	**<.001**
Testosterone (nmol/L)	1.33 ± 0.55	1.28 to 1.37	1.27 (0.65)	.767	.575	**<.001**
Triglycerides (mmol/L)	2.07 ± 0.88	2.00 to 2.14	1.90 (1.08)	.998	1.180	**<.001**

Key: BP: blood pressure; HbA1c: glycated haemoglobin; HDL-C: high-density lipoprotein cholesterol; IGF-1: insulin-like growth factor-1; LDL-C: low-density lipoprotein cholesterol; SHBG: sex hormone binding globulin; SD: standard deviation; IQR: inter-quartile range; *P*: significance value.

Table 5.7 shows the descriptive statistics of each outcome separated by cluster number. Without exception, the values of cluster one are more favourable (*i.e.*, associated with healthier status) than those presented in cluster two. In addition, when cluster membership by cohort is analysed, there is a relatively even split for women with PCOS (Cluster one: 48.7% *vs* Cluster two: 51.3%). The difference is slightly increased when cluster membership of the Age + BMI cohort is considered; ~56% have membership in Cluster One. There are a far larger proportion of Age-matched women (~78%) in Cluster One. Looking at Cluster Two membership, the data reveals that women with PCOS have the highest proportion of their cohort as members, this is followed by the Age + BMI-matched (7.4% less than PCOS) and Age-matched only (29.2% less than PCOS). Table 5.8 shows the proportion of study group that make up each cluster. Although, women with PCOS have half as many cases as the other group, they still make up nearly a third of Cluster two.

Table 5.8. Overview of cluster membership based upon study comparator group.

	PCOS	**Age-matched**	**Age + BMI matched**
Cluster One, *n* (%)	155 (15)	496 (49)	357 (35)
Cluster Two, *n* (%)	163 (28)	141 (24)	279 (48)

Key: PCOS: polycystic ovary syndrome cohort; Age-matched: cohort that are age-matched only with the PCOS cohort; Age + BMI-matched: cohort that are matched with the PCOS cohort using both age and body mass index; *n*: number from study group in cluster; %: percentage of cluster from each study group.

Figures 5.2 and 5.3 present the median and interquartile ranges for each outcome separated by cluster number. Within cluster, data are presented by cohort/group membership. Statistical analysis was not presented because clusters have been designed to be as different as possible.

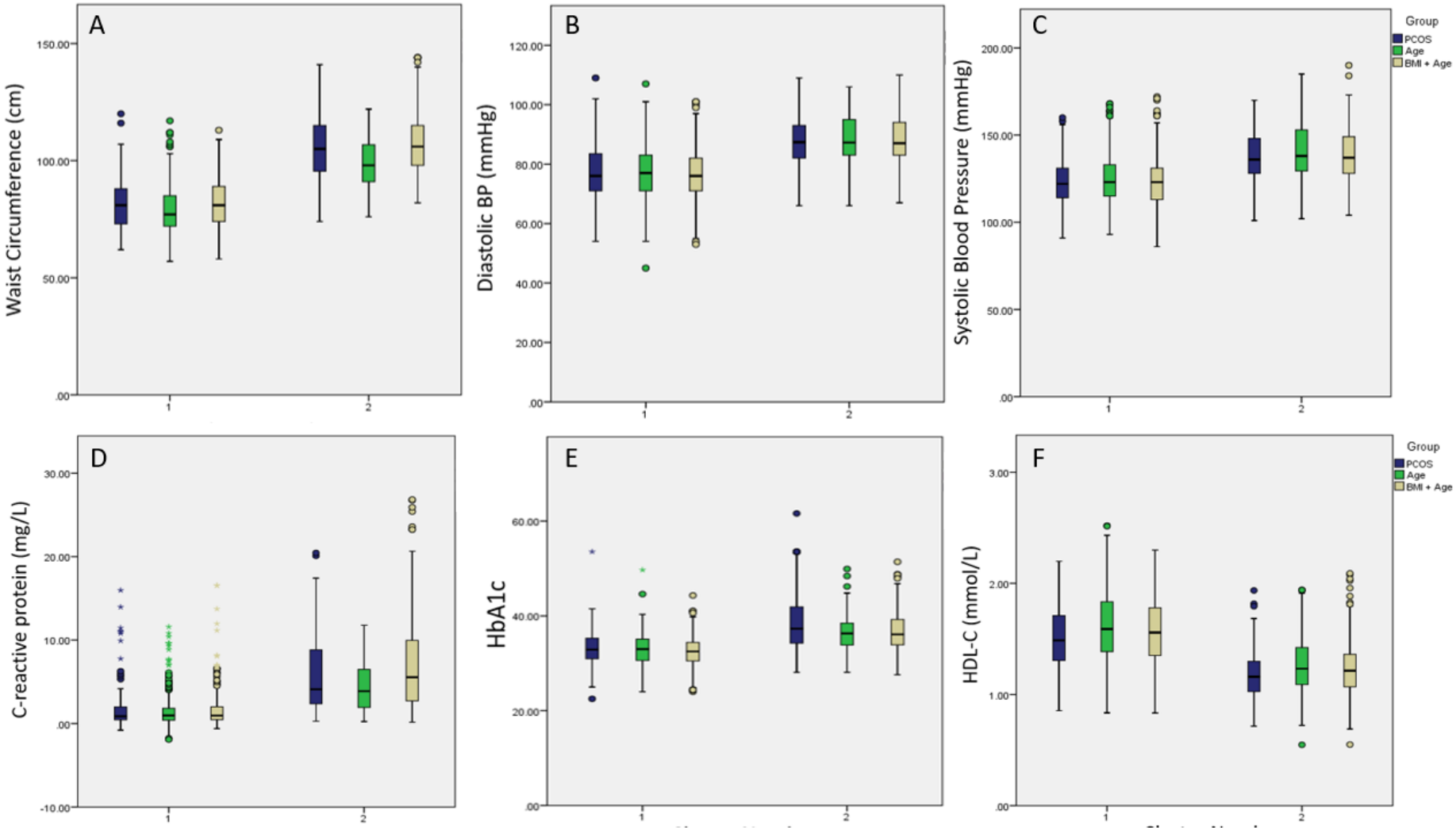


Figure 5.2. Data presented are median values of each study group separated by cluster membership. **A.** waist circumference scores; **B.** diastolic blood pressure (BP) scores; **C.** systolic blood pressure (BP) scores; **D.** C-reactive protein scores; **E.** HbA1c (mmol/mol) scores; **F.** HDL-C scores**;** Coloured box represents interquartile range; the lower whisker is the 1st quartile and upper whisker is the 4th quartile; coloured dots outside of whisker are outliers.

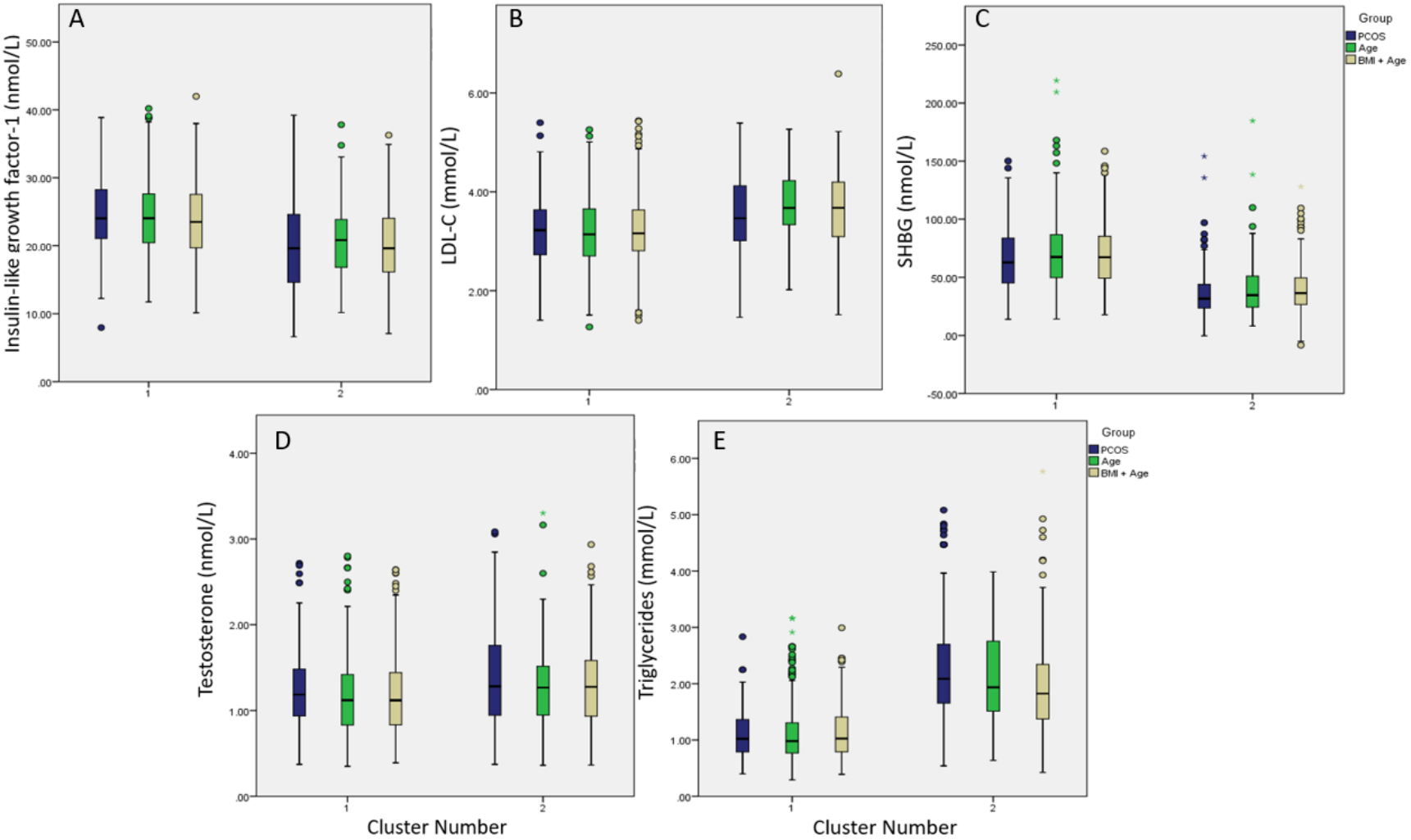


Figure 5.3. Data presented are median values of each study group separated by cluster membership. **A.** IGF-1 scores; **B**. LDL-C scores; **C.** SHBG scores; **D.** testosterone scores; **E.** triglycerides scores; coloured box represents interquartile range; the lower whisker is the 1st quartile and upper whisker is the 4th quartile. Coloured dots outside of whisker are outliers.

Using the clusters as categorical groups, descriptive statistics and tests of difference (Table 5.9) were completed for key outcomes relating to PA behaviours, namely weekly PA (MET-mins/wk) and total sedentary time (hrs/day). These analyses show that cluster one has significantly more MET-mins/wk performed (d = 0.23; P <.001) and less sedentary behaviour per day (d = 0.43; P <.001).

Using independent samples t-tests, there were no statistical differences for the number of comorbidities (mean difference: 0.08; P = .204) or metabolic morbidities (mean difference: 0.03; P = .237) between clusters. However, when data is split by cohort, within each cluster (Figure 5.4), it is clear that in cluster one, women with PCOS have a greater number of comorbidities. These data also do not account for their PCOS diagnosis.

Table 5.9. Descriptive statistics for selected variables for Cluster One and Cluster Two and test of difference using Mann-Whitney *U* independent samples.

	Cluster One	Cluster Two	Mean difference	95% CIs of difference	*U*	Cohen's *d*	*P*
MET-mins/wk	1792 (2644)	1182 (2134)	610	276 to 796	179608	0.226	<.001
Sedentary time (hrs/day)	4.00 (2.00)	5.00 (3.00)	-1.00	-1.29 to -0.80	218352	-0.434	<.001

Key: MET-mins/wk: metabolic equivalent of task minutes per week as measured by the IPAQ; SD: standard deviation; IQR: interquartile range; mean diff: mean difference between cluster; 95% CI of diff: 95% Confidence interval of difference; *U*: test statistic; *P*: significance level

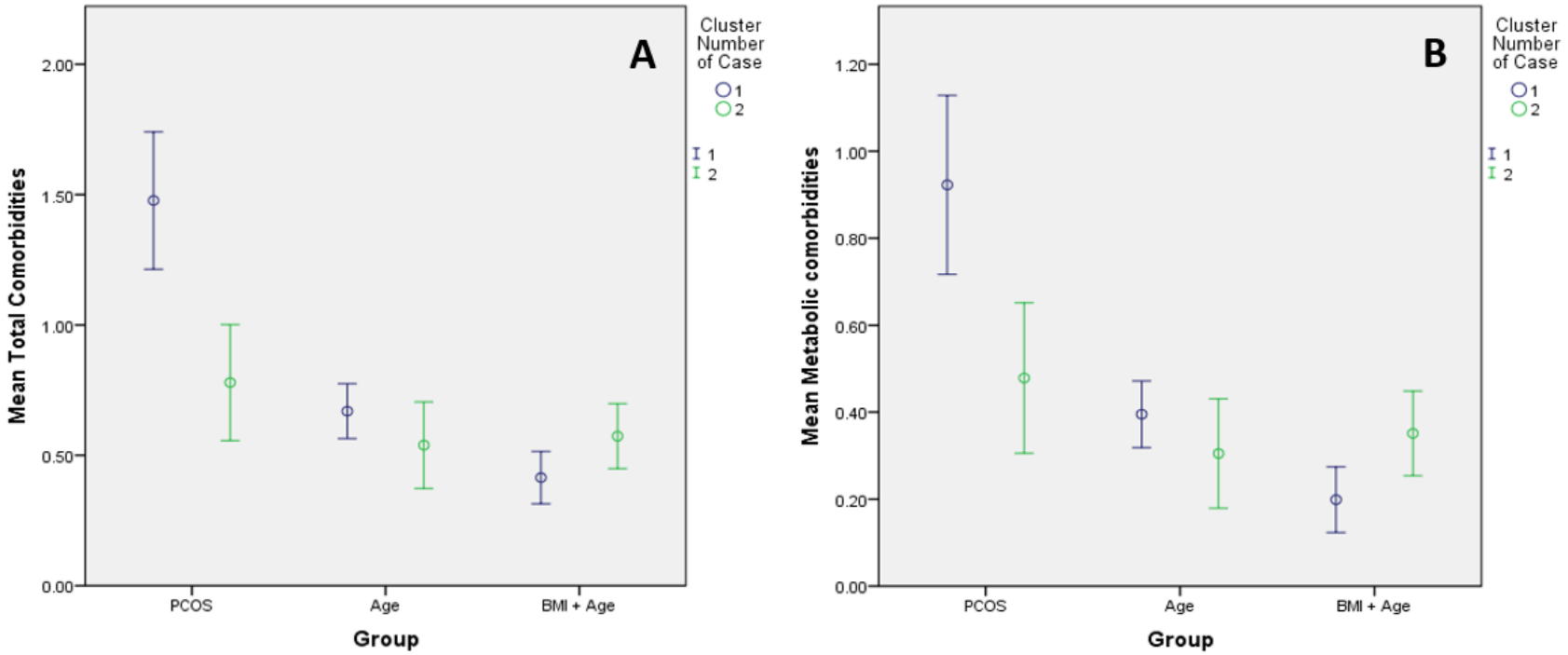


Figure 5.4. Comparison of mean number of total comorbidities (A) and metabolic comorbidities (B) separated by study group for each cluster. Error bars represent 95% confidence intervals.

5.4.5. Risk profiles of Physical activity and sitting

When the PA and sitting time data was split into 50th centiles and used to create a new categorical variable intended to reflect risk of ill health, 207 participants were excluded because data from one of the two variables was missing (*i.e.*, PA or sitting time). The majority of missing cases (n = 199) came from missing data on MET-mins/week performed. Descriptive statistics for missing cases are presented in Appendix 7.24 whilst the descriptive statistics for the remaining participants are presented in Tables 5.10 and 5.11.

Table 5.10. Descriptive statistics for physical activity and sitting outcomes in each risk profile.

Risk Group	***N***	**MET-mins/wk** Mean ± SD	**Sitting time (hrs/day)** Mean ± SD
High PA, Low Sit	394	3901 ± 2418	3.15 ± 0.82
High PA, High Sit	297	4001 ± 2391	6.20 ± 1.77
Low PA, Low Sit	325	735 ± 433	3.20 ± 0.82
Low PA, High Sit	366	632 ± 444	6.99 ± 2.76

Key: PA: physical activity (MET-mins-wk); Sit: sitting time (hrs/day); SD: standard deviation

Figure 5.5 presents the percentage of study group (*i.e.*, women with PCOS, age-matched controls, or age- and BMI-matched controls) membership that makes up each risk group based on PA and sitting time. It should be noted that the age-matched and the age- + BMI-matched study groups have twice as many participants than the women with PCOS. However, it can be seen that women with PCOS have the largest contribution in the Low PA, High Sit group (23%), which is equivalent to ~26% of the total cohort; for comparison, their next highest group membership is 7.3% lower (High PA, Low Sit).

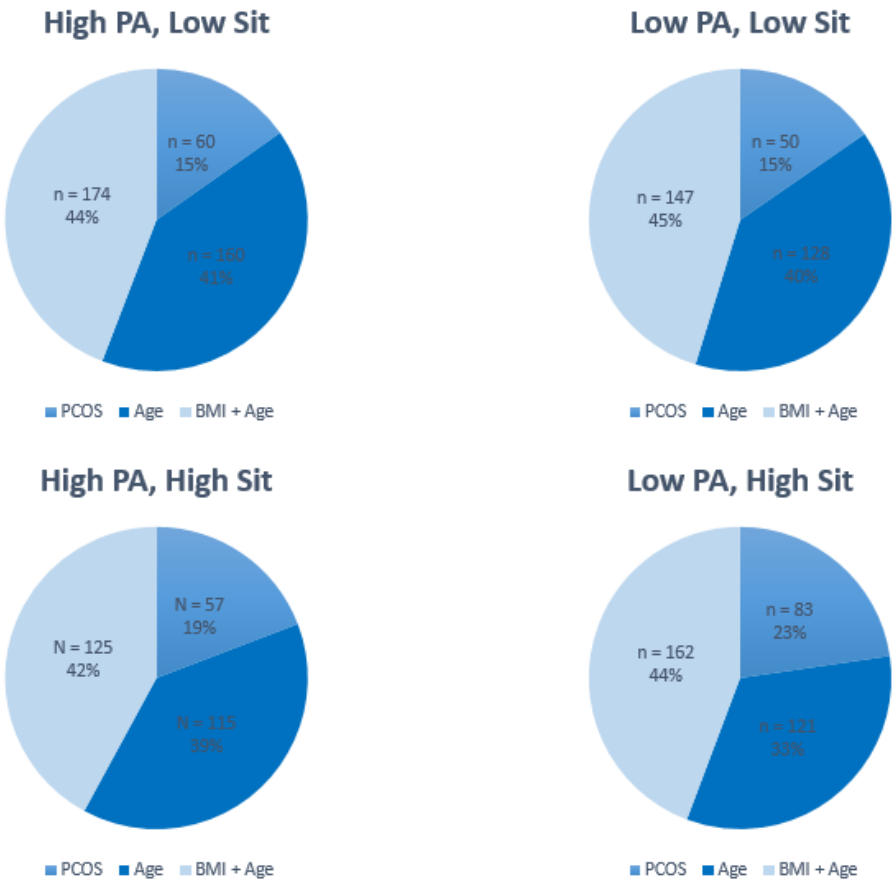


Figure 5.5. Study group membership based on physical activity (PA) and sedentary behaviour (sitting time) risk groups.

Table 5.11. Descriptive statistics for each physical activity (PA) and sedentary behaviour (sitting time; Sit) risk group and Kruskal-Wallis test results.

	High PA, Low Sit	**High PA, High Sit**	**Low PA, Low Sit**	**Low PA, High Sit**	***W***	***P***
Age (yrs)	44.00 (6.00)	45.00 (7.00)	45.17 (6.00)	45.00 (7.00)	3.188	.364
Waist circumference (cm)	79.00 (19.00)	87.00 (21.00)	85.00 (22.50)	94.50 (26.00)	99.627	**<.001**
BMI (kg/m^2)	26.40 (8.85)	25.25 (7.21)	29.65 (12.00)	29.88 (10.71)	84.952	**<.001**
Diastolic BP (mmHg)	78.41 (13.93)	82.00 (13.35)	79.00 (15.78)	83.00 (13.00)	27.887	**<.001**
Systolic BP (mmHg)	126.00 (20.00)	128.00 (21.54)	128.00 (21.00)	129.00 (23.62)	14.250	**.003**
C-reactive protein (mg/L)	1.24 (2.24)	1.67 (3.00)	1.56 (3.07)	2.41 (5.12)	46.804	**<.001**
HbA1c (mmol/mol)	33.40 (4.30)	34.00 (5.25)	33.70 (5.16)	34.75 (5.93)	27.471	**<.001**
HDL-C (mmol/L)	1.52 (0.45)	1.41 (0.44)	1.46 (0.48)	1.33 (0.45)	46.801	**<.001**
IGF-1 (nmol/L)	22.66 (7.38)	23.66 (8.29)	22.99 (7.63)	21.74 (8.83)	14.128	**.003**
LDL-C (mmol/L)	3.22 (0.99)	3.38 (1.00)	3.30 (0.97)	3.39 (1.04)	6.849	.077
SHBG (nmol/L)	63.65 (39.71)	51.66 (36.45)	53.47 (37.65)	48.26 (39.17)	42.538	**<.001**
Testosterone (nmol/L)	1.13 (0.64)	1.19 (0.59)	1.20 (0.60)	1.21 (0.67)	7.899	**.048**
Triglycerides (mmol/L)	1.08 (0.80)	1.27 (0.90)	1.16 (0.78)	1.42 (1.04)	41.959	**<.001**
MET-mins/wk	3165.00 (2626.00)	3230.00 (2816.00)	720.00 (676.75)	594.00 (773.25)	1038.250	**<.001**
Sitting time (hrs/day)	3.00 (1.00)	5.90 (2.00)	3.00 (1.10)	6.00 (3.00)	1042.807	**<.001**

Key: BMI: body mass index; BP: blood pressure; HbA1c: glycated haemoglobin; HDL-C: high-density lipoprotein cholesterol; IGF-1: insulin-like growth factor-1; LDL-C: low-density lipoprotein cholesterol; SHBG; sex-hormone binding globulin; MET-mins/wk: metabolic equivalent of task minutes PA performed per week. Data are presented as median and interquartile range; *W*: test statistic; *P*: significance value.

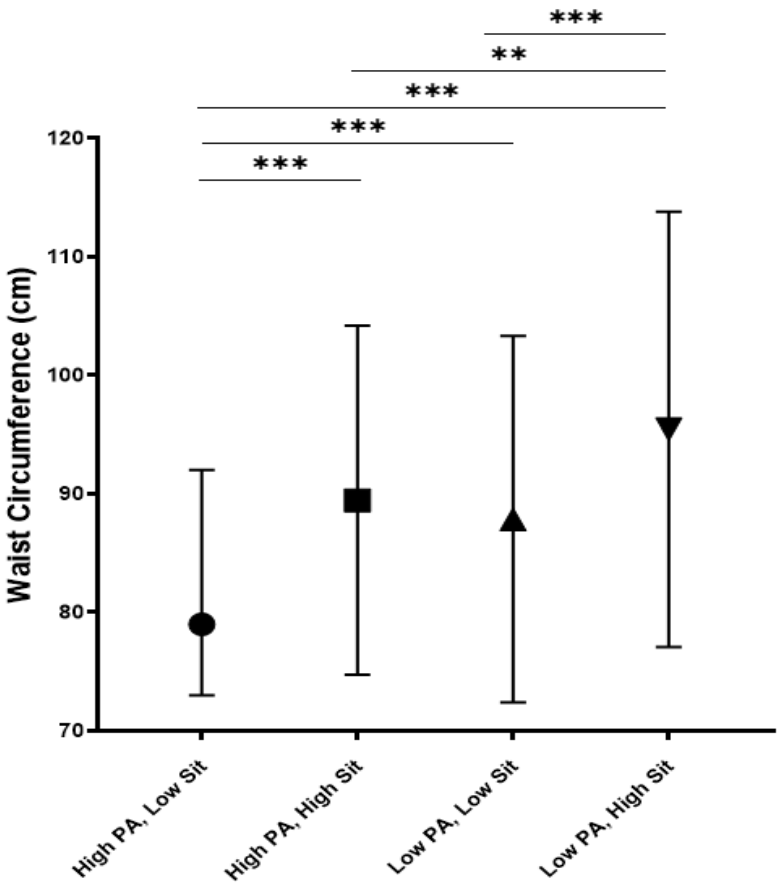


Figure 5.6. Pairwise comparison (Mann-Whitney U) of waist circumference between physical activity (PA) and sedentary time (siting time; Sit) risk groups/categories. ** = *P* <.01; *** = *P* <.001.

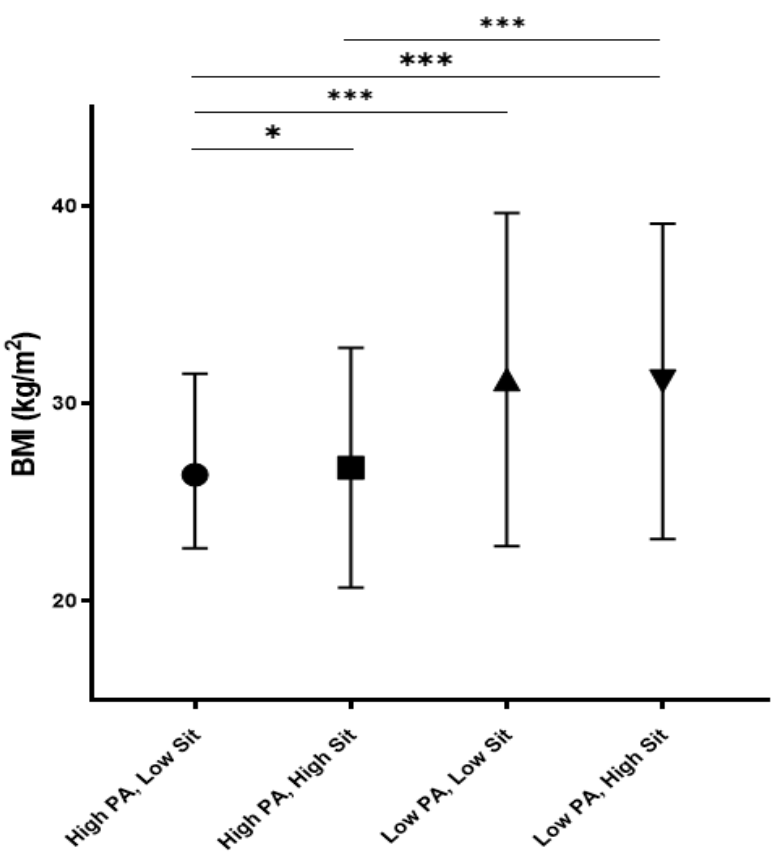


Figure 5.7. Pairwise comparison (Mann-Whitney U) of body mass index (BMI) between physical activity (PA) and sedentary time (sitting time; Sit) risk groups/categories. * = *P* <.05; *** = *P* <.001.

Table 5.12. Pairwise comparisons of physical activity (PA) and sedentary time (siting time; Sit) risk groups/categories. Analysis completed using Mann-Whitney U.

Outcome	Group One	Group Two	*U*	*P*
Diastolic BP	High PA, Low Sit	High PA, High Sit	-2.876	**.024**
	High PA, Low Sit	Low PA, High Sit	-5.005	**<.001**
	Low PA, Low Sit	Low PA, High Sit	-3.587	**.002**
Systolic BP	High PA, Low Sit	Low PA, High Sit	-3.705	**.001**
CRP	High PA, Low Sit	High PA, High Sit	-2.820	**.029**
	High PA, Low Sit	Low PA, High Sit	-6.767	**<.001**
	Low PA, Low Sit	Low PA, High Sit	-4.214	**<.001**
	High PA, High Sit	Low PA, High Sit	-3.516	**.003**
HbA1c	High PA, Low Sit	Low PA, High Sit	-5.012	**<.001**
	Low PA, Low Sit	Low PA, High Sit	-3.801	**.001**
	High PA, High Sit	Low PA, High Sit	-2.935	**.020**
HDL-C	High PA, High Sit	Low PA, High Sit	-3.407	**.004**
	Low PA, Low Sit	Low PA, High Sit	-4.328	**<.001**
	High PA, Low Sit	Low PA, High Sit	-6.742	**<.001**
	High PA, Low Sit	High PA, High Sit	-2.907	**.022**
IGF-1	High PA, Low Sit	Low PA, High Sit	-2.760	**.035**
	Low PA, Low Sit	Low PA, High Sit	-2.799	**.031**
	High PA, High Sit	Low PA, High Sit	-3.381	**.004**
Triglycerides	High PA, Low Sit	High PA, High Sit	-3.354	**.005**
	High PA, Low Sit	Low PA, High Sit	-6.214	**<.001**
	Low PA, Low Sit	Low PA, High Sit	-4.359	**<.001**

Key: CRP: C-reactive protein; HbA1c: glycated haemoglobin; HDL-C; high-density lipoprotein cholesterol; IGF-1: insulin-like growth factor-1; *U*: Mann-Whitney test statistic; *P*: statistical significance.

Table 5.11 presents the descriptive statistics [median (IQR)] of participants classified in each PA and siting time risk group. Furthermore, Kruskal-Wallis test results are presented demonstrating that all outcomes, apart from age and LDL-C, contain statistically significant between groups differences. Pairwise comparisons were performed using Mann-Whitney U tests (Table 5.12, Figures 5.6 to 5.9), showing that the greatest number (n = 11), and magnitude of differences were between the High PA, Low Sit and the Low PA, High Sit group. In addition, when the High PA, Low Sit was compared with the High PA, High Sit group (n = 7), and when the Low PA, Low Sit was compared with the Low PA, High Sit group (n = 7), there was an identical number of statistical findings. This tends to support the notion that sedentary behaviours have an integral role to play in health maintenance.

When the High PA, High Sit was compared to the Low PA, High Sit group, a further five differences were observed, and a further three when the High PA, Low Sit compared with the Low PA, Low Sit group. There were no observed differences for any outcome between the High PA, High Sit and the Low PA, Low Sit group. Without exception, outcomes favoured the risk group with more PA and less sitting time.

Finally, morbidity prevalence was reported for the lowest risk group (Figure 5.10), and comparisons made to the highest risk group. It was observed that there was a 63% greater chance of having PCOS if in the high-risk group. Furthermore, for all other measured morbidities, prevalence was greater and risk increased if in the high-risk group.

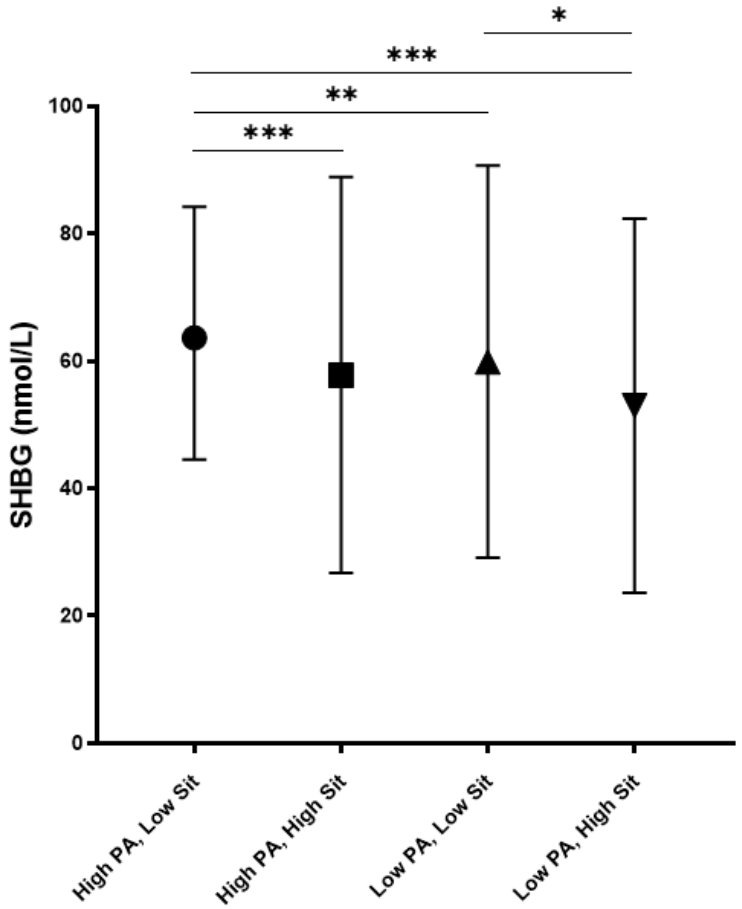


Figure 5.8. Pairwise comparison (Mann-Whitney U test) of sex hormone binding globulin (SHBG) between physical activity (PA) and sedentary time (sitting time; Sit) risk groups/categories. * = *P* <.05; ** = *P* <.01; *** = *P* <.001

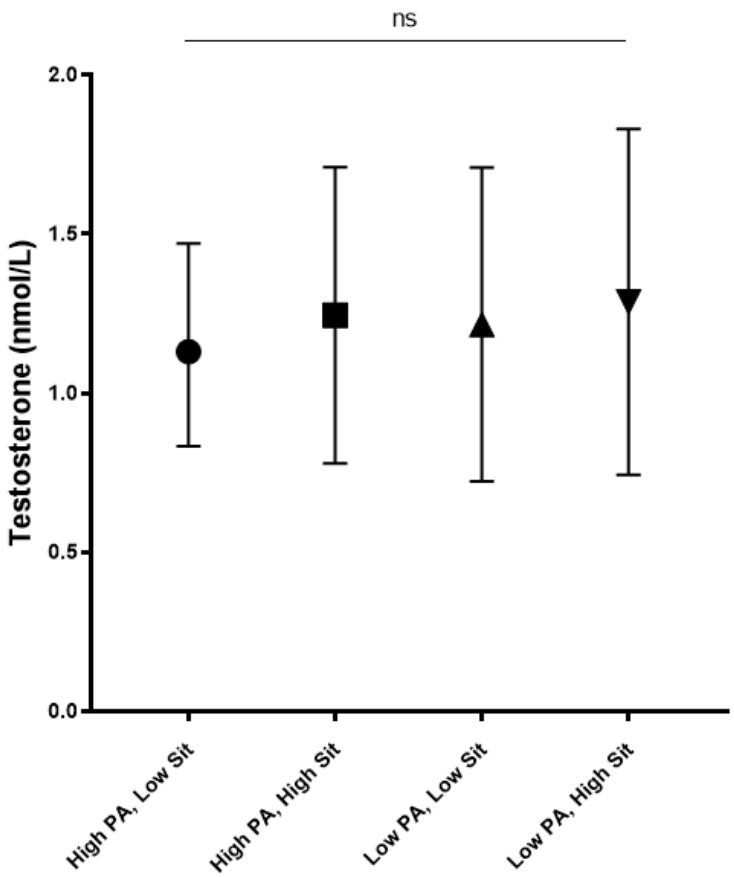


Figure 5.9. Pairwise comparison (Mann-Whitney U test) of testosterone between physical activity (PA) and sedentary time (sitting time; Sit) risk groups/categories. Ns: not significant

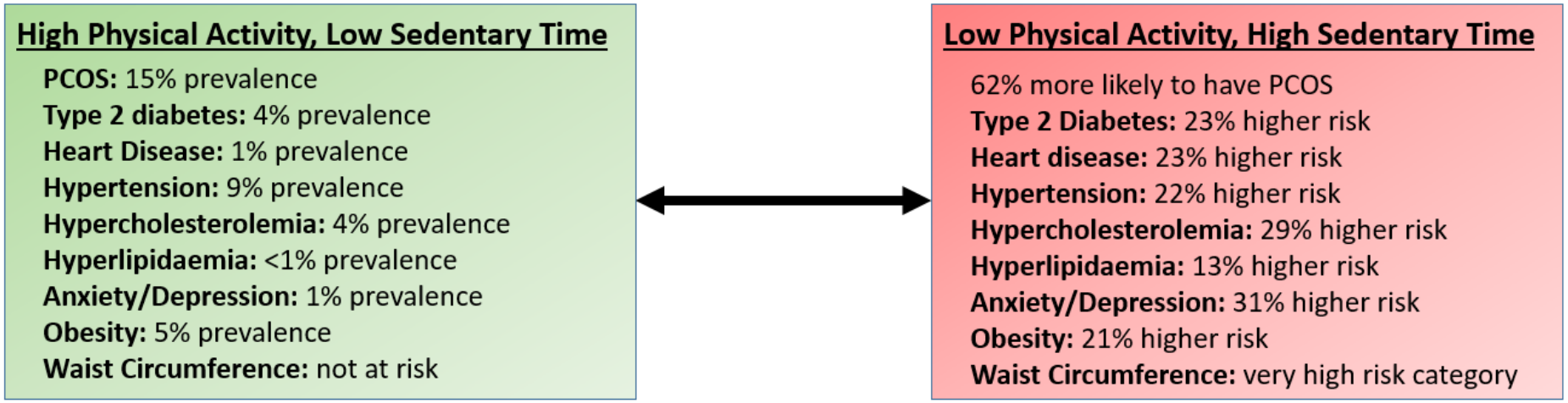


Figure 5.10. Prevalence of morbidity for those with low risk behaviours (high physical activity and low sedentary/sitting time) and relative risk of disease incidence for those engaged in high risk activities (low physical activity and high sedentary/sitting time)**.** Waist circumference is used as a surrogate for central obesity.

5.5. Discussion

5.5.1. Key Findings

The current chapter aimed to build upon the findings of the previous chapter by addressing some of the research gaps that were identified. The initial analyses looked for differences between the three study groups. For outcomes related to body weight/composition, there were statistical differences between women with PCOS and the age-matched group; women with PCOS were heavier, with higher BMI and larger waist circumferences. However, these findings were not unique to women with PCOS, since the same differences were evidenced between the age + BMI-matched cohort and the age-matched only group. Because the women in the age + BMI-matched group were BMI matched to the women with PCOS, these findings may not be particularly surprising. However, this suggests that there appear to be no significant differences in the degree, or indeed the distribution of adiposity in women with PCOS when compared to an age + BMI-matched group without PCOS.

Similarly, no observed difference was identified for numerous additional outcomes between women with PCOS and the age + BMI group; including DBP, IGF-1, testosterone, and CRP. However, both the women with PCOS and the age + BMI-matched group had statistically less-favourable values than the age-matched only cohort. Furthermore, there were certain notable differences between the women with PCOS and the age + BMI-matched cohort, with the former exhibiting a less favourable profile regarding HDL-C, triglycerides, HbA1c and SHBG. The magnitude of these less favourable effects was greater when women with PCOS were compared to the age-matched group. Although of a smaller magnitude than those with PCOS, the age + BMI-matched women also had statistically lower HDL-C and SHBG, and greater levels of triglycerides than the age-matched only group.

In agreement with the findings presented in Chapter 4, there were no statistically significant between group differences in PA performed per week (either MET-mins/week or summed mins/week). However, when PA intensity was assessed, statistical differences in MET-mins/week achieved through vigorous intensity activity was lower in women with PCOS than either comparator group. There was no corresponding difference for MET-mins/week achieved through moderate intensity activity. Differences were also observed for sedentary behaviours; screen time and summed

sedentary time was statistically higher in women with PCOS than the age-matched only group. To further support these findings, when the IPAQ categorisation data was assessed for each group, a greater proportion (~9% more) of women with PCOS were in the low activity category, and fewer in the high-activity category, than either comparator group. Women with PCOS also had the smallest proportion meeting PA guidelines (*i.e.,* 150 mins/week moderate intensity, 75 mins/week vigorous intensity or a combination of both) regardless of whether walking was accounted for in the calculation.

In addition to the outcomes noted above, all participants were also asked to rate their satisfaction on a number of psychosocial variables. Women with PCOS tended to view their own health less favourably (*i.e.,* perceived health rating and health satisfaction) than either comparator group. Similarly, women with PCOS largely reported less favourable scores for happiness, familial satisfaction, and financial satisfaction. Some of these less favourable scores may be due to incidences of comorbidity in women with PCOS. Prevalence of diagnosed type 2 diabetes, hypertension, hypercholesterolemia, obesity, kidney disease, chronic lung conditions and non-invasive cancers were statistically higher in women with PCOS compared to the other two study groups. Indeed, when total morbidities and only those associated with metabolic health (type 2 diabetes, hypercholesterolaemia, hyperlipidaemia, and obesity) were compared, they were statistically higher than either of the other groups.

When cluster analysis was completed, two distinct clusters were identified. Based on the outcomes used to cluster, the average values in Cluster One (n = 1,008) were without exception, more favourable. Cluster Two demonstrated inferior health in all of these key biological markers. The majority of the age-matched (77.9%) and the age + BMI-matched (56.1%) participants were included in Cluster One. Indeed, it was only the women with PCOS that had their majority in Cluster Two and although this was approximately half (51.3%). In fact, 28% of Cluster Two was comprised of women with PCOS, compared to only 15% in Cluster One. The derived clusters were then used as categorical variables in subsequent analyses where outcomes relating to PA levels were compared. The total MET-mins/week of PA performed were statistically higher (MD: 535.18 MET-mins/week; d =

0.226) in the cluster with a more favourable health profile (as measured by circulating factors). In addition, those with a better health profile also spent less time (MD: 1.04 hrs/day; *d* = -0.434) engaged in sedentary behaviours each day. The number of morbidities (total and metabolic) were compared between clusters; no statistical differences were observed but, when separated by study comparator group membership, women with PCOS had the highest number of morbidities even when their PCOS diagnosis was excluded from the analysis.

The final analysis grouped all participants into categorical groups based upon their total level of weekly PA (MET-mins/wk) performed and their time spent in sedentary time each day. Women with PCOS had the largest proportion of their group in the high-risk group (Low PA, High Sit), although the age + BMI-matched group closely followed this. When membership of the low-risk group was analysed (High PA, Low Sit), women with PCOS had the smallest proportion of their cohort in this group. Conversely, the age + BMI-matched cohort had the highest proportion of their group in the low-risk group. It should be noted that the age + BMI-matched cohort had only 4.4% of their total participants excluded from this analysis due to incomplete data (compared to PCOS: 21.3%, and age-matched: 17.7% missing). When biomarkers were compared between these risk groups, statistical effects were observed for all outcomes apart from LDL-C. When pairwise comparisons were made to identify where the differences lay, it became apparent that the highest number of differences were between the lowest risk group and the highest risk group. Furthermore, the magnitude of these differences, without exception, was larger than any other pairwise comparison; from a health perspective, the values observed in the low-risk group for each outcome were favourable. Similarly, if categorised as high-risk, there was a 63% greater chance of having PCOS when compared to the low-risk group and furthermore, risk factors for metabolic syndrome and CVD all carried a greater risk of morbidity.

The implications of these findings mean that:

The null hypothesis that women with PCOS will not have a less favourable physical and psychological health profile than their counterparts without PCOS can in all likelihood be rejected. Compared to the age-matched cohort, women with PCOS had worse indicators of health across the majority of explored variables. When compared with the age + BMI-matched cohort, there were more similarities (particularly those linked to anthropometry). However, even where statistical significance was not evident, recorded values tended to be less favourable.

For the second hypothesis, the null hypothesis must be rejected. The variation in metabolic health markers between women with PCOS and the BMI + age-matched cohort is negligible. Whilst there are many differences between women with PCOS and the age-matched only cohort, it is hard to ascertain the cause of these differences. Indeed, excess weight appears to be a key indicator of worsened health. However, it should be noted that when cases were grouped into two clusters, women with PCOS had the highest proportion (>51%) of their cohort as members in the cluster that would be classified as having poorer health. The BMI + age-matched cohort (~44% membership) and age-matched (~22% membership) were predominantly allocated into the cluster with more favourable health outcomes.

The null hypothesis that there would be no statistical between group differences in PA performed and time spent in sedentary behaviours must also be rejected. Whilst there were no differences in total MET-mins/wk or total minutes, women with PCOS performed significantly less vigorous PA per week, and spent more time engaged in sedentary behaviours, than their age-matched counterparts. It can however be accepted that regardless of PCOS diagnosis, those who are the most active, and spend the least amount of time sitting have the lowest cardiometabolic risk.

5.5.2. Between group differences

When outcomes associated with anthropometry and body composition were analysed, it was observed that women with PCOS had a significantly higher body weight, BMI, waist circumference

and body fat percentage than the age-matched women without PCOS. Whilst these findings agree with those presented in Chapter 4, the findings in the present study offer increased confidence because, as opposed to the self-report data captured in the previous chapter, data here were objectively measured in the UK Biobank testing centres. Whilst it has been widely reported that women with PCOS tend to be more overweight than their non-PCOS counterparts (Kiddy *et al.*, 1990; Balen *et al.*, 1995; Yildiz *et al.*, 2008), it is a reasonably novel approach to compare the body compositional outcomes of these women to (age and) BMI-matched women without PCOS. The women in the age + BMI-matched cohort within the current study were selected based upon their lack of a PCOS diagnosis, but also because their BMI was closely matched to the women with PCOS. What was not factored into their inclusion were their waist circumference or their body fat percentage. Previous studies have cited that the distribution of adipose tissue in women with PCOS tends to be more central (Talbott *et al.*, 1995; Taponen *et al.*, 2003), resulting in greater metabolic disruption. However, this does not appear to be evident across the current sample; there are negligible differences for waist circumference (MD: 1 cm) and body fat (MD: 0.01%) between the two cohorts.

These findings tend to agree with previous studies albeit in much smaller samples; Orio *et al*. (2004) compared 30 overweight women with PCOS to an age + BMI-matched group and reported no statistical differences in waist-hip ratio (WHR). Similarly, Dolfing *et al.* (2011), compared a small group (n=10) of women with PCOS with 10 matched controls and found no significant differences in WHR or fat mass in these groups. The findings of the present study add confidence to these previous studies by including a sample size that is approximately 10 times greater; although women with PCOS are more likely to be overweight or obese than women without PCOS, when BMI is matched, the differences in the volume and distribution of adipose tissue is negligible.

It is common knowledge that engaging in regular PA can help to minimise weight and fat gain and that maintenance of a physically active lifestyle is an effective strategy to improve obesity-related health outcomes irrespective of any concomitant weight loss (McArdle *et al*., 2015). However, when the PA levels of these women are compared, there are no statistical differences between each study cohort for the MET-mins/wk performed or the weekly minutes spent engaged in PA of any duration.

These findings agree with those from Chapter 4 and a number of previous studies that included measures of self-reported PA (Moran *et al.*, 2013; Douglas *et al.*, 2006; Wright *et al.*, 2004; Alvarez-Blasco *et al.*, 2011; Cutler, Pride and Cheung, 2018). Lin *et al.* (2019) further supported these findings by incorporating objective measures of PA and assessing PA intensity levels; they reported no statistical differences in duration, type or intensity of activity. Whilst the UK Biobank did not capture the type of PA (beyond walking) that participants were engaged in, the relative intensity was captured and this is where, in contrast to Lin *et al.* (2019), significant differences were identified. Women with PCOS achieved fewer MET-mins/week through vigorous PA than either the age-matched (MD: 185.8 MET-mins/week; P <.001) or BMI + age-matched (MD: 104.8 MET-mins/week; P = .005) comparator groups. No corresponding differences were observed between the age-matched and the BMI + age-matched despite the differences in BMI, body fat and waist circumference. Lower body fat percentage has previously been associated with higher levels of Vigorous PA, but not for moderate PA (Gutlin *et al.*, 2005). The importance of Vigorous PA levels in the maintenance of healthy weight is unclear.

Another contrasting factor, is time spent engaged in sedentary behaviours; compared to the age-matched cohort, women with PCOS had more screen time (TV viewing and computer usage) and more total sedentary time (0.4 hrs/day). Sedentary time is a known risk factor, independent of PA (Biswas *et al.*, 2015), for a multitude of undesirable health outcomes, and, although the sedentary behaviours may also be a contributing factor to less favourable body composition, similar differences were not evident between the BMI + age-matched and the age-matched cohort. This could suggest that sedentary time may not be a key factor in outcomes relating to body composition for this population, but the chronic effects of longer-term increased sedentary time may lead to future health complications.

However, what is not reported in the current study is the opposing side to the energy balance equation, which is calories consumed. Food diaries and other measures of caloric consumption were not reported in the present study, but differences in dietary habits almost certainly play a role in body composition (Schröder *et al.*, 2004), and metabolic health (Fung *et al.*, 2009), as well as in health-

related quality of life (Milte *et al.*, 2015). Despite the lack of reporting dietary habits in the current study, numerous studies have previously reported no differences between the caloric intake of women with PCOS compared to their non-PCOS counterparts (Douglas *et al.*, 2006; Wright *et al.*, 2004; Alvarez-Blasco *et al.*, 2011; Cutler, Pride and Cheung, 2018, Lin *et al.*, 2019). In fact, from searching the relevant literature, only three contradictory studies were located. Indeed, Moran *et al.*, (2013) cited higher daily energy [+215 KJ/day (~50 kcal); P = .02] consumption in women with PCOS compared to controls, but the women with PCOS incidentally had higher body weight and BMI than the control group. This study also factored in PA levels, reporting no statistical differences in PA performed, yet women with PCOS were sedentary for 30 mins more per day. Zhang and colleagues (2015) noted a considerably greater daily energy consumption in women with PCOS (~3664.1 KJ/day; P = .01) compared to BMI-matched healthy controls. Finally, Ahmadi *et al.* (2013) report that women with PCOS consumed ~300 kcal more per day than a healthy BMI-matched group, although, despite this increased calorie consumption, there were no statistical differences in body weight or BMI between these groups.

Collectively, the findings of this chapter (and Chapter 4), alongside the findings of previous studies, indicate that women with PCOS have increased body weight and BMI compared to age-matched women without PCOS, but it appears that PA levels are not mediators or predictors of this phenomenon. The relationship between PCOS, BMI and sedentary time may demonstrate a causal effect; women with PCOS are inactive for longer periods each day when compared to age-matched controls. However, the BMI + age-matched cohort showed no statistical variation from the age-matched group despite much less-favourable body composition; they also did not differ from the PCOS cohort. Thus, it is likely that, the magnitude of differences in PA and sedentary time are not responsible for observed differences in BMI, but as in any population, they may be contributing factors. These findings highlight the need for a comprehensive study that incorporates (preferably objective) measures of PA and sedentary time alongside analysis of dietary composition in order to gain a fuller understanding of their impact.

If factors pertaining to lifestyle are not fully implicated in overweight and obesity in PCOS, the findings of this study may be able to implicate some additional mechanisms. Teede *et al.* (2013) state that PCOS is an independent risk factor for overweight and obesity and that a bi-directional interaction exists, where PCOS causes a physiological and/or psychological environment that promotes weight gain and weight that is gained leads to increased prevalence and severity of PCOS symptoms. It is clear from the study results that the less-favourable values observed for most outcomes are likely due to excess body weight and/or increased adiposity. Indeed, for women with PCOS and the BMI + age-matched cohort, the majority of circulating biomarkers and physical measurements are significantly worse than the age-matched (and lower weight) cohort. However, four outcomes are statistically less favourable in women with PCOS than the BMI + age-matched study group, namely differences in HbA1c, HDL-C, triglycerides and SHBG levels between these groups, which will be the focus of the following paragraphs.

One potential mechanism that demonstrates the aforementioned bi-directional relationship is the elevated HbA1c observed in PCOS. The current study found that women with PCOS had significantly higher levels of HbA1c than their counterparts without PCOS, even where participants were age and BMI-matched. HbA1c develops during hyperglycaemia when glucose bonds with haemoglobin in the blood and is typically used as a diagnostic/monitoring biomarker for diabetes/glycaemic control, typically representing the of average plasma glucose over the previous 8-12 weeks (Nathan, Turgeon and Regan, 2007). Considering that a normal value for HbA1c is < 42 mmol/mol (WHO, 2011), the average HbA1c values for women within this study are normal, but those recorded for women with PCOS are closer to the threshold for pre-diabetes (42-47 mmol/mol). Despite statistical significance (P <.001), the small magnitude of these differences tends to agree with the findings of previous studies. Ollila and colleagues (2017) reported similar values to those in the current study and when statistically controlling for variation in BMI, reported a statistical (but likely not clinically relevant) difference in HbA1c (P = .029). Sadaria and Ravi (2015) reported small variation between groups, but no statistical effect for HbA1c levels between age and BMI-matched women with and without PCOS.

In fact, it appears that it is only when the PCOS and control groups are not BMI-matched (that is the women with PCOS have higher BMI) that statistical and clinically relevant variation is observed in HbA1c. In age-matched women, Bala *et al.* (2017) noted elevated HbA1c and a markedly higher prevalence of type 2 diabetes and prediabetes in women with PCOS compared to controls. Similarly, additional studies have made similar comparisons and found again that women with PCOS had higher HbA1c levels than their counterparts without. However, these women were either not age- or BMI-matched (Ansari *et al.,* 2017; Fuad and Alobaidy, 2018), or matching was not specified (Sathisha *et al.,* 2020). A further study investigated prevalence of prediabetes (using HbA1c ≥ 5.7% as a diagnostic cut-point) in obese and lean women with and without PCOS (Kim *et al.,* 2012). Whole cohort analysis (unmatched groups) revealed HbA1c to be higher (~20.9 mmol/mol) in women with PCOS and this was further supported by diagnoses of prediabetes. Of the non-obese women with PCOS, 31% were prediabetic compared to ~6% of controls. Although women with obesity and PCOS had a higher prevalence of prediabetes, there was more parity in these groups (23.5% vs 20% respectively), albeit from a smaller sample (n = 32).

It is also pertinent to note that elevated HbA1c may promote weight gain (Ridderstråle *et al.,* 2016). Higher HbA1c is indicative of a chronic elevation of blood glucose, which in turn increases insulin production. Insulin mediates facilitated diffusion of glucose into cells and when the body has met its energy demand quota, and glycogen stores are full, excess carbohydrates are converted into triglyceride molecules where they are stored as fat in adipose tissue, leading to weight gain (McArdle, Katch and Katch, 2015). It seems that HbA1c is elevated in women with PCOS and that there is a positive correlation with several anthropometric, metabolic and androgenic factors (De Medeiros *et al.,* 2014). Indeed, in the general female population, HbA1c level is associated with CVD risk (Singer *et al.,* 1992), and this is further compounded in PCOS (Kim *et al.,* 2012).

Between group differences were identified for lipid related parameters (triglycerides and HDL-C). Given the known associations that elevated triglycerides (Hjellvik, Sakshaug and Strøm, 2012) and HDL-C (Williams *et al.,* 1994) have with overweightness, obesity and cardiometabolic risk, it is unsurprising that both women with PCOS and the age + BMI-matched groups had significantly less

favourable values for these lipidaemic parameters than the age-matched only group. However, the observation that women with PCOS had less favourable values for circulating triglycerides and HDL-C than their BMI + age-matched counterparts was less expected. In the diagnostic criteria for metabolic syndrome (MetS), the diagnostic cut off point for triglycerides is ≥ 150 mg/dL (WHO, 1998; IDF, 2005; NCEP ATP III, 2005), and for HDL-C in women is < 50 mg/dL (IDF, 2005; NCEP ATP III, 2005) (Alberti and Zimmet, 1998, Alberti, Zimmet and Shaw, 2006). When using standard conversion factors (triglycerides: mmol/L * 88.57; HDL-C: mmol/L * 38.67), the mean values reported for both these factors in the PCOS study group indicate that mean circulating triglyceride levels are indeed > 150 mg/dL, and that, although just above the aforementioned cut-off point, the mean HDL-C levels (52.6 mg/dL) are also low.

Indeed, certain abnormalities in lipidaemic profiles, including low circulating levels of HDL-C and elevated triglycerides, are associated with increased CVD risk (Wilson *et al.* 1988; Bass *et al.,* 1993), and are strong independent predictors of CVD-related death in women. Another used method of assessing CVD risk utilises plasma triglyceride to HDL-C concentration ratio (TG/HDL-C), with 2.5 mg/dL^{-1} used as a cut off point for being at risk (Salazar *et al.,* 2012). It has been further reported that an elevated TG/HDL-C ratio is as effective as the MetS diagnostic criteria at predicting the development of CVD (Salazar *et al.,* 2013). Results from the current study indicate that women with PCOS are the only group that cross this threshold (TG/HDL-C ratio: 2.94 mg/dL^{-1}) indicating that they are at an increased risk of cardiometabolic illness. Moreover, dyslipidaemia in women with PCOS has been previously reported (Kousta and Franks, 2007), with previous findings in line with those of the current study. Indeed, women with PCOS have been found to have higher circulating triglycerides (Talbott *et al.,* 1995; Talbott *et al.,* 1998; Pirwany *et al.,* 2001), and lower HDL-C concentrations (Wild *et al.,* 1985; Talbott *et al.,* 1995; Talbott *et al.,* 1998) when compared to women without PCOS. Other studies also reported higher total cholesterol, LDL-C and very low-density lipoprotein cholesterol levels, which persisted when controlling for overweightness and obesity (Mather *et al.,* 2000; Vrbikova *et al.,* 2003).

Finally, the current study identified that SHBG was significantly lower in women with PCOS compared to either control group. A key feature of PCOS is hyperandrogenism; ovarian androgens, androstendione and testosterone are frequently elevated (Carmina, 2002) resulting in the common features of hyperandrogenism in PCOS, such as acne and hirsutism, androgenic alopecia and even virilisation that patients often present with (Pasquali *et al.*, 1997; Davison *et al.*, 2005; Elmlinger *et al.*, 2005). SHBG concentration plays an important role in the metabolic clearance of sex hormones, as well regulating their bioavailability to target tissues. SHBG levels can vary substantially between individuals since its hepatic production and circulating concentrations are influenced by nutritional, hormonal and metabolic factors (*e.g.*, insulin levels) (Allen *et al.*, 2002) as well as diseases such as PCOS. Accordingly, reduced SHBG concentrations are widely reported in women with PCOS compared to those without and is often associated with overweight or obesity and insulin resistance (Teede, Meyer and Norman, 2005). Thus, even when testosterone levels are within the normal range, decreases in SHBG levels result is an increase in the bioavailable (free) testosterone/androgens (Deswal, Yadav and Dang, 2018), consequently increasing the peripheral androgen activity and potentially leading to manifestation of symptoms of hyperandrogenism). Despite the lack of reported PCOS-specific symptoms in the current study, the differences in biomarkers between each group tend to support the above points. In the present study there are no statistical or clinical differences between groups for total testosterone, yet the women with higher BMI have decreased SHBG and this is further compounded by a diagnosis of PCOS.

5.5.3. Physical activity and ill health

The results from the current chapter also provide additional information about the contribution of PA to a range of health-related parameters in women with PCOS (and indeed the control groups). Whilst the previous chapter found that women with PCOS had poorer HRQoL compared to a control group, it also revealed that there was no significant relationship between PA and HRQoL. Although the measures in the current chapter are still based upon self-reported data, the presented results indicate

that the PA is indeed an important component in the reduction of CVD risk and the maintenance of health.

The performed cluster analyses grouped cases together based upon their within cluster similarities, whilst making them as varied as possible from other clusters based upon the outcomes used for clustering. In the current study, this analysis revealed that the best solution separated cases into two clusters and each outcome included was statistically different to the comparable data in the other cluster. Further than that though, the analysis has essentially created a cluster (cluster one) with a profile of more favourable measurements across all outcomes compared to the second cluster; that is cluster one members generally have reduced central adiposity, lower BP, improved lipidaemic parameters and less free circulating androgens. In order to gain a better understanding of this, cluster membership was investigated based upon PCOS and other group membership. Women with PCOS have a higher proportion of their group as members of cluster two and make up a greater proportion of cluster two than cluster one. This is not surprising, considering the aforementioned results indicating that women with PCOS had less favourable levels of circulating biomarkers than either of the other study groups. The other factors that may have influenced cluster membership were of course those relating to PA and sedentary behaviour.

Chapter 4 revealed no between group (PCOS versus control) statistical differences for either weekly PA performed or sitting time and this is partially supported in the current study. There were no between group differences for total weekly PA performed, but women with PCOS completed less vigorous PA than either of the other cohorts and, they also spent more time engaged in sedentary behaviours than the BMI + age-matched group. Given the relatively small magnitude of these differences, it is questionable whether this alone could explain variation in health markers. Total MET-mins/wk and daily sitting time were used as dependent variables in statistical analyses, and it was revealed that women in the cluster with the healthier profile (cluster one) engaged in ~550 MET-mins/wk more PA, and ~1 hr/day less sitting time than those in cluster two.

Notably, current PA guidelines (CMO, 2019) recommend that activities equivalent to 600 MET-min/wk are beneficial for good mental and physical health. In this study, the average values reported

by both clusters, far exceed this recommendation. The all-cause mortality benefits of PA are based upon an inverse linear dose-response relationship (Lee and Skerrett, 2001); that is, the more PA performed, then the longer life an individual will have. In fact, it has previously been revealed that, when compared to those engaging in no PA, individuals performing 3-5 times the PA recommendations (as is reported in the current study) have a 39% lower risk of premature all-cause mortality (Arem *et al.*,2015). Furthermore, an additional 4.5 years in life expectancy can be gained when weekly PA exceeds 1350 MET-mins/wk (Moore *et al*., 2012). When considering morbidity rather than mortality, it is those completing the greatest volume of PA that have the greatest risk reduction for chronic disease. For example, those achieving 5-6 times the PA guidelines have relative risk reductions of 28% for type 2 diabetes, 26% for ischaemic stroke, and 26% for coronary artery disease compared to those falling below the guidelines (Kyu *et al.,* 2016). From the data reported in the current study, both clusters should benefit from the protective effects of being physically active, but there is a marked difference in health status. This variation may be explained in part by the difference in weekly PA, but the magnitude of beneficial effects becomes less marked when comparing relatively high volumes; for example, those performing 22.5-<40 MET-hrs/wk or 40-<75 MET-hrs/wk have identical mortality risk (0.61) over time in a study by Arem *et al.* (2015).

Another possible explanation for health variability may concern time spent in sitting behaviours. Sedentary behaviour is an independent risk factor for chronic disease and premature mortality (Owen *et al*., 2010). In fact, it has been demonstrated that individuals must complete high levels (> 35.5 MET-hrs/wk) of PA to protect against prolonged bouts of sitting (Ekelund *et al*., 2016). In a longitudinal study (Saidj *et al.*, 2013), prolonged daily sitting has been demonstrated to be a predictor of fasting insulin, VO_2max and waist circumference, independent of PA.

Some potential mechanisms to explain the independent association of sedentary behaviours with cardiometabolic health have been purported; lipid and glucose metabolism are severely impaired during sedentary activities and this is thought to be due to the inactivity of large skeletal muscles that are usually engaged when standing/moving (Hamilton, Hamilton and Zderic, 2004; Tremblay *et al.*, 2010). Another suggested mechanism relates to the dysregulation of haemodynamic vascular

signalling; that is an absence of exercise-induced vascular dilation which promotes inflammatory-mediated atherogenesis (Raschke and Eckel, 2013; Young *et al.*, 2016) leading to a narrowing of blood vessels. In addition to the physiological responses to sitting, Bowman (2006) suggests that when engaged in sedentary behaviour, particularly TV/computer usage, diet is substantially changed; total energy intake is increased, and macronutrients consumed are less desirable. These two factors (total calorie intake and poor diet quality) are also risk factors for premature mortality and determinants of type 2 diabetes (Kant *et al.*, 2000). Finally, the timings of sedentary behaviour may also be key in determining the degree of harm; Almoosawi *et al.* (2012) suggest that prolonged screen time is more likely to occur in the evening, following most people's main meal of the day. The implication is that postprandial glucose and lipids are elevated with a potentially harmful effect to cardiovascular health.

What is clear from the current study is that, compared to the most active individuals, those who are the least physically active and spend the most amount of time in sedentary behaviours, have the least favourable markers of metabolic health regardless of their health status (*i.e.*, PCOS or control). Unfortunately, it appears that women who have PCOS are more likely to be classified as high risk based on their activity levels than low risk; this categorisation means that they have an increased risk of developing other chronic health conditions. In this context, it is reasonable to assume that unless they become more physically active and spend less time sitting, their metabolic health could further deteriorate and the sequelae of PCOS could be further exacerbated.

5.6. Limitations

Whilst the obvious strengths of the current study are the large number of participants with PCOS that are included, and that large comparator groups of matched participants were also analysed, the lack of PCOS-specific data used to define participants is a potential limitation. The UK Biobank project reports participants' health conditions, allowing for the women with PCOS to be identified, but it does not report any further diagnostic criteria and since, as previously described, there are different phenotypes within PCOS based on the applied diagnostic criteria which are not defined/reported

within the current study. Whilst Chapter 4 of this thesis reported negligible differences between phenotypic subgroups for a range of self-reported outcomes, including perceived quality of life, it was not possible to make similar comparisons with the outcomes included in the present study. It has previously been reported (Diamanti-Kandarakis and Panidis, 2007) that there are variations in the biochemical profiles of women with PCOS dependent upon the phenotype with which they have been diagnosed. One such variation is the degree of insulin resistance and/or metabolic syndrome (Legro, 2007), and its increased prevalence in those who are hyperandrogenic and anovulatory (Mehrabian *et al.,* 2011). The identification of women with PCOS who are at increased risk of type 2 diabetes and CVD could have been facilitated if phenotypic prevalence measures were completed in this study.

Another limitation of the current study surrounds the omission of several outcomes that are specific to PCOS and its treatment. Circulating hormones such as luteinising hormone (Liu *et al.,* 2012), follicle stimulating hormone (Elsheikh and Murphy, 2008), anti-Mullerian hormone (Piouka *et al.,* 2009) and progesterone (Li *et al.,* 2014) are all implicated in the pathophysiology of PCOS and the severity of its symptoms. However, these hormones are not available in the current UK Biobank data set. Furthermore, had it been possible to include PCOS specific quality of life measures such as the PCOS-Q, alongside those circulating biomarkers that have been reported in the current study, an objective assessment of key metabolic risk factors could have been conducted.

The age of included participants is another limitation of this study. UK Biobank participants are aged between 40 and 69 years; whilst PCOS symptoms are evident in reproductive-aged women and the characterisation of the related sequalae in post-menopausal women with PCOS remains understudied (Schmidt *et al.,* 2011). PCOS is indeed a heterogeneous condition with a complex pathophysiology (Kovacs and Norman, 2007), but it is also a dynamic syndrome that manifests with different clinical and metabolic traits throughout the reproductive age (Tannus *et al.,* 2018). Young women of reproductive age, tend to present with the full features of hyperandrogenism, anovulation and polycystic ovaries (depending upon diagnostic criteria). In contrast, women with PCOS over the age of 40 years tend to report reduced incidence of hyperandrogenaemia (Hsu, 2013).

There are also potential limitations in the sedentary behaviour data; because the UK Biobank data does not explicitly ask about total daily sitting time, there are some uncertainties in the data. Firstly, the data sums total minutes driving, using a computer and watching TV; this may miss other sedentary activities from participants' day-to-day routines. There is some evidence however (Prince *et al.*, 2020) that multi-item questionnaires may result in more accurate reporting of self-report sedentary time. In addition to questionnaire format, there is no evaluation of when these sedentary periods occurred making it difficult to establish the role of the aforementioned mechanisms in metabolic health. Finally, it is unclear whether the sedentary bouts were continuous or accumulated throughout the day; Healy *et al.* (2008) found that interrupting prolonged bouts of sedentary time with short activity breaks was beneficial for a range of CVD risk factors.

Finally, another potential study limitation relates to the prevalence of PCOS within the UK Biobank sample. Conservative estimates report that 6-8% of reproductive-aged women are affected by PCOS (March *et al.*, 2010), but some estimates report prevalence as high as 21% (Lizneva *et al.* 2016). Here, from a sample of ~250,000 women, only 319 were classified as having PCOS, which equates to 0.13% prevalence in this sample. The exact reasons for this reduced prevalence is unclear, but it may be partly explained by the age of the population that is included in the UK Biobank sample.

5.7. Conclusion

The findings of the current study build upon those of the previous chapter. Whereas the previous chapter identified that women with PCOS had less favourable self-reported anthropometric measures and perceived HRQoL, the current study benefitted from more objective measures. Similarly to previous findings, when compared to a random sample (age-matched cohort), more women with PCOS were classified as having overweight/obesity with poorer markers of metabolic health. The findings that women with PCOS are at a greater risk of poor cardiometabolic health are generally supported by the existing relevant research. Moreover, the current study also benefitted from a large comparator group that was matched both for BMI and age, thus removing a variable that potentially has a large mediating role in an individual's health (Trent *et al.*, 2005).

The current study did not include a specific HRQoL measurement (such as the SF-12), but it was revealed that a higher proportion of women with PCOS self-rated their health as "poor"(+10.4%), or that they were "extremely unhappy*"* with their health satisfaction (+6.5%) compared to their BMI + age-matched counterparts. This may be a fairly crude indicator that PCOS, as opposed to elevated BMI, is having a negative effect upon their own health perceptions. However, the objectively measured metabolic and hormonal biomarkers also indicate that women with PCOS have worse health profiles, or increased risk of worse health, than their counterparts without PCOS. It is true that the greatest observed differences were between women with PCOS and the age-matched cohort only, whilst it is also pertinent that statistical differences were lacking when women with PCOS were compared to the BMI-matched group. This is indicative that increased body weight and/or adiposity has a significant role in metabolic dysregulation and clearly, if women with PCOS are more susceptible to overweight and obesity (Lim, Davies and Moran, 2012), this should be a treatment priority. However, there are a number of outcomes (HDL-C, triglycerides, HbA1c and SHBG) that are less favourable for women with PCOS even when controlling for BMI. This suggests that PCOS further disrupts biological processes which has a two-fold effect; that is further exacerbation of PCOS symptoms and increased risk of CVD and or type 2 diabetes.

The current study also agrees with previous studies (and the previous chapter) in that there are no differences in the amount of PA performed between women with PCOS and the control groups. However, in contrast to many previous studies, it was apparent that women with PCOS performed less vigorous PA each week and spent more time in sedentary behaviours; the cumulative effect of this may explain some of the variability, but it is unlikely to be responsible for the large differences observed in key outcomes when compared to the age-matched only cohort. Where this study builds upon previous findings is through the use of clustering; when the study clusters were created, it became apparent that they were diametrically opposed in terms of health status. Cluster One had more favourable values for all outcomes, whereas Cluster Two was opposing. It was also apparent that the majority of participants from the non-PCOS cohort were assigned to the cluster with the healthier profile, but women with PCOS were split almost evenly between the two, supporting the

notion that a PCOS diagnosis further disrupts the health profile of a large proportion of women with PCOS.

Regardless of PCOS diagnosis, additional factors defined the two clusters. Cluster One were more physically active and spent less time engaged in sedentary behaviours than those in Cluster Two. Whilst it is feasible that some individuals may not engage in physical activity due to ill health (Macniven *et al.*, 2014), there were no statistical differences in the number of morbidities (excluding PCOS) between the two clusters. This cluster pattern indicates that being physically active and spending less time sitting are important variables in the preservation of health in these women. When women with PCOS are considered individually, these findings support current treatment recommendations for lifestyle improvements, with the women with PCOS who have the most favourable health measures being also the ones who are more physically active.

Overall, considering both the findings and the limitations of the current study, a need for further projects in women with PCOS is also highlighted. The UK Biobank is a valuable resource for population level enquiry, but it may not be the most suitable resource for disease specific enquiry, particularly for certain conditions such as PCOS. Indeed, studies such as this one, offer further insight which generate subsequent research ideas and generate informed hypotheses, but are also limited in their ability to make definitive statements/conclusions. As with previous chapters, this again reiterates the need for robustly designed, well-conducted studies in this female patient population.

Chapter 6: Conclusions and Future work

The work presented in this thesis focused on polycystic ovary syndrome (PCOS) because it represents the most common endocrinopathy in reproductive aged women which, depending upon the applied diagnostic criteria and the sampled population, affects between 2% (Chen *et al.*, 2008) and 21% (Lizneva *et al.* 2016) of reproductive-aged women, and therefore, poses an increasing problem to healthcare systems globally. The PCOS diagnostic criteria include a combination of three principle features (hyperandrogenism, menstrual irregularity and polycystic ovary morphology) (Azziz *et al.* 2009), which frequently manifest in a range of symptoms, such as hirsutism, acne, alopecia and irregularity of periods. Adding to the complexity of PCOS, there is also a range of additional health problems that are widely associated with PCOS. These include obesity and insulin resistance (leading to increased prevalence of type 2 diabetes), psychological issues (namely anxiety and depression), which, in turn, have a further negative effect upon health-related quality of life (HRQoL) (Diamanti-Kandarakis and Panidis, 2007; Moran and Teede, 2009; Barry *et al.*, 2014).

Increasing physical activity (PA) has been demonstrated in numerous populations to be an effective strategy for lowering body weight (Look *et al*., 2013), restoring insulin sensitivity (Conn *et al*., 2014), lowering cardiovascular disease (CVD) risk (Pattyn, *et al.*, 2013) and at reducing levels of psychological morbidity (Cooney *et al*., 2013). This evidence indicates that PA may be a suitable treatment strategy for improving the overall health of women with PCOS. Accordingly, the first line treatment recommendations for these women focus upon lifestyle changes, incorporating caloric restriction through dietary regulation alongside increases in the volume of PA/exercise performed (Legro *et al.*, 2013). Although the high prevalence of PCOS is widely recognised, and there is a general consensus regarding the purported benefits of PA and the relevant recommendations for women with PCOS, there is still a degree of uncertainty based on the available evidence as to its effectiveness in this female patient population. Indeed, previous systematic reviews (Harrison *et al.*, 2011; Moran *et al*., 2011; Domecq *et al*., 2013; Haqq *et al*., 2015, Benham *et al.*, 2018) have not been able to determine the beneficial effects of PA, in a range of outcomes, for women with PCOS. In addition, many reviews have overlooked the dietary components of lifestyle interventions and have included only a small number of studies that typically incorporate a low number of participants.

Alongside this, some previous reviews have included non-randomised, uncontrolled or cohort studies and the authors' methods for evaluating the quality of included studies, and the degree of confidence that can be assumed from their findings, may not be deemed as robust (Higgins and Green, 2009).

In this context, the systematic review and meta-analysis presented in Chapter 3 of this thesis aimed to address many of the issues in previous reviews and provide a current synthesis of all available evidence using the highest level of methodological rigour. It also went further than previous systematic reviews in that, where evidence was available, it included multiple permutations of 'lifestyle' as comparators; that is exercise, diet, and exercise combined with diet. It is also currently the only systematic review on the topic to use the GRADE assessment tool (Guyatt *et al.,* 2008) for evaluating the certainty of the evidence, something that is recommended by the Cochrane Collaboration (Schünemann *et al.,* 2011).

The work presented in Chapter 3 identified many statistically significant benefits of the effects of exercise interventions on individual outcomes which tend to agree with the findings of one (or more) of the preceding systematic reviews. As such, when exercise was compared to no intervention in the present systematic review, statistical effects favouring exercise were observed for only four outcomes (fasting insulin, total cholesterol, LDL-C, and VO_2 max) when both change from baseline and post-intervention comparisons were made. Similar findings were reported for fasting insulin (Moran et al., 2011; Domecq et al., 2013; Benham *et al.,* 2018), total cholesterol (Benham *et al.*, 2018), LDL-C (Benham *et al.*, 2018) and VO_2 max (Haqq *et al.,* 2015) in previous reviews, which collectively suggests that exercise may have a beneficial effect on women with PCOS. However, the magnitude of these improvements must be noted. For these four outcomes named above, the mean difference was relatively small, and when considering the width of the 95% confidence intervals (CIs), there was evidence of a lack of precision (*i.e.,* wide CIs) and differences that were close to zero were included within their bounds. Alongside this, studies included in individual analyses typically had included results from a small number of participants and a high or unclear risk of bias across multiple domains. Thus, the quality of the related evidence was downgraded to either low, or very low, indicating that true effect estimates are likely to be substantially different from those reported

(Hultcrantz *et al.*, 2017). Moreover, when comparisons including diet, or diet and exercise were made, the findings had even more uncertainty surrounding them. Of note, although exercise and dietary control are key constituents of current recommendations for the management of PCOS, very few studies have applied these as investigative interventions in randomised controlled trials (RCTs). Indeed, only six eligible trials incorporated any type of dietary intervention and the variability between the included outcomes was great, making a meta-analysis difficult. Where meta-analysis was completed, results were null or negligible and the certainty of the evidence was very low.

Building on the findings from the presented systematic review and although it appears that exercise interventions may induce favourable changes for some outcomes in women with PCOS, the evidence presented in Chapter 3 suggests that the magnitude of these changes is relatively small. In fact, it is difficult to state whether changes reported are clinically meaningful because, for many included outcomes, there is ambiguity around what constitutes a clinically meaningful change in women with PCOS (and often in the general population). This poses the question about whether is it appropriate to recommend exercise as a treatment for these women even where there are statistical improvements reported, since there is uncertainty surrounding the degree of importance that this change represents, and the body of evidence upon which it is based is graded as low, or very low quality.

In view of the question posed above, it was clearly apparent from completing the systematic review presented in Chapter 3 that there is a paucity of studies which have adequately investigated the impact of lifestyle interventions in women with PCOS. The latter requires having sufficient statistical power (*i.e.,* larger sample sizes), well conducted and reported studies (*i.e.,* evidence that risk of bias has been reduced) that utilise appropriate interventions (*i.e.,* exercise, diet and behaviour change) with appropriate follow-up assessments that assess long-term patient benefits. Given that PCOS is so prevalent in a young patient population, conducting these studies should be a priority. It is in fact the quality of the existing evidence highlighted in Chapter 3, rather than the observed statistical improvements, which highlights the need for such studies to be completed. Accordingly, in response to the findings of the systematic review, the objective of this PhD project was to also design and implement a randomised feasibility study which would compare an intervention incorporating

exercise and diet, underpinned by behaviour change techniques (BCTs), to a diet and usual care control group.

Unfortunately, it was not feasible to conduct this planned randomised controlled trial despite a comprehensive protocol design. In fact, the National Health Service (NHS) Health Research Authority and Health and Care Research Wales (HCRW) ethics committee (18/WM/0123) granted approval for this planned study in 2018. Despite ethical approval and extensive preparation, it was subsequently not possible to progress to the recruitment phase within the time and budget constraints of a PhD project due to logistical and administrative complications. Sufficiently sized RCTs are the main staple of determining the safety and efficacy of treatments and are principle in global health improvements (Collins and MacMahon, 2001). Whilst there is a need to ensure the safety of participants, increasingly stringent regulation and bureaucracy make the commencement and conduct of trials more complicated (Califf, 2006) and whilst such policy may be well-intentioned, it can cause unprecedented and irreversible delays to important patient-centred research (Reith *et al.,* 2013). One such example surrounds the requirement to obtain approval from multiple organisations before such a study can commence. Even when study regulation and ethics are centralised, permission is required from each study site (*e.g.,* universities, hospitals, *etc.*) which in many instances, leads to duplicated efforts and unnecessary administration (Duley *et al.,* 2008). Indeed, the presence of such hurdles in the current project were unassailable given the time constraints of a typical PhD, and this may be also viewed as a finding confirmed by the work conducted in the context of this PhD. Nevertheless, and despite setbacks in conducting such well-designed RCTs, the research recommendations should remain unchanged and objective, and unbiased investigations into the effectiveness of lifestyle interventions in this under-studied female patient population should be a research priority.

Although not deemed the gold standard of scientific investigation (Abel and Koch, 1999), well-designed case-control studies and observational studies utilising concurrent selection of control subjects can produce useful results and help to support/shape patient care, clinical practice guidelines and the education of healthcare professionals (Concato *et al.,* 2000). Therefore, without the incorporation of an exercise intervention, the remainder of this PhD thesis addressed some key

questions surrounding the effects of PA (as opposed to structured exercise) as a desirable health behaviour in women with PCOS and looked to seek comparisons with women without PCOS.

Accordingly, the case-control study presented in Chapter 4 recruited women with PCOS and healthy controls and found that those with PCOS had substantially worse body weight, greater waist circumference, poorer self-esteem and reduced HRQoL than the comparison group. Furthermore, it revealed that there were no statistical differences in the volume of self-reported weekly PA, or indeed time spent in sedentary behaviours. Somewhat surprisingly, when mediation analysis was completed, PA appeared to have no influence upon HRQoL in either study group, which contradicts commonly held beliefs. This suggests that the PA behaviours of these women are not responsible for these undesirable differences, and that other mechanisms, most likely a diagnosis of PCOS, are implicated. In fact, both a positive PCOS diagnosis and higher self-reported BMI were associated with poorer self-esteem and lower HRQoL. However, it would be problematic to conclude that PCOS is solely responsible for poorer health in these women, because self-reported BMI explains a similar proportion of the variance. Thus, it is more likely is that the diagnostic symptoms/manifestations of PCOS result in considerable physical and psychological burden, whilst the metabolic dysregulation also promotes weight gain leading to decreased self-esteem. This triple-threat effect is then responsible, at least in part, for the impaired HRQoL observed in the study participants.

To some extent, this is supported by the results of the study presented in Chapter 5 where, as expected, more women with PCOS had overweightness/obesity than an age-matched control group. Again, it is likely that the consequences of increased adiposity (*i.e.*, systemic inflammation and insulin resistance), alongside the PCOS diagnosis, is responsible for the less-favourable health measures that were found in this study. In addition, although the UK Biobank do not use a previously validated measure of HRQoL (such as the SF-12), a greater proportion of women without PCOS rated their health as "*excellent"* or "*good"* compared to women with PCOS (76% *versus* 52%, respectively). Chapter 5 also utilised a BMI- and age-matched control group with almost two thirds of this group (66%) self-rating their health as "*excellent"* or "*good".* This is a reasonable indicator suggesting that increased BMI may impair health, but that health may be further negatively affected

when these women have PCOS. Further evidence for this can be gained when circulating biomarkers and physiological characteristics are also considered since, without exception, women with PCOS in this study had less favourable values than the age-matched group. Moreover, the former also had less favourable values for four key CVD risk factors (HDL-C, triglycerides, HbA1c and SHBG) when compared to the BMI-matched cohort. The culmination of these less favourable outcomes for women with PCOS in this study may be associated with the increased number of comorbidities (total and metabolic-related, namely type 2 diabetes, hypertension and hypercholesterolemia) when compared to either comparator group.

It is also noteworthy that, whilst higher PA has been shown to independently improve cardiometabolic risk factors and to reduce CVD mortality (Warburton, Nicol and Bredin, 2006; Leitzmann *et al.,* 2007), the evidence from the current PhD, albeit using CVD risk factors as surrogate measures of mortality, suggests PA does not have an equally protective role in women with PCOS. Indeed, many previous studies report no statistical differences between women with PCOS and healthy comparators for PA (Moran *et al.,* 2013; Lin *et al.,* 2019; Douglas *et al.,* 2006; Wright *et al.,* 2004; Alvarez-Blasco *et al.*, 2011; Cutler, Pride and Cheung, 2018). Regarding this point, where the work presented in Chapter 4 agreed with these findings, the study detailed in Chapter 5 did uncover some statistical variation. As such, in the latter, women with PCOS performed less vigorous PA than both comparator groups and spent more time engaged in sedentary behaviours than the age-matched group. It is unlikely that the magnitude of difference for vigorous PA can explain the poorer health in women with PCOS. Indeed, in a female population, it has previously been reported, that where energy expenditure was matched, regardless of intensity (walking compared to vigorous PA), the magnitude of type 2 diabetes risk reduction was comparable (Hu *et al.,* 1999).

What the results presented in Chapter 5 allowed, was a classification of participants based upon their objectively measured health (clustering) and between group PA and sedentary behaviour comparison. This cluster analysis revealed two clusters, one where the values of the measures indicated a healthier profile (Cluster one) and one with less favourable values indicating a less healthy profile (Cluster two). A defining feature of these clusters was the PA behaviours, with

participants in cluster one reporting higher PA levels and less sitting time than those in cluster two. When classified on PA and sedentary time centiles, women who were in the upper centile for PA and the lower centile for sedentary time (low risk group/category) had more favourable values in all outcomes, and lower instances of morbidity compared to those in the lower PA and higher sedentary centiles (high risk group/category). Importantly, women with PCOS made up a greater proportion of Cluster two than Cluster one, and despite being represented by half as many study participants, also had greater membership to Cluster two than the age-matched study group. Women with PCOS in cluster one completed more PA and spent less time sitting than those in cluster two, suggesting that there is a health protective effect in these women. This is strongly supported by the centile analysis where the health of these women gradually decreases as they move from low, to moderate, to high risk based upon their PA behaviours.

Overall, the implications of these findings are that higher PA and reduced sedentary time appear to be associated with a more favourable health profile in the general female population, but also specifically in women with PCOS. The effectiveness of structured exercise interventions at improving health in PCOS is uncertain and there remains very limited information about the long-term impact of such interventions. However, there is evidence, as also indicated by the present work, that those women with PCOS who are more physically active and spend less time sitting have better health which reinforces its suitability as a treatment recommendation by healthcare professionals. However, it is possible that the healthier women with PCOS are more active *because* they have better health; eliminating reverse causation such as this is difficult to do using observational study designs and further reinforces the need for robust interventional trials.

In this context, an additional problem lies in how to support lifestyle changes in women with PCOS, who are often sedentary, in order to engage in more PA and maintain favourable PA behaviours long-term. This represents an additional area that needs well-designed large-scale research studies in this understudied female patient population. These should focus upon delivering effective lifestyle interventions, which incorporate diet and exercise, but should also place increased emphasis on the

development of targeted behaviour change interventions and evaluation of their effectiveness during long-term follow-up.

References.

Abel, U., & Koch, A. (1999). The role of randomization in clinical studies: myths and beliefs. *Journal of Clinical Epidemiology*, *52*(6), 487-497.

Abraham, C., Conner, M. Jones, F., and O'Connor, D, (2016) *Health Psychology* (2nd ed.). Oxon, UK: Routledge.

Açmaz, G., Albayrak, E., Acmaz, B., Başer, M., Soyak, M., Zararsız, G. and İpekMüderris, İ., (2013). Level of anxiety, depression, self-esteem, social anxiety, and quality of life among the women with polycystic ovary syndrome. *The Scientific World Journal*, *2013*, 1-7.

Adams, J., Polson, D. W., Abdulwahid, N., Morris, D. V., Franks, S., Mason, H. D., ... & Jacobs, H. S. (1985). Multifollicular ovaries: clinical and endocrine features and response to pulsatile gonadotropin releasing hormone. *The Lancet*, *326*(8469-8470), 1375-1379.

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, *16*(2), 270-301.

Ahmadi, A., Akbarzadeh, M., Mohammadi, F., Akbari, M., Jafari, B., & Tolide-Ie, H. R. (2013). Anthropometric characteristics and dietary pattern of women with polycystic ovary syndrome. *Indian Journal of Endocrinology and Metabolism*, *17*(4), 672.

Ainsworth, B. E., Haskell, W. L., Herrmann, S. D., Meckes, N., Bassett Jr, D. R., Tudor-Locke, C., ... & Leon, A. S. (2011). Compendium of Physical Activities: a second update of codes and MET values. *Medicine & Science in Sports & Exercise*, *43*(8), 1575-1581.

Ainsworth, B. E., Haskell, W. L., Leon, A. S., Jacobs Jr, D. R., Montoye, H. J., Sallis, J. F., & Paffenbarger Jr, R. S. (1993). Compendium of physical activities: classification of energy costs of human physical activities. *Medicine and Science in Sports and Exercise*, *25*(1), 71-80.

Ainsworth, B. E., Haskell, W. L., Whitt, M. C., Irwin, M. L., Swartz, A. M., Strath, S. J., ... & Jacobs, D. R. (2000). Compendium of physical activities: an update of activity codes and MET intensities. *Medicine and Science in Sports and Exercise*, *32*(9; SUPP/1), S498-S504.

Albers, C. & Lakens, D. (2017). Biased sample size estimates in a-priori power analysis due to the choice of the effect size index and follow-up bias. *Journal of Experimental Social Psychology*.

Alberti, K. G. M. M., & Zimmet, P. Z. (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.

Alberti, K. G. M. M., Zimmet, P., & Shaw, J. (2006). Metabolic syndrome—a new world-wide definition. A consensus statement from the international diabetes federation. *Diabetic Medicine*, *23*(5), 469-480.

Allen, N. E., Appleby, P. N., Davey, G. K., & Key, T. J. (2002). Lifestyle and nutritional determinants of bioavailable androgens and related hormones in British men. *Cancer Causes & Control*, *13*(4), 353-363.

Almenning, I., Rieber-mohn, A., Garnaes, K. K., Lundgren, K. M., Lovik, T.S., Moholdt, T.T. (2015) High intensity interval training versus strength training to improve cardiovascular risk factors in women with polycystic ovary syndrome. A randomized controlled trial. *European Journal of Preventative Cardiology*. *22*, S23.

Almenning, I., Rieber-Mohn, A., Lundgren, K. M., Løvvik, T. S., Garnæs, K. K., & Moholdt, T. (2015). Effects of high intensity interval training and strength training on metabolic, cardiovascular and hormonal outcomes in women with polycystic ovary syndrome: a pilot study. *PLoS One*, *10*(9), e0138793.

Almoosawi, S., Winter, J., Prynne, C. J., Hardy, R., & Stephen, A. M. (2012). Daily profiles of energy and nutrient intakes: are eating profiles changing over time? *European Journal of Clinical Nutrition*, *66*(6), 678-686.

Altuntas, Y., Bilir, M., Ucak, S., & Gundogdu, S. (2005). Reactive hypoglycemia in lean young women with PCOS and correlations with insulin sensitivity and with beta cell function. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, *119*(2), 198-205.

Alvarez-Blasco, F., Botella-Carretero, J. I., San Millán, J. L., & 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., Luque-Ramírez, M., & Escobar-Morreale, H. F. (2011). Diet composition and physical activity in overweight and obese premenopausal women with or without polycystic ovary syndrome. *Gynecological Endocrinology*, *27*(12), 978-981.

Anandacoomarasamy, A., Caterson, I. D., Leibman, S., Smith, G. S., Sambrook, P. N., Fransen, M., & March, L. M. (2009). Influence of BMI on health-related quality of life: Comparison between an obese adult cohort and age-matched population norms. *Obesity*, *17*(11), 2114-2118.

Anderberg, M. R. (2014). *Cluster analysis for applications: probability and mathematical statistics: a series of monographs and textbooks* (Vol. 19). Academic press.

Annagür, B. B., Tazegül, A., & Akbaba, N. (2014). Body image, self-esteem and depressive symptomatology in women with polycystic ovary syndrome. *Nöro Psikiyatri Arşivi*, *51*(2), 129.

Ansari, J. N., Banyameen, I., Nikam, S. R., & Siddique, A. K. (2017). A case control study to asceetain the relationship of metabolic, biochemical and hormonal changes in woman suffering from polycystic ovary syndrome. *European Journal of Pharmaceutical and Medical Research*, *4*(6), 603-608.

Ansarin, H., Azizi, J. M., Rasi, A., & Soltani, A. R. (2007). Clinical presentation and etiologic factors of hirsutism in premenopausal Iranian women. *Archives of Iranian Medicine*, *10*(1), 7-13.

Arem, H., Moore, S. C., Patel, A., Hartge, P., De Gonzalez, A. B., Visvanathan, K., ... & Linet, M. S. (2015). Leisure time physical activity and mortality: a detailed pooled analysis of the dose-response relationship. *JAMA Internal Medicine*, *175*(6), 959-967.

Asante, A., Konopka, A. R., Stewart, E. A., Coddington, C. C., & Nair, S. K. (2014). Impact of aerobic exercise on mitochondrial function and insulin sensitivity in women with polycystic ovary syndrome-a randomized controlled trial. *Reproductive Sciences*, *21*(3), 126A.

Ascaso, J. F., Pardo, S., Real, J. T., Lorente, R. I., Priego, A., & Carmena, R. (2003). Diagnosing insulin resistance by simple quantitative methods in subjects with normal glucose metabolism. *Diabetes Care*, *26*(12), 3320-3325.

Avery, J. C., & Braunack-Mayer, A. J. (2007). The information needs of women diagnosed with Polycystic Ovarian Syndrome–implications for treatment and health outcomes. *BMC Women's Health*, *7*(1), 9.

Azziz, R., Carmina, E., Chen, Z., Dunaif, A., Laven, J. S., Legro, R. S., ... & Yildiz, B. O. (2016). Polycystic ovary syndrome. *Nature Reviews Disease Primers*, *2*(1), 1-18.

Azziz, R., Carmina, E., Dewailly, D., Diamanti-Kandarakis, E., Escobar-Morreale, H. F., Futterweit, W., ... & Witchel, S. F. (2009). The Androgen Excess and PCOS Society criteria for the polycystic ovary syndrome: the complete task force report. *Fertility and Sterility*, *91*(2), 456-488.

Azziz, R., Carmina, E., Dewailly, D., Diamanti-Kandarakis, E., Escobar-Morreale, H. F., Futterweit, W., ... & Witchel, S. F. (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.

Bacchi, E., Negri, C., Di Sarra, D., Tosi, F., Tarperi, C., Moretta, R., ... & Moghetti, P. (2015). Serum testosterone predicts cardiorespiratory fitness impairment in normal-weight women with polycystic ovary syndrome. *Clinical Endocrinology*, *83*(6), 895-901.

Bachani, S., Malla, V. G., & Singh, P. (2012). M110 to study the efficacy of intervention strategies on features of metabolic syndrome in Asian urban adolescent girls with PCOS. *International Journal of Gynecology and Obstetrics*, (119), S567.

Bala, M., Meenakshi, M. K., & Gupta, A. (2017). Correlation of HbA1C levels with body mass index in newly diagnosed polycystic ovary syndrome. *Electronic Journal of the International Federation of Clinical Chemistry and Laboratory*, *28*(3), 196.

Balen, A. H., Conway, G. S., Kaltsas, G., Techatraisak, K., Manning, P. J., West, C., & Jacobs, H. S. (1995). Andrology: Polycystic ovary syndrome: the spectrum of the disorder in 1741 patients. *Human Reproduction*, *10*(8), 2107-2111.

Balen, A. H., Laven, J. S., Tan, S. L., & Dewailly, D. (2003). Ultrasound assessment of the polycystic ovary: international consensus definitions. *Human Reproduction Update*, *9*(6), 505-514.

Balen, A. H., Morley, L. C., Misso, M., Franks, S., Legro, R. S., Wijeyaratne, C. N., ... & Teede, H. (2016). The management of anovulatory infertility in women with polycystic ovary syndrome: an analysis of the evidence to support the development of global WHO guidance. *Human Reproduction Update*, *22*(6), 687-708.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191.

Bandura, A. (1986). *Social foundations of thought and action.* Englewood Cliffs, NJ, 1986.

Banting, L. K., Gibson-Helm, M., Polman, R., Teede, H. J., & Stepto, N. K. (2014). Physical activity and mental health in women with polycystic ovary syndrome. *BMC Women's Health*, *14*(1), 51.

Baranowski, T. (1988). Validity and reliability of self-report measures of physical activity: an information-processing perspective. *Research Quarterly for Exercise and Sport*, *59*(4), 314-327.

Barber, T. M., McCarthy, M. I., Wass, J. A. H., & Franks, S. (2006). Obesity and polycystic ovary syndrome. *Clinical Endocrinology*, *65*(2), 137-145.

Barnard, L., Ferriday, D., Guenther, N., Strauss, B., Balen, A.H. and Dye, L., (2007). Quality of life and psychological well-being in polycystic ovary syndrome. *Human Reproduction*, *22*(8), 2279-2286.

Barr, S., Hart, K., Reeves, S., Sharp, K., Jeanes, Y.M, (2011). Habitual dietary intake, eating pattern and physical activity of women with polycystic ovary syndrome. *European Journal of Clinical Nutrition*, *65*(10), 1126-32.

Barry, J. A. (2019). *Psychological Aspects of Polycystic Ovary Syndrome*. Switzerland: Springer Nature.

Barry, J. A., Azizia, M. M., & 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., Bouloux, P., & Hardiman, P. J. (2011a). The impact of eating behavior on psychological symptoms typical of reactive hypoglycemia. A pilot study comparing women with polycystic ovary syndrome to controls. *Appetite*, *57*(1), 73-76.

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

Barry, J. A., Qu, F., & Hardiman, P. J. (2018). An exploration of the hypothesis that testosterone is implicated in the psychological functioning of women with polycystic ovary syndrome (PCOS). *Medical Hypotheses*, *110*, 42-45.

Bass, K. M., Newschaffer, C. J., Klag, M. J., & Bush, T. L. (1993). Plasma lipoprotein levels as predictors of cardiovascular death in women. *Archives of Internal Medicine*, *153*(19), 2209-2216.

Batool, S., Ul Ain Ahmed, F., Ambreen, A., Sheikh, A., & Faryad, N. (2016). Depression and anxiety in women with polycystic ovary syndrome and its biochemical associates. *Journal of South Asian Federation of Obstetrics and Gynecology*, *8*, 44-47.

Bazarganipour, F., Taghavi, S.A., Montazeri, A., Ahmadi, F., Chaman, R. and Khosravi, A., (2015). The impact of polycystic ovary syndrome on the health-related quality of life: A systematic review and meta-analysis. *Iranian Journal of Reproductive Medicine*, *13*(2), 61.

Bazarganipour, F., Ziaei, S., Montazeri, A., Foroozanfard, F., Kazemnejad, A. and Faghihzadeh, S., (2013). Predictive factors of health-related quality of life in patients with polycystic ovary syndrome: a structural equation modeling approach. *Fertility and Sterility*, *100*(5), pp.1389-1396.

Bazarganipour, F., Ziaei, S., Montazeri, A., Foroozanfard, F., Kazemnejad, A., & Faghihzadeh, S. (2013). Body image satisfaction and self-esteem status among the patients with polycystic ovary syndrome. *Iranian Journal of Reproductive Medicine*, *11*(10), 829.

Bazarganipour, F., Ziaei, S., Montazeri, A., Foroozanfard, F., Kazemnejad, A. and Faghihzadeh, S. (2014). Health-related quality of life in patients with polycystic ovary syndrome (PCOS): a model-based study of predictive factors. *The Journal of Sexual Medicine*, *11*(4), 1023-1032.

Becerra, M. O., Muros, J. J., Cuadros, J. P., Sánchez, J. M., & González, M. C. (2015). Influence of BMI on self-esteem of children aged 12-14 years. *Anales de Pediatría (English Edition)*, *83*(5), 311-317.

Becker, M. H., & Maiman, L. A. (1975). Sociobehavioral determinants of compliance with health and medical care recommendations. *Medical Care*, 10-24.

Benham, J. L., Yamamoto, J. M., Friedenreich, C. M., Rabi, D. M., & Sigal, R. J. (2018). Role of exercise training in polycystic ovary syndrome: a systematic review and meta-analysis. *Clinical Obesity*, *8*(4), 275-284.

Benham, J., Yamamoto, J., Sigal, R., Corenblum, B., Friedenreich, C. and Rabi, D. The role of exercise training in polycystic ovary syndrome: a systematic review. PROSPERO 2017: CRD42017065324 Available from http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42017065324

Benson, S., Arck, P. C., Tan, S., Hahn, S., Mann, K., Rifaie, N., ... & Elsenbruch, S. (2009). Disturbed stress responses in women with polycystic ovary syndrome. *Psychoneuroendocrinology*, *34*(5), 727-735.

Berger, B., Weinberg, R., and Eklund, R. (2015). *Exercise Psychology.* Morgantown, WV: Fitness Information Technology.

Bigaard, J., Spanggaard, I., Thomsen, B. L., Overvad, K., & Tjønneland, A. (2005). Self-reported and technician-measured waist circumferences differ in middle-aged men and women. *The Journal of Nutrition*, *135*(9), 2263-2270.

Biro, F. M., Striegel-Moore, R. H., Franko, D. L., Padgett, J., & Bean, J. A. (2006). Self-esteem in adolescent females. *Journal of Adolescent Health*, *39*(4), 501-507.

Biswas, A., Oh, P. I., Faulkner, G. E., Bajaj, R. R., Silver, M. A., Mitchell, M. S., & Alter, D. A. (2015). Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis. *Annals of Internal Medicine*, *162*(2), 123-132.

Blair, S. N., Kohl, H. W., Paffenbarger, R. S., Clark, D. G., Cooper, K. H., & Gibbons, L. W. (1989). Physical fitness and all-cause mortality: a prospective study of healthy men and women. *Jama*, *262*(17), 2395-2401.

Bland, C. (2015) *An Introduction to Medical Statistics* (4th ed). Oxford, UK: Oxford University Press.

Bongaard, B. (2010). Fact or fiction? Is a low glycemic index diet the best choice for your PCOS patients? *Alternative Medicine Alert*, *13*(8), 91-93.

Bonheur, B., & Young, S. W. (1991). Exercise as a health-promoting lifestyle choice. *Applied Nursing Research*, *4*(1), 2-6.

Borbón-Castro, N. A., Castro-Zamora, A. A., Cruz-Castruita, R. O. S. A., Takahiro, S., Banda-Sauceda, N. C., & De La Cruz-Ortega, M. F. (2019). The effects of a multidimensional exercise program on health behavior and biopsychological factors in Mexican older adults. *Frontiers in Psychology*, *10*, 2668.

Borghi, L., Leone, D., Vegni, E., Galiano, V., Lepadatu, C., Sulpizio, P., & Garzia, E. (2018). Psychological distress, anger and quality of life in polycystic ovary syndrome: associations with biochemical, phenotypical and socio-demographic factors. *Journal of Psychosomatic Obstetrics & Gynecology*, *39*(2), 128-137.

Boulé, N. G., Haddad, E., Kenny, G. P., Wells, G. A., & Sigal, R. J. (2001). Effects of exercise on glycemic control and body mass in type 2 diabetes mellitus: a meta-analysis of controlled clinical trials. *Jama*, *286*(10), 1218-1227.

Boyle, J. A., Cunningham, J., O'Dea, K., Dunbar, T., & Norman, R. J. (2012). Prevalence of polycystic ovary syndrome in a sample of Indigenous women in Darwin, Australia. *Medical Journal of Australia*, *196*(1), 62-66.

Boyns, D. R., Crossley, J. N., Abrams, M. E., Jarrett, R. J., & Keen, H. (1969). Oral glucose tolerance and related factors in a normal population sample—I. Blood sugar, plasma insulin, glyceride, and cholesterol measurements and the effects of age and sex. *British Medical Journal*, *1*(5644), 595-598.

Bozdag, G., Mumusoglu, S., Zengin, D., Karabulut, E., & Yildiz, B. O. (2016). The prevalence and phenotypic features of polycystic ovary syndrome: a systematic review and meta-analysis. *Human Reproduction*, *31*(12), 2841-2855.

Briesch, R, & Rajagopal, P. (2010). Neural network applications in consumer behavior. *Journal of Consumer Psychology*, 20, 381‑389

Brook, C., Marshall, N. (1996). The Endocrine System, In: *Essential Endocrinology* (3rd ed). Oxford, UK: Blackwell Science Ltd. pp. 1-13.

Brown, A. J. (2005). Effect of exercise without weight loss on insulin resistance in women with polycystic ovary syndrome: a randomized, controlled study. In *3rd Annual Meeting of the Androgen Excess Society San Diego, California. June 2005, 9 28*.

Brown, A. J., Setji, T. L., Sanders, L. L., Lowry, K. P., Otvos, J. D., Kraus, W. E., & Svetkey, L. P. (2009). Effects of exercise on lipoprotein particles in women with polycystic ovary syndrome. *Medicine and Science in Sports and Exercise*, *41*(3), 497.

Brueton, V., Tierney, J., Stenning, S., Nazareth, I., Meredith, S., Harding, S., & Rait, G. (2011). Strategies to reduce attrition in randomised trials. *Trials*, *12*(S1), A128.

Bruner, B., Chad, K., & Chizen, D. (2006). Effects of exercise and nutritional counseling in women with polycystic ovary syndrome. *Applied Physiology, Nutrition, and Metabolism*, *31*(4), 384-391.

Burger, H. G. (2002). Androgen production in women. *Fertility and Sterility*, *77*, 3-5.

Califf, R. M. (2006). Clinical trials bureaucracy: unintended consequences of well-intentioned policy. *Clinical Trials*, *3*(6), 496-502.

Campo, S., Felli, A., Lamanna, M. A., Barini, A., & Garcea, N. (1993). Endocrine changes and clinical outcome after laparoscopic ovarian resection in women with polycystic ovaries. *Human Reproduction*, *8*(3), 359-363.

Carmina, E., Bucchieri, S., Mansueto, P., Rini, G., Ferin, M., & Lobo, R. A. (2009). Circulating levels of adipose products and differences in fat distribution in the ovulatory and anovulatory phenotypes of polycystic ovary syndrome. *Fertility and Sterility*, *91*(4), 1332-1335.

Carmina, E., Chu, M. C., Longo, R. A., Rini, G. B., & Lobo, R. A. (2005). Phenotypic variation in hyperandrogenic women influences the findings of abnormal metabolic and cardiovascular risk parameters. *The Journal of Clinical Endocrinology & Metabolism*, *90*(5), 2545-2549.

Carmina, E., Koyama, T., Chang, L., Stanczyk, F. Z., & Lobo, R. A. (1992). Does ethnicity influence the prevalence of adrenal hyperandrogenism and insulin resistance in polycystic ovary syndrome? *American Journal of Obstetrics and Gynecology*, *167*(6), 1807-1812.

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health*, *100*(2), 126-131.

Chen, X., Yang, D., Mo, Y., Li, L., Chen, Y., & 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.

Chevenne, D., Trivin, F., & Porquet, D. (1999). Insulin assays and reference values. *Diabetes & Metabolism*, *25*(6), 459-476.

Ching, H.L., Burke, V. and Stuckey, B.G.A., (2007). Quality of life and psychological morbidity in women with polycystic ovary syndrome: body mass index, age and the provision of patient information are significant modifiers. *Clinical Endocrinology*, 66(3), 373-379.

Chizen, D. R., Serrao, S., Rooke, J., McBreairty, L., Pierson, R. A., Chilibeck, P., Zello, G. (2014). The "pulse" diet & PCOS. *Fertility and Sterility*. *102*(3): e267

Christiansen, S. C., Vanky, E., Klungland, H., Stafne, S. N., Mørkved, S., Salvesen, K. Å., ... & Carlsen, S. M. (2014). The effect of exercise and metformin treatment on circulating free DNA in pregnancy. *Placenta*, *35*(12), 989-993.

Chung-Yan Chan, J. (2014). Psychological determinants of exercise behavior of nursing students. *Contemporary Nurse*, *49*(1), 60-67.

Cibula, D., Kuzel, D., Rezabek, K., Vrbikova, J., Zivny, J., & Starka, P. L. (2000). Multiple ovarian biopsy in the treatment of women with polycystic ovary syndrome (PCOS). *Clinical and Experimental Obstetrics & Gynecology*, *27*(3-4), 194-196.

Clark, A. M., Ledger, W., Galletly, C., Tomlinson, L., Blaney, F., Wang, X., & Norman, R. J. (1995). Weight loss results in significant improvement in pregnancy and ovulation rates in anovulatory obese women. *Human Reproduction*, *10*(10), 2705-2712.

Clark, A. M., Thornley, B., Tomlinson, L., Galletley, C., & Norman, R. J. (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.

Clark, B. J., Soo, S. C., Caron, K. M., Ikeda, Y., Parker, K. L., & Stocco, D. M. (1995). Hormonal and developmental regulation of the steroidogenic acute regulatory protein. *Molecular Endocrinology*, *9*(10), 1346-1355.

Clark, T. G., & Altman, D. G. (2003). Developing a prognostic model in the presence of missing data: an ovarian cancer case study. *Journal of Clinical Epidemiology*, *56*(1), 28-37.

Coffey, S., Bano, G., & 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.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (second ed.). Lawrence Erlbaum Associates.

Collins J. (2016). Your Hormones; everything affects your hormones – Steroidogenic Pathway. [online] Available at: https://www.yourhormones.com/everything-affects-your-hormones/. Accessed November 6th 2020.

Collins, R., & MacMahon, S. (2001). Reliable assessment of the effects of treatment on mortality and major morbidity, I: clinical trials. *The Lancet*, *357*(9253), 373-380.

Columbia Public Health (2020). *Cluster Analysis Using K-means* [online]. Available at: https://www.publichealth.columbia.edu/research/population-health-methods/cluster-analysis-using-k-means [last accessed 12th August 2020].

Colwell, K., Lujan, M. E., Lawson, K. L., Pierson, R. A., & Chizen, D. R. (2010). Women's perceptions of polycystic ovary syndrome following participation in a clinical research study: implications for knowledge, feelings, and daily health practices. *Journal of Obstetrics and Gynaecology Canada*, *32*(5), 453-459.

Concato, J., Shah, N., & Horwitz, R. I. (2000). Randomized, controlled trials, observational studies, and the hierarchy of research designs. *New England Journal of Medicine*, *342*(25), 1887-1892.

Conn, V. S., Koopman, R. J., Ruppar, T. M., Phillips, L. J., Mehr, D. R., & Hafdahl, A. R. (2014). Insulin Sensitivity Following Exercise Interventions: Systematic Review and Meta-Analysis of Outcomes Among Healthy Adults. *Journal of Primary Care and Community Health*, 5(3), 211-222.

Consensus Panel, N. I. H. (1993). Triglyceride, HDL and coronary heart disease. *Journal of the American Medical Association*, *269*, 505-510.

Conway, G. S., Avey, C., & Rumsby, G. (1994). Genetics: The tyrosine kinase domain of the insulin receptor gene is normal in women with hyperinsulinaemia and polycystic ovary syndrome. *Human Reproduction*, *9*(9), 1681-1683.

Conway, G., Dewailly, D., Diamanti-Kandarakis, E., Escobar-Morreale, H. F., Franks, S., Gambineri, A., ... & Pfeifer, M. (2014). The polycystic ovary syndrome: a position statement from the European Society of Endocrinology. *European Journal of Endocrinology*, *171*(4), P1-P29.

Cooney, G. M., Dwan, K., Greig, C. A., Lawlor, D. A., Rimer, J., Waugh, . . . Mead, G. E. (2013). Exercise for depression. *The Cochrane database of systematic reviews*.

Cooney, L. G., Lee, I., Sammel, M. D., & Dokras, A. (2017). High prevalence of moderate and severe depressive and anxiety symptoms in polycystic ovary syndrome: a systematic review and meta-analysis. *Human Reproduction*, *32*(5), 1075-1091.

Copay, A. G., Subach, B. R., Glassman, S. D., Polly Jr, D. W., & Schuler, T. C. (2007). Understanding the minimum clinically important difference: a review of concepts and methods. *The Spine Journal*, *7*(5), 541-546.

Cornelissen, V. A., & Smart, N. A. (2013). Exercise training for blood pressure: a systematic review and meta-analysis. *Journal of the American Heart Association*, *2*(1), e004473.

Costa, E. C., DE, J. S., Stepto, N. K., Costa, I. B. B., Farias-Junior, L. F., Moreira, S. D. N. T., ... & Azevedo, G. D. (2018). Aerobic Training Improves Quality of Life in Women with Polycystic Ovary Syndrome. *Medicine and Science in Sports and Exercise*, *50*(7), 1357-1366.

Costello, M. F., Shrestha, B., Eden, J., Johnson, N., & Moran, L. J. (2007). Insulin-sensitising drugs versus the combined oral contraceptive pill for hirsutism, acne and risk of diabetes, cardiovascular

disease, and endometrial cancer in polycystic ovary syndrome. *Cochrane Database of Systematic Reviews*, (1).

Cott, C. A., Gignac, M. A., & Badley, E. M. (1999). Determinants of self-rated health for Canadians with chronic disease and disability. *Journal of Epidemiology & Community Health*, *53*(11), 731-736.

Couse, J. F., Yates, M. M., Sanford, R., Nyska, A., Nilson, J. H., & Korach, K. S. (2004). Formation of cystic ovarian follicles associated with elevated luteinizing hormone requires estrogen receptor-β. *Endocrinology*, *145*(10), 4693-4702.

Craig, C.L., Marshall, A.L., Sjöström, M., Bauman, A.E., Booth, M.L., Ainsworth, B.E., Pratt, M., Ekelund, U.L.F., Yngve, A., Sallis, J.F. and Oja, P. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine & Science in Sports & Exercise*, *35*(8), 1381-1395.

Cronin, L., Guyatt, G., Griffith, L., Wong, E., Azziz, R., Futterweit, W., ... & Dunaif, A. (1998). Development of a health-related quality-of-life questionnaire (PCOSQ) for women with polycystic ovary syndrome (PCOS). *The Journal of Clinical Endocrinology & Metabolism*, *83*(6), 1976-1987.

Crosignani, P. G., Colombo, M., Vegetti, W., Somigliana, E., Gessati, A., & Ragni, G. (2003). Overweight and obese anovulatory patients with polycystic ovaries: parallel improvements in anthropometric indices, ovarian physiology and fertility rate induced by diet. *Human Reproduction*, *18*(9), 1928-1932.

Curi, D. D., Fonseca, A. M., Marcondes, J. A. M., Almeida, J. A. M., Bagnoli, V. R., Soares Jr, J. M., & Baracat, E. C. (2012). Metformin versus lifestyle changes in treating women with polycystic ovary syndrome. *Gynecological Endocrinology*, *28*(3), 182-185.

Cussons, A. J., Stuckey, B. G., & Watts, G. F. (2007). Metabolic syndrome and cardiometabolic risk in PCOS. *Current Diabetes Reports*, *7*(1), 66.

Cutler, D. A., Pride, S. M., & Cheung, A. P. (2019). Low intakes of dietary fiber and magnesium are associated with insulin resistance and hyperandrogenism in polycystic ovary syndrome: A cohort study. *Food Science & Nutrition*, *7*(4), 1426-1437.

Davison, S. L., Bell, R., Donath, S., Montalto, J. G., & Davis, S. R. (2005). Androgen levels in adult females: changes with age, menopause, and oophorectomy. *The Journal of Clinical Endocrinology & Metabolism*, *90*(7), 3847-3853.

De Medeiros, S. F., Yamamoto, M. M. W., Bueno, H. B., Belizario, D., & Barbosa, J. S. (2014). Prevalence of elevated glycated hemoglobin concentrations in the polycystic ovary syndrome: anthropometrical and metabolic relationship in Amazonian women. *Journal of Clinical Medicine Research*, *6*(4), 278.

De Vries, L., Karasik, A., Landau, Z., Phillip, M., Kiviti, S., & Goldberg-Stern, H. (2007). Endocrine effects of valproate in adolescent girls with epilepsy. *Epilepsia*, *48*(3), 470-477.

Deeks, A. A., Gibson-Helm, M. E., Paul, E., & Teede, H. J. (2011). Is having polycystic ovary syndrome a predictor of poor psychological function including anxiety and depression? *Human Reproduction*, *26*(6), 1399-1407.

Deeks, J. J., Higgins, J. P., Altman, D. G., & Cochrane Statistical Methods Group. (2019). Analysing data and undertaking meta-analyses. *Cochrane Handbook for Systematic Reviews of Interventions*, 241-284.

Department of Health. UK Physical Activity Guidelines (2011). [online] Available at: https://www.gov.uk/government/publications/uk-physical-activity-guidelines. Accessed November 6th 2020.

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, *7*(3), 177-188.

Deswal, R., Yadav, A., & Dang, A. S. (2018). Sex hormone binding globulin-an important biomarker for predicting PCOS risk: A systematic review and meta-analysis. *Systems Biology in Reproductive Medicine*, *64*(1), 12-24.

DeUgarte, C. M., Bartolucci, A. A., & Azziz, R. (2005). Prevalence of insulin resistance in the polycystic ovary syndrome using the homeostasis model assessment. *Fertility and Sterility*, *83*(5), 1454-1460.

Dewailly, D., Catteau-Jonard, S., Reyss, A. C., Leroy, M., & Pigny, P. (2006). Oligoanovulation with polycystic ovaries but not overt hyperandrogenism. *The Journal of Clinical Endocrinology & Metabolism*, *91*(10), 3922-3927.

Diamanti-Kandarakis, E., & Dunaif, A. (2012). Insulin resistance and the polycystic ovary syndrome revisited: an update on mechanisms and implications. *Endocrine Reviews*, *33*(6), 981-1030.

Diamanti-Kandarakis, E., & Panidis, D. (2007). Unravelling the phenotypic map of polycystic ovary syndrome (PCOS): a prospective study of 634 women with PCOS. *Clinical Endocrinology*, *67*(5), 735-742.

Dokras, A., Sarwer, D. B., Allison, K. C., Milman, L., Kris-Etherton, P. M., Kunselman, A. R., ... & Fleming, J. (2016). Weight loss and lowering androgens predict improvements in health-related quality of life in women with PCOS. *The Journal of Clinical Endocrinology & Metabolism*, *101*(8), 2966-2974.

Dokras, A., Stener-Victorin, E., Yildiz, B. O., Li, R., Ottey, S., Shah, D., ... & Teede, H. (2018). Androgen Excess-Polycystic Ovary Syndrome Society: Position statement on depression, anxiety, quality of life, and eating disorders in polycystic ovary syndrome. *Fertility and Sterility*, *109*(5), 888-899.

Dolfing, J. G., Stassen, C. M., van Haard, P. M., Wolffenbuttel, B. H., & Schweitzer, D. H. (2011). Comparison of MRI-assessed body fat content between lean women with polycystic ovary syndrome (PCOS) and matched controls: less visceral fat with PCOS. *Human Reproduction*, *26*(6), 1495-1500.

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

Donà, S., Bacchi, E., & Moghetti, P. (2017). Is cardiorespiratory fitness impaired in PCOS women? A review of the literature. *Journal of Endocrinological Investigation*, *40*(5), 463-469.

Douglas, C. C., Norris, L. E., Oster, R. A., Darnell, B. E., Azziz, R., & Gower, B. A. (2006). Difference in dietary intake between women with polycystic ovary syndrome and healthy controls. *Fertility and Sterility*, *86*(2), 411-417.

Duley, L., Antman, K., Arena, J., Avezum, A., Blumenthal, M., Bosch, J., ... & Sleight, P. (2008). Specific barriers to the conduct of randomized trials. *Clinical Trials*, *5*(1), 40-48.

Dunaif, A. (1995). Hyperandrogenic anovulation (PCOS): a unique disorder of insulin action associated with an increased risk of non-insulin-dependent diabetes mellitus. *The American Journal of Medicine*, *98*(1), S33-S39.

Dunaif, A., & Book, C. B. (1997). Insulin resistance in the polycystic ovary syndrome. In *Clinical Research in Diabetes and Obesity*. Humana Press, Totowa, NJ. pp. 249-274

Dunaif, A., & Thomas, A. (2001). Current concepts in the polycystic ovary syndrome. *Annual Review of Medicine*, *52*(1), 401-419.

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

Dunaif, A., Segal, K. R., Shelley, D. R., Green, G., Dobrjansky, A., & Licholai, T. (1992). Evidence for distinctive and intrinsic defects in insulin action in polycystic ovary syndrome. *Diabetes*, *41*(10), 1257-1266.

Dwyer, T., Hosmer, D., Hosmer, T., Venn, A. J., Blizzard, C. L., Granger, R. H., ... & Dunstan, D. (2007). The inverse relationship between number of steps per day and obesity in a population-based sample–the AusDiab study. *International Journal of Obesity*, *31*(5), 797.

Ehrmann, D. A., Barnes, R. B., Rosenfield, R. L., Cavaghan, M. K., & Imperial, J. (1999). Prevalence of impaired glucose tolerance and diabetes in women with polycystic ovary syndrome. *Diabetes Care*, *22*(1), 141-146.

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

Ehrmann, D. A., Sturis, J., Byrne, M. M., Karrison, T., Rosenfield, R. L., & Polonsky, K. S. (1995). Insulin secretory defects in polycystic ovary syndrome. Relationship to insulin sensitivity and family history of non-insulin-dependent diabetes mellitus. *The Journal of Clinical Investigation*, *96*(1), 520-527.

Eichenberger, P.A., Diener, S.N., Kofmehl, R. and Spengler, C.M. (2013). Effects of exercise training on airway hyperreactivity in asthma: a systematic review and meta-analysis. *Sports Medicine*, *43*(11), pp.1157-1170.

Ekeland, E., Heian, F., & Hagen, K. B. (2005). Can exercise improve self-esteem in children and young people? A systematic review of randomised controlled trials. *British Journal of Sports Medicine*, *39*(11), 792-798.

Ekelund, U., Ward, H. A., Norat, T., Luan, J. A., May, A. M., Weiderpass, E., ... & Johnsen, N. F. (2015). Physical activity and all-cause mortality across levels of overall and abdominal adiposity in

European men and women: the European Prospective Investigation into Cancer and Nutrition Study (EPIC). *The American Journal of Clinical Nutrition*, *101*(3), 613-621.

El-Kotob, R., & Giangregorio, L. M. (2018). Pilot and feasibility studies in exercise, physical activity, or rehabilitation research. *Pilot and Feasibility Studies*, *4*(1), 137.

Elmlinger, M. W., Kühnel, W., Wormstall, H., & Döller, P. C. (2005). Reference intervals for testosterone, androstenedione and SHBG levels in healthy females and males from birth until old age. *Clinical Laboratory*, *51*(11-12), 625.

Elsheikh, M., and Murphy, C. (2008). What causes polycystic ovary syndrome. In: *Polycystic Ovary Syndrome*. Oxford University Press: Oxford, UK; pp. 9-25.

Emoto, M., Nishizawa, Y., Maekawa, K., Hiura, Y., Kanda, H., Kawagishi, T., ... & Morii, H. (1999). Homeostasis model assessment as a clinical index of insulin resistance in type 2 diabetic patients treated with sulfonylureas. *Diabetes Care*, *22*(5), 818-822.

Enders, C. K. and Bandalos, D. L. (2001) The Relative Performance of Full Information Maximum Likelihood Estimation for Missing Data in Structural Equation Models. *Educational Psychology Papers and Publications*. 64.

Enea C, Boisseau N, Fargeas-Gluck MA, Diaz V, Dugue B. Circulating androgens in women: exercise-induced changes. *Sports Medicine*, *41*(1), 1-15.

Ennos, R. (2007) *Statistical and Data Handling Skills in Biology* (2nd ed). Essex, UK: Pearson Education Limited, pp. 122.

Ennour-Idrissi, K., Maunsell, E., & Diorio, C. (2015). Effect of physical activity on sex hormones in women: a systematic review and meta-analysis of randomized controlled trials. *Breast Cancer Research*, *17*(1), 139.

ESHRE, T. R., & ASRM-Sponsored PCOS Consensus Workshop Group. (2004). Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertility and Sterility*, *81*(1), 19-25.

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis* (5th ed.). John Wiley and Sons Ltd: Chichester,UK.

Ezeh, U., Yildiz, B. O., & Azziz, R. (2013). Referral bias in defining the phenotype and prevalence of obesity in polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, *98*(6), E1088-E1096.

Farquhar, C. (2007). *Introduction and history of polycystic ovary syndrome*. In, Kovacs, G. T. and Norman, R. (eds). *Polycystic Ovary Syndrome.* Cambridge, UK: Cambridge University Press.

Farquhar, C. M., Birdsall, M., Manning, P., Mitchell, J. M., & France, J. T. (1994). The prevalence of polycystic ovaries on ultrasound scanning in a population of randomly selected women. *Australian and New Zealand Journal of Obstetrics and Gynaecology*, *34*(1), 67-72.

Farrell, K., & Antoni, M. H. (2010). Insulin resistance, obesity, inflammation, and depression in polycystic ovary syndrome: biobehavioral mechanisms and interventions. *Fertility and Sterility*, *94*(5), 1565-1574.

Faulkner J. H. (2003). Terminology for contractions of muscles during shortening, while isometric, and during lengthening. *Journal of Applied Physiology*, 95, 455-459.

Ferrari, P., Friedenreich, C., Matthews, C. E. (2007). The role of measurement error in estimating levels of physical activity**.** *American Journal of Epidemiology*, 166, 832-840.

Fewtrell, M. S., Kennedy, K., Singhal, A., Martin, R. M., Ness, A., Hadders-Algra, M., ... & Lucas, A. (2008). How much loss to follow-up is acceptable in long-term randomised trials and prospective studies? *Archives of Disease in Childhood*, *93*(6), 458-461.

Field, A. (2014) *Discovering Statistics Using IBM SPSS Statistics* (4th ed). London, UK: Sage.

Fielding, S., Fayers, P. M., McDonald, A., McPherson, G., Campbell, M. K., & RECORD Study Group (2008). Simple imputation methods were inadequate for missing not at random (MNAR) quality of life data. *Health and Quality of Life Outcomes*, *6*(1), 57.

Fontaine, K. R., Cheskin, L. J., & Barofsky, I. (1996). Health-related quality of life in obese persons seeking treatment. *Journal of Family Practice*, *43*(3), 265-271.

Fox, K. R. (1997). *The physical self: From motivation to well-being*. Chapaign, IL, USA: Human Kinetics.

Fox, K. R. (1999). The influence of physical activity on mental well-being. *Public Health Nutrition*, *2*(3a), 411-418.

Franks S. (1995). Polycystic ovary syndrome. *New England Journal of Medicine*, *333*(13), 853-861.

Fuad, Z. F., & Alobaidy, E. J. (2018). Association of HbA1c Levels with Body Mass Index in a Patient Diagnosed with Polycystic Ovary Syndrome. *Indian Journal of Public Health Research & Development*, *9*(7), 298-303.

Fung, T. T., Rexrode, K. M., Mantzoros, C. S., Manson, J. E., Willett, W. C., & Hu, F. B. (2009). Mediterranean diet and incidence and mortality of coronary heart disease and stroke in women. *Circulation*, *119*(8), 1093.

Furmaniak, A. C., Menig, M., & Markes, M. H. (2016). Exercise for women receiving adjuvant therapy for breast cancer. *Cochrane Database of Systematic Reviews*, (9).

Fux Otta, C., Wior, M., Iraci, G. S., Kaplan, R., Torres, D., Gaido, M. I., & Wyse, E. P. (2010). Clinical, metabolic, and endocrine parameters in response to metformin and lifestyle intervention in women with polycystic ovary syndrome: a randomized, double-blind, and placebo control trial. *Gynecological Endocrinology*, *26*(3), 173-178.

Galletly, C., Clark, A., Tomlinson, L., & Blaney, F. (1996). A group program for obese, infertile women: weight loss and improved psychological health. *Journal of Psychosomatic Obstetrics & Gynecology*, *17*(2), 125-128.

Galletly, C., Moran, L., Noakes, M., Clifton, P., Tomlinson, L., & Norman, R. (2007). Psychological benefits of a high-protein, low-carbohydrate diet in obese women with polycystic ovary syndrome—a pilot study. *Appetite*, *49*(3), 590-593.

Gambineri, A., Patton, L., Vaccina, A., Cacciari, M., Morselli-Labate, A. M., Cavazza, C., ... & Pasquali, R. (2006). Treatment with flutamide, metformin, and their combination added to a hypocaloric diet in overweight-obese women with polycystic ovary syndrome: a randomized, 12-

month, placebo-controlled study. *The Journal of Clinical Endocrinology & Metabolism*, *91*(10), 3970-3980.

Gandek, B., Ware, J.E., Aaronson, N.K., Apolone, G., Bjorner, J.B., Brazier, J.E., Bullinger, M., Kaasa, S., Leplege, A., Prieto, L. and Sullivan, M., (1998). Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. *Journal of clinical epidemiology*, *51*(11), 1171-1178.

García-Escudero, L. Á., & Gordaliza, A. (1999). Robustness properties of k means and trimmed k means. *Journal of the American Statistical Association*, *94*(447), 956-969.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2014). *Bayesian data analysis* (Vol. 2). Boca Raton, FL: Chapman.

Georgopoulos, N. A., Papadakis, E., Armeni, A. K., Katsikis, I., Roupas, N. D., & Panidis, D. (2014). Elevated serum androstenedione is associated with a more severe phenotype in women with polycystic ovary syndrome (PCOS). *Hormones*, *13*(2), 213-221.

Giallauria, F., Palomba, S., Maresca, L., Vuolo, L., Tafuri, D., Lombardi, G., ... & Francesco Orio. (2008). Exercise training improves autonomic function and inflammatory pattern in women with polycystic ovary syndrome (PCOS). *Clinical Endocrinology*, *69*(5), 792-798.

Gibson-Helm, M. E., Lucas, I. M., Boyle, J. A., & Teede, H. J. (2014). Women's experiences of polycystic ovary syndrome diagnosis. *Family Practice*, *31*(5), 545-549.

Gibson-Helm, M., Teede, H., Dunaif, A., & Dokras, A. (2016). Delayed diagnosis and a lack of information associated with dissatisfaction in women with polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, *102*(2), 604-612.

Gilbert, E. W., Tay, C. T., Hiam, D. S., Teede, H. J., & Moran, L. J. (2018). Comorbidities and complications of polycystic ovary syndrome: An overview of systematic reviews. *Clinical Endocrinology*, *89*(6), 683-699.

Glueck, C. J., Dharashivkar, S., Wang, P., Zhu, B., Gartside, P. S., Tracy, T., & Sieve, L. (2005). Obesity and extreme obesity, manifest by ages 20–24 years, continuing through 32–41 years in women, should alert physicians to the diagnostic likelihood of polycystic ovary syndrome as a reversible underlying endocrinopathy. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, *122*(2), 206-212.

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

Głuszak, O., Stopińska-Głuszak, U., Glinicki, P., Kapuścińska, R., Snochowska, H., Zgliczyński, W., & Dębski, R. (2012). Phenotype and metabolic disorders in polycystic ovary syndrome. *ISRN Endocrinology*, *2012*.

Goldzieher, J. W., & Green, J. A. (1962). The polycystic ovary. I. Clinical and histologic features. *The Journal of Clinical Endocrinology & Metabolism*, *22*(3), 325-338.

Gómez-Ambrosi, J., Silva, C., Galofré, J. C., Escalada, J., Santos, S., Gil, M. J., ... & Frühbeck, G. (2011). Body adiposity and type 2 diabetes: increased risk with a high body fat percentage even having a normal BMI. *Obesity*, *19*(7), 1439-1444.

Goodarzi, M. O., Quiñones, M. J., Azziz, R., Rotter, J. I., Hsueh, W. A., & 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.

Goodman, N. F., Cobin, R. H., Futterweit, W., Glueck, J. S., Legro, R. S., & Carmina, E. (2015). American Association of Clinical Endocrinologists, American College of Endocrinology, and androgen excess and PCOS society disease state clinical review: guide to the best practices in the evaluation and treatment of polycystic ovary syndrome-part 1. *Endocrine Practice*, *21*(11), 1291-1300.

Gorber, S. C., Tremblay, M., Moher, D., & Gorber, B. (2007). A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. *Obesity Reviews*, *8*(4), 307-326.

Gosman, G. G., King, W. C., Schrope, B., Steffen, K. J., Strain, G. W., Courcoulas, A. P., ... & Simhan, H. N. (2010). Reproductive health of women electing bariatric surgery. *Fertility and Sterility*, *94*(4), 1426-1431.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, *60*, 549-576.

Gray-Little, B., Williams, V.S. and Hancock, T.D., (1997). An item response theory analysis of the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin*, *23*(5), 443-451.

Greenwood, E. A., Noel, M. W., Kao, C. N., Shinkai, K., Pasch, L. A., Cedars, M. I., & Huddleston, H. G. (2016). Vigorous exercise is associated with superior metabolic profiles in polycystic ovary syndrome independent of total exercise expenditure. *Fertility and Sterility*, *105*(2), 486-493.

Grubbs, L., & Carter, J. (2002). The relationship of perceived benefits and barriers to reported exercise behaviors in college undergraduates. *Family & Community Health*, *25*(2), 76-84.

Gul, O. B., Somunkiran, A., Yucel, O., Demirci, F., & Ozdemir, I. (2008). The effect of ethinyl estradiol–cyproterone acetate treatment on homocysteine levels in women with polycystic ovary syndrome. *Archives of Gynecology and Obstetrics*, *277*(1), 25-30.

Guo, W., Fensom, G. K., Reeves, G. K., & Key, T. J. (2020). Physical activity and breast cancer risk: results from the UK Biobank prospective cohort. *British Journal of Cancer*, *122*(5), 726-732.

Guo, W., Key, T. J., & Reeves, G. K. (2019). Accelerometer compared with questionnaire measures of physical activity in relation to body size and composition: a large cross-sectional analysis of UK Biobank. *BMJ Open*, *9*(1), 1-9.

Gutin, B., Yin, Z., Humphries, M. C., & Barbeau, P. (2005). Relations of moderate and vigorous physical activity to fitness and fatness in adolescents. *The American Journal of Clinical Nutrition*, *81*(4), 746-750.

Guyatt, G. H., Oxman, A. D., Kunz, R., Brozek, J., Alonso-Coello, P., Rind, D., ... & Jaeschke, R. (2011). GRADE guidelines 6. Rating the quality of evidence—imprecision. *Journal of Clinical Epidemiology*, *64*(12), 1283-1293.

Guyatt, G. H., Oxman, A. D., Kunz, R., Vist, G. E., Falck-Ytter, Y., & Schünemann, H. J. (2008). What is "quality of evidence" and why is it important to clinicians? *BMJ*, *336*(7651), 995-998.

Guyatt, G., Oxman, A. D., Sultan, S., Brozek, J., Glasziou, P., Alonso-Coello, P., ... & Rind, D. (2013). GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. *Journal of Clinical Epidemiology*, *66*(2), 151-157.

Guzick, D. S., Wing, R., Smith, D., Berga, S. L., & Winters, S. J. (1994). Endocrine consequences of weight loss in obese, hyperandrogenic, anovulatory women. *Fertility and Sterility*, *61*(4), 598-604.

Hahn, S., Janssen, O.E., Tan, S., Pleger, K., Mann, K., Schedlowski, M., Kimmig, R., Benson, S., Balamitsa, E. and Elsenbruch, S., (2005). Clinical and psychological correlates of quality-of-life in polycystic ovary syndrome. *European Journal of Endocrinology*, *153*(6), 853-860.

Hallal, P. C., Simoes, E., Reichert, F. F., Azevedo, M. R., Ramos, L. R., Pratt, M., & Brownson, R. C. (2010). Validity and reliability of the telephone-administered international physical activity questionnaire in Brazil. *Journal of Physical Activity and Health*, *7*(3), 402-409.

Hamilton, M. T., Hamilton, D. G., & Zderic, T. W. (2004). Exercise physiology versus inactivity physiology: an essential concept for understanding lipoprotein lipase regulation. *Exercise and Sport Sciences Reviews*, *32*(4), 161.

Haqq, L., McFarlane, J., Dieberg, G., & Smart, N. (2015). The effect of lifestyle intervention on body composition, glycemic control, and cardiorespiratory fitness in polycystic ovarian syndrome: a systematic review and meta-analysis. *International Journal of Sport Nutrition and Exercise Metabolism*, *25*(6), 533-540.

Harman, M. S., Louvet, J. P., & Ross, G. T. (1975). Interaction of estrogen and gonadotrophins on follicular atresia. *Endocrinology*, *96*(5), 1145-1152.

Harris-Glocker M, Davidson K, Kochman L, Guzick D, Hoeger K. (2010). Improvement in quality-of-life questionnaire measures in obese adolescent females with polycystic ovary syndrome treated with lifestyle changes and oral contraceptives, with or without metformin. *Fertility and Sterility*, *93*(3), 1016–1019.

Harrison, C. L., Lombard, C. B., Moran, L. J., & Teede, H. J. (2011). Exercise therapy in polycystic ovary syndrome: a systematic review. *Human Reproduction Update*, *17*(2), 171-183.

Harter, S. (1989). Causes, correlates, and the functional role of global self-worth: A lifespan perspective. In, Kolligian, J. and Sternberg, R. (Eds.), *Perceptions of Competence and Incompetence Across the Lifespan*. New Haven, CT: Yale University Press, pp. 67-87.

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (applied statistics)*, *28*(1), 100-108.

Hay, J. L., Ford, J. S., Klein, D., Primavera, L. H., Buckley, T. R., Stein, T. R., ... & Ostroff, J. S. (2003). Adherence to colorectal cancer screening in mammography-adherent older women. *Journal of Behavioral Medicine*, *26*(6), 553-576.

Healy, G. N., Clark, B. K., Winkler, E. A., Gardiner, P. A., Brown, W. J., & Matthews, C. E. (2011). Measurement of adults' sedentary time in population-based studies. *American Journal of Preventive Medicine*, *41*(2), 216-227.

Healy, G. N., Dunstan, D. W., Salmon, J., Cerin, E., Shaw, J. E., Zimmet, P. Z., & Owen, N. (2008). Breaks in sedentary time: beneficial associations with metabolic risk. *Diabetes Care*, *31*(4), 661-666.

Hennig, C. (2003). Clusters, outliers, and regression: fixed point clusters. *Journal of Multivariate Analysis*, *86*(1), 183-212.

Hesketh, K., Wake, M., & Waters, E. (2004). Body mass index and parent-reported self-esteem in elementary school children: evidence for a causal relationship. *International Journal of Obesity*, *28*(10), 1233.

Heymans, M. W. & Eekhout, I. (2019). *Applied Missing Data Analysis with SPSS and (R) Studio* [online]. Available at: https://bookdown.org/mwheymans/bookmi/missing-data-evaluation.html#missing-data-patterns [last accessed 14th August 2020].

Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2019). *Cochrane Handbook for Systematic Reviews of Interventions*. John Wiley & Sons.

Higgins, J.P.T., Altman, D.G. (2011). Assessing risk of bias in included studies. In: Higgins JPT, Green S, (eds). *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester, UK: The Cochrane Collaboration and John Wiley and Sons Ltd. pp. 187-241.

Higgins, T. J., Middleton, K. R., Winner, L., & Janelle, C. M. (2014). Physical activity interventions differentially affect exercise task and barrier self-efficacy: A meta-analysis. *Health Psychology*, *33*(8), 891–903.

Himelein, M. J., & Thatcher, S. S. (2006). Depression and body image among women with polycystic ovary syndrome. *Journal of Health Psychology*, *11*(4), 613-625.

Hjellvik, V., Sakshaug, S., & Strøm, H. (2012). Body mass index, triglycerides, glucose, and blood pressure as predictors of type 2 diabetes in a middle-aged Norwegian cohort of men and women. *Clinical Epidemiology*, *4*, 213.

Hochberg, Y., & Tamhane, A. C. (1987) *Multiple comparison procedures*. John Wiley & Sons: Chicago

Hoeger, K. M., Kochman, L., Wixom, N., Craig, K., Miller, R. K., & Guzick, D. S. (2004). A randomized, 48-week, placebo-controlled trial of intensive lifestyle modification and/or metformin therapy in overweight women with polycystic ovary syndrome: a pilot study. *Fertility and Sterility*, *82*(2), 421-429.

Hoeger, K., Davidson, K., Kochman, L., Cherry, T., Kopin, L., & Guzick, D. S. (2008). The impact of metformin, oral contraceptives, and lifestyle modification on polycystic ovary syndrome in obese adolescent women in two randomized, placebo-controlled clinical trials. *The Journal of Clinical Endocrinology & Metabolism*, *93*(11), 4299-4306.

Hollander, M., Wolfe, D. A., & Chicken, E. (2014). *Nonparametric Statistical Methods* (3rd ed). John Wiley & Sons: New Jersey.

Hollinrake, E., Abreu, A., Maifeld, M., Van Voorhis, B. J., & Dokras, A. (2007). Increased risk of depressive disorders in women with polycystic ovary syndrome. *Fertility and Sterility*, *87*(6), 1369-1376.

Holte, J., Bergh, T., Berne, C. H., Wide, L., & 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.

Hopkins, W., Marshall, S., Batterham, A., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. *Medicine & Science in Sports & Exercise*, *41*(1), 3.

Hsu, M. I. (2013). Changes in the PCOS phenotype with age. *Steroids*, *78*(8), 761-766.

Hu, F. B., Sigal, R. J., Rich-Edwards, J. W., Colditz, G. A., Solomon, C. G., Willett, W. C., ... & Manson, J. E. (1999). Walking compared with vigorous physical activity and risk of type 2 diabetes in women: a prospective study. *JAMA*, *282*(15), 1433-1439.

Hu, G., Tuomilehto, J., Silventoinen, K., Barengo, N., & Jousilahti, P. (2004). Joint effects of physical activity, body mass index, waist circumference and waist-to-hip ratio with the risk of cardiovascular disease among middle-aged Finnish men and women. *European Heart Journal*, *25*(24), 2212-2219.

Hupin, D., Roche, F., Gremeaux, V., Chatard, J. C., Oriol, M., Gaspoz, J. M., ... & Edouard, P. (2015). Even a low-dose of moderate-to-vigorous physical activity reduces mortality by 22% in adults aged≥ 60 years: a systematic review and meta-analysis. *British Journal of Sports Medicine*, *49*(19), 1262-1267.

Huxley, R., Mendis, S., Zheleznyakov, E., Reddy, S., & Chan, J. (2010). Body mass index, waist circumference and waist: hip ratio as predictors of cardiovascular risk—a review of the literature. *European Journal of Clinical Nutrition*, *64*(1), 16-22.

Imani, B., Eijkemans, M. J., de Jong, F. H., Payne, N. N., Bouchard, P., Giudice, L. C., & Fauser, B. C. (2000). Free androgen index and leptin are the most prominent endocrine predictors of ovarian response during clomiphene citrate induction of ovulation in normogonadotropic oligoamenorrheic infertility. *The Journal of Clinical Endocrinology & Metabolism*, *85*(2), 676-682.

Imani, B., Eijkemans, M. J., te Velde, E. R., Habbema, J. D. F., & Fauser, B. C. (1999). Predictors of chances to conceive in ovulatory patients during clomiphene citrate induction of ovulation in normogonadotropic oligoamenorrheic infertility. *The Journal of Clinical Endocrinology & Metabolism*, *84*(5), 1617-1622.

IPAQ Research Committee. (2005). Guidelines for data processing and analysis of the International Physical Activity Questionnaire (IPAQ)-short and long forms. *http://www. ipaq. ki. se/scoring. pdf*.

Iuorno, M.J. (2007). Role of hyperinsulinemic insulin resistance in polycystic ovary syndrome. In, Kovacs, G.T., Norman, R. (eds). *Polycystic Ovary Syndrome*, 2nd ed. Cambridge, UK: Cambridge University Press. pp. 204-214.

Jabbour, H. N., Kelly, R. W., Fraser, H. M., & Critchley, H. O. (2006). Endocrine regulation of menstruation. *Endocrine Reviews*, *27*(1), 17-46.

Jaffe, R. B. (2006). Combined lifestyle modification and metformin in obese patients with polycystic ovary syndrome: a randomized, placebo-controlled, double blind multicentre study. *Obstetrical & Gynecological Survey*, *61*(2), 108–109.

Jakobsen, J. C., Gluud, C., Wetterslev, J., & Winkel, P. (2017). When and how should multiple imputation be used for handling missing data in randomised clinical trials–a practical guide with flowcharts. *BMC Medical Research Methodology*, *17*(1), 162.

Jakobsen, J. C., Gluud, C., Winkel, P., Lange, T., & Wetterslev, J. (2014). The thresholds for statistical and clinical significance–a five-step procedure for evaluation of intervention effects in randomised clinical trials. *BMC Medical Research Methodology*, *14*(1), 1-12.

Janssen, I., Katzmarzyk, P. T., & Ross, R. (2004). Waist circumference and not body mass index explains obesity-related health risk. *The American Journal of Clinical Nutrition*, *79*(3), 379-384.

Janz, N. K., & Becker, M. H. (1984). The health belief model: A decade later. *Health Education Quarterly*, *11*(1), 1-47.

Jedel, E., Holm, G., Labrie, F., Oden, A., Nilsson, L., Janson, P. O., ... & Stener-Victorin, E. (2010). Effect of Low-Frequency Electro-Acupuncture on Serum Testosterone and Menstrual Pattern in Women with Polycystic Ovary Syndrome Compared to Physical Exercise: Randomised Controlled Trial. *Endocrine Reviews*, *31*(3).

Jedel, E., Labrie, F., Odén, A., Holm, G., Nilsson, L., Janson, P. O., ... & Stener-Victorin, E. (2011). Impact of electro-acupuncture and physical exercise on hyperandrogenism and oligo/amenorrhea in women with polycystic ovary syndrome: a randomized controlled trial. *American Journal of Physiology-Endocrinology and Metabolism*, *300*(1), E37-E45.

Jedel, E., Waern, M., Gustafson, D., Landen, M., Eriksson, E., Holm, G., ... & Stener-Victorin, E. (2009). Anxiety and depression symptoms in women with polycystic ovary syndrome compared with controls matched for body mass index. *Human Reproduction*, *25*(2), 450-456.

Jennersjö, P., Ludvigsson, J., Länne, T., Nystrom, F. H., Ernerudh, J., & Östgren, C. J. (2012). Pedometer-determined physical activity is linked to low systemic inflammation and low arterial stiffness in type 2 diabetes. *Diabetic Medicine*, *29*(9), 1119-1125.

Johansson, J., Redman, L., Veldhuis, P. P., Sazonova, A., Labrie, F., Holm, G., ... & Stener-Victorin, E. (2013). Acupuncture for ovulation induction in polycystic ovary syndrome: a randomized controlled trial. *American Journal of Physiology-Endocrinology and Metabolism*, *304*(9), E934-E943.

Jonard, S., Robert, Y., Ardaens, Y., Dewailly, D. (2007). Imaging polycystic ovaries. In, Kovacs, G. T., Norman, R. (eds). *Polycystic ovary syndrome* (2nd ed). Cambridge, UK: Cambridge University Press. pp. 48-64.

Jones, G.L., Hall, J.M., Balen, A.H. and Ledger, W.L., (2007). Health-related quality of life measurement in women with polycystic ovary syndrome: a systematic review. *Human Reproduction Update*, *14*(1), 15-25.

Jones, G.L., Palep-Singh, M., Ledger, W.L., Balen, A.H., Jenkinson, C., Campbell, M.J. and Lashen, H., (2010). Do South Asian women with PCOS have poorer health-related quality of life than Caucasian women with PCOS? A comparative cross-sectional study. *Health and Quality of Life Outcomes*, *8*(1), 149.

Jones, M., & Nies, M. A. (1996). The relationship of perceived benefits of and barriers to reported exercise in older African American women. *Public Health Nursing*, *13*(2), 151-158.

Joseph, R. P., Royse, K. E., Benitez, T. J., & Pekmezi, D. W. (2014). Physical activity and quality of life among university students: exploring self-efficacy, self-esteem, and affect as potential mediators. *Quality of Life Research*, *23*(2), 659-667.

Kahn, S. E., Hull, R. L., & Utzschneider, K. M. (2006). Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature*, *444*(7121), 840-846.

Kant, A. K., Schatzkin, A., Graubard, B. I., & Schairer, C. (2000). A prospective study of diet quality and mortality in women. *JAMA*, *283*(16), 2109-2115.

Karabulut, A., Yaylali, G. F., Demirlenk, S., Sevket, O., & Acun, A. (2012). Evaluation of body fat distribution in PCOS and its association with carotid atherosclerosis and insulin resistance. *Gynecological Endocrinology*, *28*(2), 111-114.

Karimzadeh, M. A., & Javedani, M. (2010). An assessment of lifestyle modification versus medical treatment with clomiphene citrate, metformin, and clomiphene citrate–metformin in patients with polycystic ovary syndrome. *Fertility and Sterility*, *94*(1), 216-220.

Kassambara, A. (2017). *Practical guide to cluster analysis in R: Unsupervised machine learning* (Vol. 1). STHDA.

Kaufman, L. R., & Rousseeuw, P. PJ (1990) *Finding groups in data: An introduction to cluster analysis.* Hoboken, NJ: John Wiley & Sons Inc.

Kelley, G.A., Kelley, K.S. and Hootman, J.M., (2015). Effects of exercise on depression in adults with arthritis: a systematic review with meta-analysis of randomized controlled trials. *Arthritis Research & Therapy*, *17*(1), 21.

Kelly, P., Kahlmeier, S., Götschi, T., Orsini, N., Richards, J., Roberts, N., ... & Foster, C. (2014). Systematic review and meta-analysis of reduction in all-cause mortality from walking and cycling and shape of dose response relationship. *International Journal of Behavioral Nutrition and Physical Activity*, *11*(1), 132.

Kendall, M. G. (1970) *Rank Correlation Methods.* London, UK: Charles Griffin.

Kiddy, D. S., Hamilton-Fairley, D., Bush, A., Short, F., Anyaoku, V., Reed, M. J., & Franks, S. (1992). Improvement in endocrine and ovarian function during dietary treatment of obese women with polycystic ovary syndrome. *Clinical Endocrinology*, *36*(1), 105-111.

Kiddy, D. S., Sharp, P. S., White, D. M., Scanlon, M. F., Mason, H. D., Bray, C. S., ... & Franks, S. (1990). Differences in clinical and endocrine features between obese and non-obese subjects with polycystic ovary syndrome: an analysis of 263 consecutive cases. *Clinical Endocrinology*, *32*(2), 213-220.

Kim, J. J., Choi, Y. M., Cho, Y. M., Jung, H. S., Chae, S. J., Hwang, K. R., ... & Moon, S. Y. (2012). Prevalence of elevated glycated hemoglobin in women with polycystic ovary syndrome. *Human Reproduction*, *27*(5), 1439-1444.

Kim, J. J., Hwang, K. R., Choi, Y. M., Moon, S. Y., Chae, S. J., Park, C. W., ... & Lee, B. S. (2014). Complete phenotypic and metabolic profiles of a large consecutive cohort of untreated Korean women with polycystic ovary syndrome. *Fertility and Sterility*, *101*(5), 1424-1430.

Kim, O., & Kim, K. (2001). Body weight, self-esteem, and depression in Korean female adolescents. *Adolescence*, *36*, 315-322.

King, A. C., Haskell, W. L., Young, D. R., Oka, R. K., & Stefanick, M. L. (1995). Long-term effects of varying intensities and formats of physical activity on participation rates, fitness, and lipoproteins in men and women aged 50 to 65 years. *Circulation*, *91*(10), 2596-2604.

Kirschner, M. A., & Bardin, C. W. (1972). Androgen production and metabolism in normal and virilized women. *Metabolism*, *21*(7), 667-688.

Kite, C., Lahart, I. M., Afzal, I., Broom, D. R., Randeva, H., Kyrou, I., & Brown, J. E. (2019). Exercise, or exercise and diet for the management of polycystic ovary syndrome: a systematic review and meta-analysis. *Systematic Reviews*, *8*(1), 51.

Knuttgen, H. (1978). Force, work, power, and exercise. *Medicine and Science in Sports*, *10*(3), 227-228.

Kodama, S., Saito, K., Tanaka, S., Maki, M., Yachi, Y., Asumi, M., ... & Yamada, N. (2009). Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis. *JAMA*, *301*(19), 2024-2035.

Koistinen, H. A., Galuska, D., Chibalin, A. V., Yang, J., Zierath, J. R., Holman, G. D., & Wallberg-Henriksson, H. (2003). 5-amino-imidazole carboxamide riboside increases glucose transport and cell-surface GLUT4 content in skeletal muscle from subjects with type 2 diabetes. *Diabetes*, *52*(5), 1066-1072.

Konopka, A. R., Asante, A., Lanza, I. R., Robinson, M. M., Johnson, M. L., Dalla Man, C., ... & Nair, K. S. (2015). Defects in mitochondrial efficiency and H2O2 emissions in obese women are restored to a lean phenotype with aerobic exercise training. *Diabetes*, *64*(6), 2104-2115.

Kousta, E., & Franks, S. (2007). Long-term health consequences of polycystic ovary syndrome. In, Kovacs, G. and Norman, R. (eds) *Polycystic ovary syndrome* (2nd Ed), Cambridge, UK: Cambridge University Press.

Kumar, P., & Arora, S. (2014). Orlistat in polycystic ovarian syndrome reduces weight with improvement in lipid profile and pregnancy rates. *Journal of Human Reproductive Sciences*, *7*(4), 255.

Kyrou, I., Weickert, M. O., & Randeva, H. (2014). Diagnosis and management of PCOS. *Endocrinology and Diabetes: Case Studies, Questions and Commentaries. UK: springer*, 9.

Kyu, H. H., Bachman, V. F., Alexander, L. T., Mumford, J. E., Afshin, A., Estep, K., ... & Cercy, K. (2016). Physical activity and risk of breast cancer, colon cancer, diabetes, ischemic heart disease, and ischemic stroke events: systematic review and dose-response meta-analysis for the Global Burden of Disease Study 2013. *BMJ*, *354*, i3857.

Läärä, E. (2011) Study Designs for Biobank-Based Epidemiologic Research on Chronic Diseases. In, Dillner, J. (Ed) *Methods in Biobanking*. Humana Press: New York.

Ladson, G., Dodson, W. C., Sweet, S. D., Archibong, A. E., Kunselman, A. R., Demers, L. M., ... & Legro, R. S. (2011). Effects of metformin in adolescents with polycystic ovary syndrome undertaking lifestyle therapy: a pilot randomized double-blind study. *Fertility and Sterility*, *95*(8), 2595-2598.

Ladson, G., Dodson, W. C., Sweet, S. D., Archibong, A. E., Kunselman, A. R., Demers, L. M., ... & Legro, R. S. (2011). The effects of metformin with lifestyle therapy in polycystic ovary syndrome: a randomized double-blind study. *Fertility and Sterility*, *95*(3), 1059-1066.

Laitinen, J., Taponen, S., Martikainen, H., Pouta, A., Millwood, I., Hartikainen, A. L., ... & Järvelin, M. R. (2003). Body size from birth to adulthood as a predictor of self-reported polycystic ovary syndrome symptoms. *International Journal of Obesity*, *27*(6), 710-715.

Lakens, D. (2015). Why you should use omega-squared instead of eta-squared [blog]. From http://daniellakens.blogspot.com [last accessed April 6th 2020)

Lakoski, S. G., Willis, B. L., Barlow, C. E., Leonard, D., Gao, A., Radford, N. B., ... & Jones, L. W. (2015). Midlife cardiorespiratory fitness, incident cancer, and survival after cancer in men: the cooper center longitudinal study. *JAMA Oncology*, *1*(2), 231-237.

Lamb, J. D., Johnstone, E. B., Rousseau, J. A., Jones, C. L., Pasch, L. A., Cedars, M. I., & Huddleston, H. G. (2011). Physical activity in women with polycystic ovary syndrome: prevalence, predictors, and positive health associations. *American Journal of Obstetrics and Gynecology*, *204*(4), 352-e1.

Langer, G., Meerpohl, J. J., Perleth, M., Gartlehner, G., Kaminski-Hartenthaler, A., & Schünemann, H. (2012). GRADE guidelines: 1. Introduction–GRADE evidence profiles and summary of findings tables. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, *106*(5), 357-368.

Le Donne, M., Alibrandi, A., Giarrusso, R., Lo, M. I., & Muraca, U. (2012). Diet, metformin and inositol in overweight and obese women with polycystic ovary syndrome: effects on body composition. *Minerva Ginecologica*, *64*(1), 23.

Lean, M. E. J., Han, T. S., & Morrison, C. E. (1995). Waist circumference as a measure for indicating need for weight management. *BMJ*, *311*(6998), 158-161.

Lee, I. M., & Skerrett, P. J. (2001). Physical activity and all-cause mortality: what is the dose-response relation? *Medicine and Science in Sports and Exercise*, *33*(6; SUPP), S459-S471.

Legro, R. S. (2000). The genetics of obesity Lessons for polycystic ovary syndrome. *Annals of the New York Academy of Sciences*, *900*(1), 193-202.

Legro, R. S. (2012). Obesity and PCOS: implications for diagnosis and treatment. *Seminars in Reproductive Medicine*, *30*(6). NIH Public Access.

Legro, R. S., Arslanian, S. A., Ehrmann, D. A., Hoeger, K. M., Murad, M. H., Pasquali, R., & 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., & Kauffman, R. P. (2004). Detecting insulin resistance in polycystic ovary syndrome: purposes and pitfalls. *Obstetrical & Gynecological Survey*, *59*(2), 141-154.

Legro, R. S., Chiu, P., Kunselman, A. R., Bentley, C. M., Dodson, W. C., & Dunaif, A. (2005). Polycystic ovaries are common in women with hyperandrogenic chronic anovulation but do not predict metabolic or reproductive phenotype. *The Journal of Clinical Endocrinology & Metabolism*, *90*(5), 2571-2579.

Legro, R. S., Dodson, W. C., Kris-Etherton, P. M., Kunselman, A. R., Stetter, C. M., Williams, N. I., ... & Sarwer, D. B. (2015). Randomized controlled trial of preconception interventions in infertile women with polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, *100*(11), 4048-4058.

Legro, R. S., Dodson, W. C., Kunselman, A. R., Kris-Etherton, P. M., Allison, K. C., Sarwer, D. B., ... & Coutifaris, C. (2014). Effects of preconception intervention on the PCOS phenotype, ovulation, and live birth rates: a multi-center, multi-phase RCT. *Fertility and Sterility*, *102*(3), e2.

Legro, R. S., Kunselman, A. R., & Dunaif, A. (2001). Prevalence and predictors of dyslipidemia in women with polycystic ovary syndrome. *The American Journal of Medicine*, *111*(8), 607-613.

Legro, R. S., Kunselman, A. R., Dodson, W. C., & 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. *The Journal of Clinical Endocrinology & Metabolism*, *84*(1), 165-169.

Leitzmann, M. F., Park, Y., Blair, A., Ballard-Barbash, R., Mouw, T., Hollenbeck, A. R., & Schatzkin, A. (2007). Physical activity recommendations and decreased risk of mortality. *Archives of Internal Medicine*, *167*(22), 2453-2460.

Leonhardt, H., Hellström, M., Gull, B., Lind, A. K., Nilsson, L., Janson, P. O., & Stener-Victorin, E. (2015). Serum anti-Müllerian hormone and ovarian morphology assessed by magnetic resonance imaging in response to acupuncture and exercise in women with polycystic ovary syndrome: secondary analyses of a randomized controlled trial. *Acta Obstetricia et Gynecologica Scandinavica*, *94*(3), 279-287.

Li, C. (2013). Little's test of missing completely at random. *The Stata Journal*, *13*(4), 795-809.

Li, T., Wei, S., Shi, Y., Pang, S., Qin, Q., Yin, J., ... & Liu, L. (2016). The dose–response effect of physical activity on cancer mortality: findings from 71 prospective cohort studies. *British Journal of Sports Medicine*, *50*(6), 339-345.

Li, X., Feng, Y., Lin, J. F., Billig, H., & Shao, R. (2014). Endometrial progesterone resistance and PCOS. *Journal of Biomedical Science*, *21*(1), 1-7.

Li, Y., Li, Y., Ng, E.H.Y., Stener-Victorin, E., Hou, L., Wu, T., Han, F. and Wu, X., (2011). 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.

Liao, L. M., Nesic, J., Chadwick, P. M., Brooke-Wavell, K., & Prelevic, G. M. (2008). Exercise and body image distress in overweight and obese women with polycystic ovary syndrome: a pilot investigation. *Gynecological Endocrinology*, *24*(10), 555-561.

Lim, S. S., Davies, M. J., Norman, R. J., & Moran, L. J. (2012). Overweight, obesity and central obesity in women with polycystic ovary syndrome: a systematic review and meta-analysis. *Human Reproduction Update*, *18*(6), 618-637.

Lim, S.S., Hutchison, S.K., Van Ryswyk, E., Norman, R.J., Teede, H.J. and Moran, L.J., (2019). Lifestyle changes in women with polycystic ovary syndrome. *Cochrane Database of Systematic Reviews*, (3).

Lin, A.W., Kazemi, M., Jarrett, B. Y., Vanden Brink, H., Hoeger, K. M., Spandorfer, S. D., Lujan, M. E. (2019). Dietary and Physical Activity Behaviours in Women with Polycystic Ovary Syndrome per the New International Evidence-Based Guideline. *Nutrients, 11*(2271).

Lindholm, Å., Bixo, M., Björn, I., Wölner-Hanssen, P., Eliasson, M., Larsson, A., ... & Poromaa, I. S. (2008). Effect of sibutramine on weight reduction in women with polycystic ovary syndrome: a randomized, double-blind, placebo-controlled trial. *Fertility and Sterility*, *89*(5), 1221-1228.

Liou, T. H., Yang, J. H., Hsieh, C. H., Lee, C. Y., Hsu, C. S., & Hsu, M. I. (2009). Clinical and biochemical presentations of polycystic ovary syndrome among obese and non-obese women. *Fertility and Sterility*, *92*(6), 1960-1965.

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198-1202.

Little, R. J. A., Rubin, D. B. (1987). *Statistical Analysis with Missing Data*. New York, NY: Wiley.

Littlejohns, T. J., Sudlow, C., Allen, N. E., & Collins, R. (2019). UK Biobank: opportunities for cardiovascular research. *European Heart Journal*, *40*(14), 1158-1166.

Liu, L., Shi, Y., Li, T., Qin, Q., Yin, J., Pang, S., ... & Wei, S. (2016). Leisure time physical activity and cancer risk: evaluation of the WHO's recommendation based on 126 high-quality epidemiological studies. *British Journal of Sports Medicine*, *50*(6), 372-378.

Liu, N., Ma, Y., Wang, S., Zhang, X., Zhang, Q., Zhang, X., ... & Qiao, J. (2012). Association of the genetic variants of luteinizing hormone, luteinizing hormone receptor and polycystic ovary syndrome. *Reproductive Biology and Endocrinology*, *10*(1), 36.

Liu, S. Y., Wrosch, C., Miller, G. E., & Pruessner, J. C. (2014). Self-esteem change and diurnal cortisol secretion in older adulthood. *Psychoneuroendocrinology*, *41*, 111-120.

Liu, Y. X., & Hsueh, A. J. (1986). Synergism between granulosa and theca-interstitial cells in estrogen biosynthesis by gonadotropin-treated rat ovaries: studies on the two-cell, two-gonadotropin hypothesis using steroid antisera. *Biology of Reproduction*, *35*(1), 27-36.

Lizneva, D., Kirubakaran, R., Mykhalchenko, K., Suturina, L., Galina, C., & Azziz, R. (2016). Phenotypes and body mass in women with PCOS identified in clinical versus unselected populations: systematic review and meta-analysis. *Fertility and Sterility*, *106*, 1510-20.

Lizneva, D., Suturina, L., Walker, W., Brakta, S., Gavrilova-Jordan, L., & Azziz, R. (2016). Criteria, prevalence, and phenotypes of polycystic ovary syndrome. *Fertility and Sterility*, *106*(1), 6-15.

Look, A. R. G., Wing, R. R., Bolin, P., Brancati, F. L., Bray, G. A., Clark, J. M., . . . Yanovski, S. Z. (2013). Cardiovascular effects of intensive lifestyle intervention in type 2 diabetes. *New England Journal of Medicine*, *369*(2), 145-154.

Lord, J., Thomas, R., Fox, B., Acharya, U., & Wilkin, T. (2006). The central issue? Visceral fat mass is a good marker of insulin resistance and metabolic disturbance in women with polycystic ovary syndrome. *BJOG: An International Journal of Obstetrics & Gynaecology*, *113*(10), 1203-1209.

Love-Koh, J., & Taylor, M. (2018). *Physical Activity and the Environment: Final Report*. [online] https://www.nice.org.uk/guidance/ng90/evidence/march-2018-economic-modelling-report-pdf-4788819757. [Last accessed: 29th October 2020].

Lundgren, K. M., Moholdt, T. and Riphagen, I. I. (2016). Exercise training in polycystic ovary syndrome: a systematic review. PROSPERO 2016: CRD42016034048 Available from: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016034048.

Luque-Ramírez, M., Alpañés, M., Sanchón, R., Fernández-Durán, E., Ortiz-Flores, A. E., & Escobar-Morreale, H. F. (2015). Referral bias in female functional hyperandrogenism and polycystic ovary syndrome. *European Journal of Endocrinology*, *173*(5), 603.

Ma, L. K., Jin, L. N., Yu, Q., & Xu, L. (2007). Effect of lifestyle adjustment, metformin and rosiglitazone in polycystic ovary syndrome. *Zhonghua fu chan ke za zhi*, *42*(5), 294-297.

Ma, Y.M., Li, R., Jie, Q. I. A. O., Zhang, X. W., Wang, S. Y., Zhang, Q. F., Li, L. I., ... & Zhang, X. (2010). Characteristics of abnormal menstrual cycle and polycystic ovary syndrome in community and hospital populations. *Chinese Medical Journal*, *123*(16), 2185-2189.

Machlitt, A. R., Steinheim, E., & Pfueller, B. F. (2007). Effects of metformin plus lifestyle modification on menstrual cycle and body mass index are not influenced by insulin resistance: a 16 week, double-blind, controlled, randomised study in PCOS. *Fertility and Sterility*, *88*, S15-S16.

Macniven, R., Pye, V., Merom, D., Milat, A., Monger, C., Bauman, A., & Van Der Ploeg, H. (2014). Barriers and enablers to physical activity among older Australians who want to increase their physical activity levels. *Journal of Physical Activity and Health*, *11*(7), 1420-1429.

Maher, C. G., Sherrington, C., Herbert, R. D., Moseley, A. M., & Elkins, M. (2003). Reliability of the PEDro scale for rating quality of randomized controlled trials. *Physical Therapy*, *83*(8), 713-721.

Mani, H., Chudasama, Y., Hadjiconstantinou, M., Bodicoat, D. H., Edwardson, C., Levy, M. J., ... & Khunti, K. (2018). Structured education programme for women with polycystic ovary syndrome: a randomised controlled trial. *Endocrine Connections*, *7*(1), 26-35.

Mani, H., Henson, J., Gray, L. J., Sutton, S., Levy, M., Howlett, T., Khunti, K. and Davies, M. J. (2013). A systematic review and meta-analysis comparing the pharmacological (insulin sensitizers) or lifestyle management in women with polycystic ovary syndrome and their effect on weight reduction. PROSPERO 2013: CRD42013004169 Available from: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42013004169.

Manjoo, P., Joseph, L., & Dasgupta, K. (2012). Abdominal adiposity and daily step counts as determinants of glycemic control in a cohort of patients with type 2 diabetes mellitus. *Nutrition & Diabetes*, *2*(1), e25.

Manjoo, P., Joseph, L., Pilote, L., & Dasgupta, K. (2010). Sex differences in step count-blood pressure association: a preliminary study in type 2 diabetes. *PLoS One*, *5*(11), e14086.

Manna, P. R., Dyson, M. T., & Stocco, D. M. (2009). Regulation of the steroidogenic acute regulatory protein gene expression: present and future perspectives. *Molecular Human Reproduction*, *15*(6), 321-333.

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

Marin, C. M., Bartolucci, A. A., & Azziz, R. (2003). Prevalence of insulin resistance in polycystic ovary syndrome (PCOS) patients using the homeostatic measurement assessment (HOMA-IR). *Fertility and Sterility*, *80*, 274-275.

Mario, F. M., do Amarante, F., Toscani, M. K., & Spritzer, P. M. (2012). Lean muscle mass in classic or ovulatory PCOS: association with central obesity and insulin resistance. *Experimental and Clinical Endocrinology & Diabetes*, *120*(09), 511-516.

Mario, F. M., Graff, S. K., & Spritzer, P. M. (2017). Habitual physical activity is associated with improved anthropometric and androgenic profile in PCOS: a cross-sectional study. *Journal of Endocrinological Investigation*, *40*(4), 377-384.

Marsh, H. W. (1993). Relations between global and specific domains of self: The importance of individual importance, certainty, and ideals. *Journal of Personality and Social Psychology*, *65*(5), 975.

Marzouk, T. M., & Ahmed, W. A. S. (2015). Effect of dietary weight loss on menstrual regularity in obese young adult women with polycystic ovary syndrome. *Journal of Pediatric and Adolescent Gynecology*, *28*(6), 457-461.

Mather, K. J., Kwan, F., & Corenblum, B. (2000). Hyperinsulinemia in polycystic ovary syndrome correlates with increased cardiovascular risk independent of obesity. *Fertility and Sterility*, *73*(1), 150-156.

Matthews, D. R., Hosker, J. P., Rudenski, A. S., Naylor, B. A., Treacher, D. F., & Turner, R. C. (1985). Homeostasis model assessment: insulin resistance and β-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia*, *28*(7), 412-419.

McAllister, J. M., Legro, R. S., Modi, B. P., & Strauss III, J. F. (2015). Functional genomics of PCOS: from GWAS to molecular mechanisms. *Trends in Endocrinology & Metabolism*, *26*(3), 118-124.

McArdle, W. D., Katch, F. I., & Katch, V. L. (2010). *Exercise Physiology: Nutrition, Energy, and Human Performance* (8th ed). Lippincott Williams & Wilkins.

McAuley, E., & Blissmer, B. (2000). Self-efficacy determinants and consequences of physical activity. *Exercise and Sport Sciences Reviews*, *28*(2), 85-88.

McAuley, E., & Jacobson, L. (1991). Self-Efficacy and Exercise Participation in Sedentary Adult Females. *American Journal of Health Promotion*, *5*(3), 185-207.

McAuley, E., Lox, C., & Duncan, T. E. (1993). Long-term maintenance of exercise, self-efficacy, and physiological change in older adults. *Journal of Gerontology*, *48*(4), P218-P224.

McAuley, W. (1990) *Self-efficacy measures*. Unpublished raw data.

McBreairty, L., Zello, G., Rooke, J., Serrao, S., Pierson, R., Chizen, D., & Chilibeck, P. (2015). Long-term effect of a pulse-based diet and exercise training intervention on body composition and dietary intake in women with polycystic ovarian syndrome. *The FASEB Journal*, *29*(S1), 912-915.

McCook, J.G., Reame, N.E. and Thatcher, S.S., 2005. Health-related quality of life issues in women with polycystic ovary syndrome. *Journal of Obstetric, Gynecologic, & Neonatal Nursing*, *34*(1), 12-20.

McKercher, C. M., Schmidt, M. D., Sanderson, K. A., Patton, G. C., Dwyer, T., & Venn, A. J. (2009). Physical activity and depression in young adults. *American Journal of Preventive Medicine*, *36*(2), 161-164.

Mehrabani, H. H., Salehpour, S., Amiri, Z., Farahani, S. J., Meyer, B. J., & Tahbaz, F. (2012). Beneficial effects of a high-protein, low-glycemic-load hypocaloric diet in overweight and obese women with polycystic ovary syndrome: a randomized controlled intervention study. *Journal of the American College of Nutrition*, *31*(2), 117-125.

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

Mehrabian, F., Khani, B., Kelishadi, R., & Kermani, N. (2011). The prevalence of metabolic syndrome and insulin resistance according to the phenotypic subgroups of polycystic ovary syndrome in a representative sample of Iranian females. *Journal of Research in Medical Sciences: The Official Journal of Isfahan University of Medical Sciences*, *16*(6), 763.

Miller, C. T., and Downey, K. T. (1999) A meta–analysis of heavyweight and self–esteem. *Journal of Personality and Social Psychology Reviews*. *3*(1):68-84

Milligan, G. W., & Cooper, M. C. (1988). A study of standardization of variables in cluster analysis. *Journal of Classification*, *5*(2), 181-204.

Milte, C. M., Thorpe, M. G., Crawford, D., Ball, K., & McNaughton, S. A. (2015). Associations of diet quality with health-related quality of life in older Australian men and women. *Experimental Gerontology*, *64*, 8-16.

Mishra, S.I., Scherer, R.W., Geigle, P.M., Berlanstein, D.R., Topaloglu, O., Gotay, C.C. and Snyder, C., (2012). Exercise interventions on health-related quality of life for cancer survivors. *Cochrane Database of Systematic Reviews*, (8).

Mishra, S.I., Scherer, R.W., Snyder, C., Geigle, P.M., Berlanstein, D.R. and Topaloglu, O., (2012). Exercise interventions on health-related quality of life for people with cancer during active treatment. *Cochrane Database of Systematic Reviews*, (8).

Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., & Stroup, D. F. (2000). Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Oncology Research and Treatment*, *23*(6), 597-602.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine*, *6*(7), e1000097.

Moody, A., & Neave, A. (2016). *Health Survey for England 2015 Adult overweight and obesity*. Leeds, UK: Health and Social Care Information Centre.

Moon, T. K. (1996). The expectation-maximization algorithm. *IEEE Signal Processing Magazine*, *13*(6), 47-60.

Moore, S. C., Patel, A. V., Matthews, C. E., de Gonzalez, A. B., Park, Y., Katki, H. A., ... & Thun, M. (2012). Leisure time physical activity of moderate to vigorous intensity and mortality: a large pooled cohort analysis. *PLoS Medicine*, *9*(11), e1001335.

Moran, L. J., Hutchison, S. K., Norman, R. J., & Teede, H. J. (2011). Lifestyle changes in women with polycystic ovary syndrome. *Cochrane Database of Systematic Reviews*, (7).

Moran, L. J., Noakes, M., Clifton, P. M., Tomlinson, L., & Norman, R. J. (2003). Dietary composition in restoring reproductive and metabolic physiology in overweight women with polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, *88*(2), 812-819.

Moran, L. J., Noakes, M., Clifton, P. M., Wittert, G. A., Williams, G., & Norman, R. J. (2006). Short-term meal replacements followed by dietary macronutrient restriction enhance weight loss in polycystic ovary syndrome. *The American Journal of Clinical Nutrition*, *84*(1), 77-87.

Moran, L. J., Noakes, M., Norman, R. J., Clifton, P. M., & Tomlinson, L. (2002). Dietary composition and weight loss in improving reproductive and metabolic physiology in overweight women with polycystic ovary syndrome. *Asia Pacific Journal of Clinical Nutrition*, *11*, S271-S271.

Moran, L. J., Ranasinha, S., Zoungas, S., McNaughton, S. A., Brown, W. J., & Teede, H. J. (2013). The contribution of diet, physical activity and sedentary behaviour to body mass index in women with and without polycystic ovary syndrome. *Human Reproduction*, *28*(8), 2276-2283.

Moran, L., & Teede, H. (2009). Metabolic features of the reproductive phenotypes of polycystic ovary syndrome. *Human Reproduction Update*, *15*(4), 477-488.

Morin-Papunen, L., Vauhkonen, I., Koivunen, R., Ruokonen, A., Martikainen, H., & Tapanainen, J. S. (2003). Metformin versus ethinyl estradiol-cyproterone acetate in the treatment of nonobese women with polycystic ovary syndrome: a randomized study. *The Journal of Clinical Endocrinology & Metabolism*, *88*(1), 148-156.

Morris, J. N., Heady, J. A., Raffle, P. A. B., Roberts, C. G., & Parks, J. W. (1953). Coronary heart-disease and physical activity of work. *The Lancet*, *262*(6796), 1111-1120.

Nasrekani, Z., & Fathi, M. (2016). Efficacy of 12 weeks aerobic training on body composition, aerobic power and some women-hormones in polycystic ovary syndrome infertile women. *The Iranian Journal of Obstetrics, Gynecology and Infertility*, *19*(5), 1-10.

Nathan, D. M., Turgeon, H., & Regan, S. (2007). Relationship between glycated haemoglobin levels and mean glucose levels over time. *Diabetologia*, *50*(11), 2239-2244.

National Cholesterol Education Program (1993). Second Report of the Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. *NIH Publication*, *93*, 3095.

National Cholesterol Education Program (US). Expert Panel on Detection, & Treatment of High Blood Cholesterol in Adults. (2005). *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)* (No. 2). National Cholesterol Education Program, National Heart, Lung, and Blood Institute, National Institutes of Health.

National Institute for Health and Care Excellence (2013). Scenario: Management of Polycystic Ovary Syndrome [online]. Available at: https://cks.nice.org.uk/polycystic-ovary-syndrome#!scenariorecommendation:5 [Last accessed 6th November 2020].

National Institute of Health (2014) *Quality Assessment of Case-Control Studies* [online] Available at: https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools. [Last accessed 16th November 2020].

National Institutes of Health. (2012). Evidence-based methodology workshop on polycystic ovary syndrome, December 3-5. Executive summary. [online] Available at: https://prevention.nih.gov/docs/programs/pcos/FinalReport.pdf. [Last accessed November 6th 2020].

Nawaz, H., Chan, W., Abdulrahman, M., Larson, D., Katz, D.L. (2001). Self-reported weight and height: implications for obesity research. *American Journal of Preventative Medicine, 20*, 294-298.

Nestler, J. E. (2008). Metformin for the treatment of the polycystic ovary syndrome. *New England Journal of Medicine*, *358*(1), 47-54.

Nestler, J. E., Jakubowicz, D. J., Evans, W. S., & Pasquali, R. (1998). Effects of metformin on spontaneous and clomiphene-induced ovulation in the polycystic ovary syndrome. *New England Journal of Medicine*, *338*(26), 1876-1880.

Newman, D. A. (2014). Missing data: Five practical guidelines. *Organizational Research Methods*, *17*(4), 372-411.

Ng, E. H. Y., Chan, C. C. W., and Ho, P. C. (2007). Polycystic ovary syndrome in Asian women. In, Kovacs, G. T., and Norman, R. (eds). *Polycystic ovary syndrome*, 2nd ed. Cambridge, UK: Cambridge University Press. pp. 316-330.

Ng, E. H. Y., Chan, C. C. W., Yeung, W. S. B., & Ho, P. C. (2004). Effect of age on ovarian stromal flow measured by three-dimensional ultrasound with power Doppler in Chinese women with proven fertility. *Human reproduction*, *19*(9), 2132-2137.

Ng, E. H. Y., Wat, N. M. S., & Ho, P. C. (2001). Effects of metformin on ovulation rate, hormonal and metabolic profiles in women with clomiphene-resistant polycystic ovaries: a randomized, double-blinded placebo-controlled trial. *Human Reproduction*, *16*(8), 1625-1631.

Ng, N. Y. H., Jiang, G., Cheung, L. P., Zhang, Y., Tam, C. H. T., Luk, A. O. Y., ... & Ma, R. C. W. (2019). Progression of glucose intolerance and cardiometabolic risk factors over a decade in Chinese women with polycystic ovary syndrome: A case-control study. *PLoS Medicine*, *16*(10), e1002953.

NICE (2014) Obesity: identification, assessment and management: Clinical guideline CG189 [online] https://www.nice.org.uk/guidance/cg189/ifp/chapter/Assessing-your-weight. [Last accessed: 18th November 2020].

Nidhi, R., Padmalatha, V., Nagarathna, R., & Amritanshu, R. (2013). Effect of yoga program on quality of life in adolescent polycystic ovarian syndrome: a randomized control trial. *Applied Research in Quality of Life*, *8*(3), 373-383.

Nidhi, R., Padmalatha, V., Nagarathna, R., & Amritanshu, R. (2013). Effects of a holistic yoga program on endocrine parameters in adolescents with polycystic ovarian syndrome: a randomized controlled trial. *The Journal of Alternative and Complementary Medicine*, *19*(2), 153-160.

Nidhi, R., Padmalatha, V., Nagarathna, R., & Ram, A. (2012). Effect of a yoga program on glucose metabolism and blood lipid levels in adolescent girls with polycystic ovary syndrome. *International Journal of Gynecology & Obstetrics*, *118*(1), 37-41.

Nieuwenhuis-Ruifrok, A. E., Kuchenbecker, W. K. H., Hoek, A., Middleton, P., & Norman, R. J. (2009). Insulin sensitizing drugs for weight loss in women of reproductive age who are overweight or obese: systematic review and meta-analysis. *Human Reproduction Update*, *15*(1), 57-68.

Norman, R. J., Masters, L., Milner, C. R., Wang, J. X., & Davies, M. J. (2001). Relative risk of conversion from normoglycaemia to impaired glucose tolerance or non-insulin dependent diabetes mellitus in polycystic ovarian syndrome. *Human Reproduction*, *16*(9), 1995-1998.

Nussdorfer, G. G., Mazzocchi, G., & Meneghelli, V. (1978). Cytophysiology of the adrenal zona fasciculata. *International Review of Cytology*, *55*, 291-365).

Nuttall, F. Q. (2015). Body mass index: obesity, BMI, and health: a critical review. *Nutrition Today*, *50*(3), 117.

Nybacka, Å., Carlström, K., Fabri, F., Hellström, P. M., & Hirschberg, A. L. (2013). Serum antimüllerian hormone in response to dietary management and/or physical exercise in overweight/obese women with polycystic ovary syndrome: secondary analysis of a randomized controlled trial. *Fertility and Sterility*, *100*(4), 1096-1102.

Nybacka, Å., Carlström, K., Ståhle, A., Nyrén, S., Hellström, P. M., & Hirschberg, A. L. (2011). Randomized comparison of the influence of dietary management and/or physical exercise on ovarian function and metabolic parameters in overweight women with polycystic ovary syndrome. *Fertility and Sterility*, *96*(6), 1508-1513.

Nybacka, Å., Fabri, F., Hellström, P. M., Hirschberg, A. L., & Stahle, A. (2012). LB011-Sun Plasma Anti-mullerian Hormone (AMH) Levels Correlate and Predict the Response to Lifestyle Intervention in Polycystic Ovarian Syndrome. *Clinical Nutrition Supplements*, *1*(7), 255.

Nybacka, Å., Hellström, P. M., & Hirachberg, A. L. (2014). LB021-SUN: Nutritional Counseling Significantly Improves Metabolic Variables in Overweight/Obese Women with Polycystic Ovary Syndrome. *Clinical Nutrition*, *33*, S246-S247.

Olefsky, J. M., & Kolterman, O. G. (1981). Mechanisms of insulin resistance in obesity and noninsulin-dependent (type II) diabetes. *The American Journal of Medicine*, *70*(1), 151-168.

Olefsky, J., Farquhar, J. W., & Reaven, G. (1973). Relationship between fasting plasma insulin level and resistance to insulin-mediated glucose uptake in normal and diabetic subjects. *Diabetes*, *22*(7), 507-513.

Oman, R. F., & King, A. C. (1998). Predicting the adoption and maintenance of exercise participation using self-efficacy and previous exercise participation rates. *American Journal of Health Promotion*, *12*(3), 154-161.

Omar, A. N., Habib, F., Mojaddidi, M., & El-Bab, M. F. (2013). Body weight reduction and metformin: Roles in polycystic ovary syndrome. *Pathophysiology*, *20*(2), 131-137.

Orio Jr, F., Giallauria, F., Palomba, S., Cascella, T., Manguso, F., Vuolo, L., ... & Vigorito, C. (2006). Cardiopulmonary impairment in young women with polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, *91*(8), 2967-2971.

Orio Jr, F., Palomba, S., Spinelli, L., Cascella, T., Tauchmanovà, L., Zullo, F., ... & Colao, A. (2004). The cardiovascular risk of young women with polycystic ovary syndrome: an observational, analytical, prospective case-control study. *The Journal of Clinical Endocrinology & Metabolism*, *89*(8), 3696-3701.

Orio, F., Muscogiuri, G., Giallauria, F., Savastano, S., Bottiglieri, P., Tafuri, D., ... & Palomba, S. (2016). Oral contraceptives versus physical exercise on cardiovascular and metabolic risk factors in women with polycystic ovary syndrome: a randomized controlled trial. *Clinical Endocrinology*, *85*(5), 764-771.

Ornstein, R. M., Copperman, N. M., & Jacobson, M. S. (2011). Effect of weight loss on menstrual function in adolescents with polycystic ovary syndrome. *Journal of Pediatric and Adolescent Gynecology*, *24*(3), 161-165.

Orth, U., and Robins, R. W. (2014). The development of self-esteem. *Current Directions in Psychological Science*, *23*, 381-387

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

Padwal, R., Li, S. K., & Lau, D. C. W. (2003). Long-term pharmacotherapy for overweight and obesity: a systematic review and meta-analysis of randomized controlled trials. *International Journal of Obesity*, *27*(12), 1437-1446.

Paffenbarger Jr, R. S., Hyde, R., Wing, A. L., & Hsieh, C. C. (1986). Physical activity, all-cause mortality, and longevity of college alumni. *New England Journal of Medicine*, *314*(10), 605-613.

Paffenbarger Jr, R. S., Laughlin, M. E., Gima, A. S., & Black, R. A. (1970). Work activity of longshoremen as related to death from coronary heart disease and stroke. *New England Journal of Medicine*, *282*(20), 1109-1114.

Palomba, S., Falbo, A., Giallauria, F., Russo, T., Rocca, M., Tolino, A., ... & Orio, F. (2010). Six weeks of structured exercise training and hypocaloric diet increases the probability of ovulation after clomiphene citrate in overweight and obese patients with polycystic ovary syndrome: a randomized controlled trial. *Human Reproduction*, *25*(11), 2783-2791.

Palomba, S., Giallauria, F., Falbo, A., Russo, T., Oppedisano, R., Tolino, A., ... & Orio, F. (2008). Structured exercise training programme versus hypocaloric hyperproteic diet in obese polycystic ovary syndrome patients with anovulatory infertility: a 24-week pilot study. *Human Reproduction*, *23*(3), 642-650.

Papakonstantinou, E., Kechribari, I., Mitrou, P., Trakakis, E., Vassiliadi, D., Georgousopoulou, E., ... & Dimitriadis, G. (2016). Effect of meal frequency on glucose and insulin levels in women with polycystic ovary syndrome: a randomised trial. *European Journal of Clinical Nutrition*, *70*(5), 588-594.

Parke, C. (2013). Module 5: Identifying and addressing outliers. *Essential First Steps to Data Analysis: Scenario-based Examples Using SPSS*, 81-101.

Parker, S. L., Godil, S. S., Shau, D. N., Mendenhall, S. K., & McGirt, M. J. (2013). Assessment of the minimum clinically important difference in pain, disability, and quality of life after anterior cervical discectomy and fusion. *Journal of Neurosurgery: Spine, 18(2), 154–160.*

Pasquali, R., Casimirri, F., Venturoli, S., Antonio, M., Morselli, L., Reho, S., ... & Paradisi, R. (1994). Body fat distribution has weight-independent effects on clinical, hormonal, and metabolic features of women with polycystic ovary syndrome. *Metabolism*, *43*(6), 706-713.

Pasquali, R., Fabbri, R., Venturoli, S., Paradisi, R., Antenucci, D., & Melchionda, N. (1986). Effect of weight loss and antiandrogenic therapy on sex hormone blood levels and insulin resistance in obese patients with polycystic ovaries. *American Journal of Obstetrics and Gynecology*, *154*(1), 139-144.

Pasquali, R., Gambineri, A., Biscotti, D., Vicennati, V., Gagliardi, L., Colitta, D., ... & Morselli-Labate, A. M. (2000). Effect of long-term treatment with metformin added to hypocaloric diet on body composition, fat distribution, and androgen and insulin levels in abdominally obese women with and without the polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, *85*(8), 2767-2774.

Pasquali, R. (2006). Obesity and androgens: facts and perspectives. *Fertility and Sterility*, *85*(5), 1319-1340.

Pasquali, R., Vicennati, V., Bertazzo, D., Casimirri, F., Pascal, G., Tortelli, O., & Labate, A. M. M. (1997). Determinants of sex hormone—binding globulin blood concentrations in premenopausal and postmenopausal women with different estrogen status. *Metabolism*, *46*(1), 5-9.

Pastore, L. M., Patrie, J. T., Morris, W. L., Dalal, P., & Bray, M. J. (2011). Depression symptoms and body dissatisfaction association among polycystic ovary syndrome women. *Journal of Psychosomatic Research*, *71*(4), 270-276.

Pattyn, N., Cornelissen, V. A., Eshghi, S. R., & Vanhees, L. (2013). The effect of exercise on the cardiovascular risk factors constituting the metabolic syndrome: a meta-analysis of controlled trials. *Sports Medicine*, *43*(2), 121-133.

Pehlivanov, B., & Orbetzova, M. (2007). Characteristics of different phenotypes of polycystic ovary syndrome in a Bulgarian population. *Gynecological Endocrinology*, *23*(10), 604-609.

Pentecost, C., & Taket, A. (2011). Understanding exercise uptake and adherence for people with chronic conditions: a new model demonstrating the importance of exercise identity, benefits of attending and support. *Health Education Research*, *26*(5), 908-922.

Peppard, H. R., Marfori, J., Iuorno, M. J., & Nestler, J. E. (2001). Prevalence of polycystic ovary syndrome among premenopausal women with type 2 diabetes. *Diabetes Care*, *24*(6), 1050-1052.

Petrányi, G., & Zaoura-Petrányi, M. (2011). Metformin treatment with or without life style changes in patients with polycystic ovary syndrome. *Orvosi Hetilap*, *152*(16), 628-632.

Phillips, D. I. W., Clark, P. M., Hales, C. N., & Osmond, C. (1994). Understanding oral glucose tolerance: comparison of glucose or insulin measurements during the oral glucose tolerance test with specific measurements of insulin resistance and insulin secretion. *Diabetic Medicine*, *11*(3), 286-292.

Piouka, A., Farmakiotis, D., Katsikis, I., Macut, D., Gerou, S., & Panidis, D. (2009). Anti-Mullerian hormone levels reflect severity of PCOS but are negatively influenced by obesity: relationship with increased luteinizing hormone levels. *American Journal of Physiology-Endocrinology and Metabolism*, *296*(2), E238-E243.

Pirwany, I. R., Fleming, R., Greer, I. A., Packard, C. J., & Sattar, N. (2001). Lipids and lipoprotein subfractions in women with PCOS: relationship to metabolic and endocrine parameters. *Clinical Endocrinology*, *54*(4), 447-453.

Polson, D. W., Wadsworth, J., Adams, J., & Franks, S. (1988). Polycystic ovaries—a common finding in normal women. *The Lancet*, *331*(8590), 870-872.

Popova, P., Ryasantseva, E., Zazerskaya, I., Ivanova, L., & Grineva, E. (2010). The role of weight loss and metformin in the improvement of menstrual function in overweight and obese women with polycystic ovary syndrome: PP090. *Diabetes, Obesity and Metabolism*, *12*.

Poretsky, L., Cataldo, N. A., Rosenwaks, Z., and Giudice, L. C. (1999). The insulin-related ovarian regulatory system in health and disease. *Endocrine Reviews*, *20*(4), 535-582.

Prince, S. A., Adamo, K. B., Hamel, M. E., Hardt, J., Gorber, S. C., & Tremblay, M. (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, *5*(1), 56.

Prince, S. A., Butler, G. P., Roberts, K. C., Lapointe, P., MacKenzie, A. M., Colley, R. C., ... & Thompson, W. (2019). Developing content for national population health surveys: an example using a newly developed sedentary behaviour module. *Archives of Public Health*, *77*(1), 53.

Prince, S. A., Cardilli, L., Reed, J. L., Saunders, T. J., Kite, C., Douillette, K., ... & Buckley, J. P. (2020). A comparison of self-reported and device measured sedentary behaviour in adults: a systematic review and meta-analysis. *International Journal of Behavioral Nutrition and Physical Activity*, *17*(1), 1-17.

Puska, P. (2010). From Framingham to North Karelia: from descriptive epidemiology to public health action. *Progress in Cardiovascular Diseases*, *53*(1), 15-20.

Randeva, H. S., Lewandowski, K. C., Drzewoski, J., Brooke-Wavell, K., O'Callaghan, C., Czupryniak, L., ... & Prelevic, G. M. (2002). Exercise decreases plasma total homocysteine in overweight young women with polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, *87*(10), 4496-4501.

Raschke, S., & Eckel, J. (2013). Adipo-myokines: two sides of the same coin—mediators of inflammation and mediators of exercise. *Mediators of Inflammation*, *2013*.

Rausch, M. E., Legro, R. S., Barnhart, H. X., Schlaff, W. D., Carr, B. R., Diamond, M. P., ... & Gosman, G. G. (2009). Predictors of pregnancy in women with polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, *94*(9), 3458-3466.

Rea, L. M., & Parker, R. A. (1992). *Designing and conducting survey research: a comprehensive guide*. Jossey-Bass Publishers: San Francisco

Redman, L. M., Elkind-Hirsch, K., & Ravussin, E. (2011). Aerobic exercise in women with polycystic ovary syndrome improves ovarian morphology independent of changes in body composition. *Fertility and Sterility*, *95*(8), 2696-2699.

Reiche, E. M. V., Nunes, S. O. V., & Morimoto, H. K. (2004). Stress, depression, the immune system, and cancer. *The Lancet Oncology*, *5*(10), 617-625.

Reith, C., Landray, M., Devereaux, P. J., Bosch, J., Granger, C. B., Baigent, C., Califf, R. M., Collins, R. & Yusuf, S. (2013). Randomized clinical trials--removing unnecessary obstacles. *The New England Journal of Medicine*, *369*(11), 1061.

Reitzes, D. C., & Mutran, E. J. (2006). Self and health: Factors that encourage self-esteem and functional health. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *61*(1), S44-S51.

Resnick, B. and Jenkins, L.S. (2000). Testing the reliability and validity of the self-efficacy for exercise scale. *Nursing Research*, *49*(3), 154-159.

Resnick, B., & Spellbring, A. (2000). Understanding what motivates older adults to exercise. *Journal of Gerontological Nursing, 26,* 17-21.

Ridderstråle, M., Evans, L. M., Jensen, H. H., Bøgelund, M., Jensen, M. M., Ericsson, Å., & Jendle, J. (2016). Estimating the impact of changes in HbA 1c, body weight and insulin injection regimen

on health-related quality-of-life: a time trade off study. *Health and Quality of Life Outcomes*, *14*(1), 13.

Riley, R. D., Higgins, J. P., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. *BMJ*, *342*.

Rippe, J. M., Price, J. M., Hess, S. A., Kline, G., DeMers, K. A., Damitz, S., ... & Freedson, P. (1998). Improved psychological well-being, quality of life, and health practices in moderately overweight women participating in a 12-week structured weight loss program. *Obesity Research*, *6*(3), 208-218.

Robinson, S., Chan, S. P., Spacey, S., Anyaoku, V., Johnston, D. G., & Franks, S. (1992). Postprandial thermogenesis is reduced in polycystic ovary syndrome and is associated with increased insulin resistance. *Clinical Endocrinology*, *36*(6), 537-543.

Robinson, S., Kiddy, D., Gelding, S. V., Willis, D., Niththyananthan, R., Bush, A., ... & Franks, S. (1993). The relationship of insulin insensitivity to menstrual pattern in women with hyperandrogenism and polycystic ovaries. *Clinical Endocrinology*, *39*(3), 351-355.

Rodgers, M. M., & Cavanagh, P. R. (1984). Glossary of biomechanical terms, concepts, and units. *Physical Therapy*, *64*(12), 1886-1902.

Rodino, I. S., Byrne, S., & Sanders, K. A. (2016). Obesity and psychological wellbeing in patients undergoing fertility treatment. *Reproductive Biomedicine Online*, *32*(1), 104-112.

Roessler, K. K., Andersen, M., Glintborg, D., Ravn, P., & Birkebaek, C. (2011). F417 Group Counselling and Physical Activity of Patients with Polycystic Ovarian Syndrome (PCOS). *European Journal of Pain Supplements*, *5*(S1), 156-156.

Roessler, K. K., Birkebaek, C., Ravn, P., Andersen, M. S., & Glintborg, D. (2013). Effects of exercise and group counselling on body composition and VO 2max in overweight women with polycystic ovary syndrome. *Acta Obstetricia et Gynecologica Scandinavica*, *92*(3), 272-277.

Romesburg, H. C. (1984). *Cluster analysis for researchers*. Wadsworth. Inc: Belmont, CA.

Rosenbaum, D., Haber, R. S., & Dunaif, A. (1993). Insulin resistance in polycystic ovary syndrome: decreased expression of GLUT-4 glucose transporters in adipocytes. *American Journal of Physiology-Endocrinology and Metabolism*, *264*(2), E197-E202.

Rosenstock, I. M., Strecher, V. J., & Becker, M. H. (1988). Social learning theory and the health belief model. *Health Education Quarterly*, *15*(2), 175-183.

Rowland, M. L. (1990). Self-reported weight and height. *The American Journal of Clinical Nutrition*, *52*(6), 1125-1133.

Rubin, D. B. (1976). Inference and Missing Data. *Biometrika 63*(3), 581-90.

Sá, J. C., Costa, E. C., Da Silva, E., Tamburús, N. Y., Porta, A., Medeiros, L. F., ... & Azevedo, G. D. (2016). Aerobic exercise improves cardiac autonomic modulation in women with polycystic ovary syndrome. *International Journal of Cardiology*, *202*, 356-361.

Sadaria, R. G., & Ravi, B. V. (2015). A study on assessment of testosterone, insulin resistance and HbA1c in women with polycystic ovarian syndrome. *International Journal of Health Sciences and Research*, *5*, 180-185.

Safita, N., Islam, S. M. S., Chow, C. K., Niessen, L., Lechner, A., Holle, R., & Laxy, M. (2016). The impact of type 2 diabetes on health-related quality of life in Bangladesh: results from a matched study comparing treated cases with non-diabetic controls. *Health and Quality of Life Outcomes*, *14*(1), 129.

Saidj, M., Jørgensen, T., Jacobsen, R. K., Linneberg, A., Oppert, J. M., & Aadahl, M. (2016). Work and leisure time sitting and inactivity: effects on cardiorespiratory and metabolic health. *European Journal of Preventive Cardiology*, *23*(12), 1321-1329.

Salazar, M. R., Carbajal, H. A., Espeche, W. G., Aizpurua, M., Leiva Sisnieguez, C. E., March, C. E., ... & Reaven, G. M. (2013). Identifying cardiovascular disease risk and outcome: use of the plasma triglyceride/high-density lipoprotein cholesterol concentration ratio versus metabolic syndrome criteria. *Journal of Internal Medicine*, *273*(6), 595-601

Salazar, M. R., Carbajal, H. A., Espeche, W. G., Sisnieguez, C. E. L., Balbín, E., Dulbecco, C. A., ... & Reaven, G. M. (2012). Relation among the plasma triglyceride/high-density lipoprotein cholesterol concentration ratio, insulin resistance, and associated cardio-metabolic risk factors in men and women. *The American Journal of Cardiology*, *109*(12), 1749-1753.

Sallis, J. F., & Saelens, B. E. (2000). Assessment of physical activity by self-report: status, limitations, and future directions. *Research Quarterly for Exercise and Sport*, *71*(S2), 1-14.

Sallis, J. F., Haskell, W. L., Fortmann, S. P., Vranizan, K. M., Taylor, C. B., & Solomon, D. S. (1986). Predictors of adoption and maintenance of physical activity in a community sample. *Preventive Medicine*, *15*(4), 331-341.

Sam, S. (2007). Obesity and polycystic ovary syndrome. *Obesity Management*, *3*(2), 69-73.

Samitz, G., Egger, M., & Zwahlen, M. (2011). Domains of physical activity and all-cause mortality: systematic review and dose–response meta-analysis of cohort studies. *International Journal of Epidemiology*, *40*(5), 1382-1400.

Santos, A. C., & Barros, H. (2003). Prevalence and determinants of obesity in an urban sample of Portuguese adults. *Public Health*, *117*(6), 430-437.

Santos, I. K., Oliveira Maranhão, T. M., Soares, G. M. and Dantas, P. M. S. (2017), Exercise and reproductive function in polycystic ovary syndrome: a systematic review. PROSPERO 2017: CRD42017058869 Available from: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42017058869.

Saremi, A., & Yaghoubi, M. S. (2016). Effect of resistance exercises with calcium consumption on level of anti-mullerian hormone and some metabolic indices in women with polycystic ovarian syndrome. *The Iranian Journal of Obstetrics, Gynecology and Infertility*, *18*(180), 7-15.

Saremi, A., Shavandi, N., Karamali, M., & Kazemi, M. (2013). Serum level of anti-mullerian hormone after exercise training in women with polycystic ovary syndrome: a randomized controlled trial. *The Iranian Journal of Obstetrics, Gynecology and Infertility*, *16*(64), 10-18.

Sarwono, J. (2018) *Path Analysis Using AMOS*. Washington, USA: Amazon.com, Inc.

Sathisha, T. G., Bhargavi, S. K., Riyaz Ahmad Kalaburgi, S. H., Kavitha, B. S., & Arifuddin, M. K. Study of Biochemical Profile of HbA1c, Vitamin D and AMH in Infertile Women of a Population of South India. *Scholars International Journal of Biochemistry,* 122-126.

Saxton, J. (2011). *Exercise and Chronic Disease.* Routledge: Abingdon, UK.

Schafer, J. L., Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147–177.

Schmidt, J., Landin-Wilhelmsen, K., Brännström, M., & Dahlgren, E. (2011). Cardiovascular disease and risk factors in PCOS women of postmenopausal age: a 21-year controlled follow-up study. *The Journal of Clinical Endocrinology & Metabolism*, *96*(12), 3794-3803.

Schröder, H., Marrugat, J., Vila, J., Covas, M. I., & Elosua, R. (2004). Adherence to the traditional Mediterranean diet is inversely associated with body mass index and obesity in a Spanish population. *The Journal of Nutrition*, *134*(12), 3355-3361.

Schünemann, H., Brożek, J., Guyatt, G., and Oxman, A. (2013). *GRADE Handbook for Grading Quality of Evidence and Strength of Recommendations.* The GRADE Working Group, 2013.

Schwarzer, R., Luszczynska, A., Ziegelmann, J. P., Scholz, U., & Lippke, S. (2008). Social-cognitive predictors of physical exercise adherence: Three longitudinal studies in rehabilitation. *Health Psychology, 27*(1), 54–63.

Sechrist, K.R., Walker, S.N. and Pender, N.J. (1987*b*). Development and psychometric evaluation of the exercise benefits/barriers scale. *Research in Nursing & Health, 10*(6), 357-365.

Sechrist, K.R., Walker, S.N. and Pender, N.J., (1987*a*). Health promotion model-instruments to measure HPM behavioral determinants: exercise benefits/barriers scale [EBBS] (Adult Version). [online] available at: https://deepblue.lib.umich.edu/bitstream/handle/2027.42/8535 4 /EBBS-English_Version.pdf?sequence=2. [Accessed November 10th 2020].

Seow, K. M., Juan, C. C., Hwang, J. L., & Ho, L. T. (2008, January). Laparoscopic surgery in polycystic ovary syndrome: reproductive and metabolic effects. *Seminars in Reproductive Medicine, 26*(1), 101-110).

Shapiro, S. S., Wilk, M. B., & Chen, H. J. (1968). A comparative study of various tests for normality. *Journal of the American Statistical Association*, *63*(324), 1343-1372.

Sills, E. S., Perloe, M., Tucker, M. J., Kaplan, C. R., Genton, M. G., & Schattman, G. L. (2001). Diagnostic and treatment characteristics of polycystic ovary syndrome: descriptive measurements of patient perception and awareness from 657 confidential self-reports. *BMC Women's Health, 1*(1), 3.

Silva Dantas, W., Antonio Miguel Marcondes, J., Katsuyuki Shinjo, S., Augusto Perandini, L., Olzon Zambelli, V., Das Neves, W., ... & Hisashi Murai, I. (2015). GLUT4 translocation is not impaired after acute exercise in skeletal muscle of women with obesity and polycystic ovary syndrome. *Obesity*, *23*(11), 2207-2215.

Singer, D. E., Nathan, D. M., Anderson, K. M., Wilson, P. W., & Evans, J. C. (1992). Association of HbA1c with prevalent cardiovascular disease in the original cohort of the Framingham Heart Study. *Diabetes*, *41*(2), 202-208.

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

Sir-Petermann, T., Codner, E., Pérez, V., Echiburú, B., Maliqueo, M., Ladrón de Guevara, A., Preisler, J., Crisosto, N., Sánchez, F., Cassorla, F., Bhasin, S. (2009). Metabolic and reproductive features before and during puberty in daughters of women with polycystic ovary syndrome. *Journal of Clinical Endocrinology and Metabolism*. *94,* 1923-1930.

Sivitz, W. I., DeSautel, S. L., Kayano, T., Bell, G. I., & Pessin, J. E. (1989). Regulation of glucose transporter messenger RNA in insulin-deficient states. *Nature*, *340*(6228), 72-74.

Skidmore, S. T., & Thompson, B. (2013). Bias and precision of some classical ANOVA effect sizes when assumptions are violated. *Behavior Research Methods*, *45*(2), 536-546.

Sniderman, A. D., Pedersen, T., & Kjekshus, J. (1997). Putting low-density lipoproteins at center stage in atherogenesis. *The American Journal of Cardiology*, *79*(1), 64-67.

Solomon, C. G., Hu, F. B., Dunaif, A., Rich-Edwards, J. E., Stampfer, M. J., Willett, W. C., ... & Manson, J. E. (2002). Menstrual cycle irregularity and risk for future cardiovascular disease. *The Journal of Clinical Endocrinology & Metabolism*, *87*(5), 2013-2017.

Sonstroem, R. J. (1997). Physical activity and self-esteem. In, Morgan, W. P. (Ed). *Physical activity mental health*. Washington DC, USA: Hemisphere, pp. 128-143.

Sørensen, L. B., Søe, M., Halkier, K. H., Stigsby, B., & Astrup, A. (2012). Effects of increased dietary protein-to-carbohydrate ratios in women with polycystic ovary syndrome. *The American Journal of Clinical Nutrition*, *95*(1), 39-48.

Spence, J. C., McGannon, K. R., & Poon, P. (2005). The effect of exercise on global self-esteem: A quantitative review. *Journal of Sport and Exercise Psychology*, *27*(3), 311-334.

Stein, I. F., and Leventhal, M. L. (1935). Amenorrhea associated with bilateral polycystic ovaries. *American Journal of Obstetrics and Gynecology*, *29*, 181-189.

Steinbach, M., Kumar, V., & Tan, P. (2005). Cluster analysis: basic concepts and algorithms. *Introduction to data mining (*1st ed*).* Pearson Addison Wesley.

Stener-Victorin, E., Baghaei, F., Holm, G., Janson, P. O., Olivecrona, G., Lönn, M., & Mannerås-Holm, L. (2012). Effects of acupuncture and exercise on insulin sensitivity, adipose tissue characteristics, and markers of coagulation and fibrinolysis in women with polycystic ovary syndrome: secondary analyses of a randomized controlled trial. *Fertility and Sterility*, *97*(2), 501-508.

Stener-Victorin, E., Holm, G., Janson, P. O., Gustafson, D., & Waern, M. (2013). Acupuncture and physical exercise for affective symptoms and health-related quality of life in polycystic ovary syndrome: secondary analysis from a randomized controlled trial. *BMC Complementary and Alternative Medicine*, *13*(1), 131.

Stener-Victorin, E., Jedel, E., Janson, P. O., & Sverrisdottir, Y. B. (2009). Low-frequency electroacupuncture and physical exercise decrease high muscle sympathetic nerve activity in polycystic ovary syndrome. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *297*, R387-R395.

Stepto, N. K., Cassar, S., Joham, A. E., Hutchison, S. K., Harrison, C. L., Goldstein, R. F., & Teede, H. J. (2013). Women with polycystic ovary syndrome have intrinsic insulin resistance on euglycaemic–hyperinsulaemic clamp. *Human Reproduction*, *28*(3), 777-784.

Sterne, J. A. C., Gavaghan, D. and Egger, M. (2000). Publication and related bias I meta-analysis: Power of statistical tests and prevalence in the literature. *Journal of Clinical Epidemiology, 53,* 1119-1129.

Sterne, J. A., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., ... & Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ, 338,* 2393.

Strauss, R. (2000). Childhood. Obesity and Self-Esteem. *Pediatrics, 105,* E15.

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., ... & Liu, B. (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *Plos Medicine, 12*(3), e1001779.

Swenson, I., & Havens, B. (1987). Menarche and menstruation: a review of the literature. *Journal of Community Health Nursing, 4*(4), 199-210.

Swift, D. L., Johannsen, N. M., Tudor-Locke, C., Earnest, C. P., Johnson, W. D., Blair, S. N., ... & Church, T. S. (2012). Exercise training and habitual physical activity: a randomized controlled trial. *American Journal of Preventive Medicine, 43*(6), 629-635.

Taghavi, S.A., Bazarganipour, F., Montazeri, A., Kazemnejad, A., Chaman, R. and Khosravi, A., (2015). Health-related quality of life in polycystic ovary syndrome patients: A systematic review. *Iranian Journal of Reproductive Medicine, 13*(8), 473.

Talbott, E., Clerici, A., Berga, S. L., Kuller, L., Guzick, D., Detre, K., ... & Engberg, R. A. (1998). Adverse lipid and coronary heart disease risk profiles in young women with polycystic ovary syndrome: results of a case-control study. *Journal of Clinical Epidemiology, 51*(5), 415-422.

Talbott, E., Guzick, D., Clerici, A., Berga, S., Detre, K., Weimer, K., & Kuller, L. (1995). Coronary heart disease risk factors in women with polycystic ovary syndrome. *Arteriosclerosis, Thrombosis, and Vascular Biology, 15*(7), 821-826.

Tang, T., Glanville, J., Hayden, C. J., White, D., Barth, J. H., & Balen, A. H. (2006). Combined lifestyle modification and metformin in obese patients with polycystic ovary syndrome. A randomized, placebo-controlled, double-blind multicentre study. *Human Reproduction, 21*(1), 80-89.

Tang, T., Lord, J. M., Norman, R. J., Yasmin, E., & Balen, A. H. (2012). Insulin-sensitising drugs (metformin, rosiglitazone, pioglitazone, D-chiro-inositol) for women with polycystic ovary syndrome, oligo amenorrhoea and subfertility. *Cochrane Database of Systematic Reviews*, (5).

Taponen, S., Martikainen, H., Järvelin, M. R., Laitinen, J., Pouta, A., Hartikainen, A. L., ... & Ruokonen, A. (2003). Hormonal profile of women with self-reported symptoms of oligomenorrhea and/or hirsutism: Northern Finland birth cohort 1966 study. *The Journal of Clinical Endocrinology & Metabolism, 88*(1), 141-147.

Tarkun, I., Çetinarslan, B., Türemen, E., Sahin, T., Cantürk, Z., & Komsuoglu, B. (2005). Effect of rosiglitazone on insulin resistance, C-reactive protein and endothelial function in non-obese young women with polycystic ovary syndrome. *European Journal of Endocrinology, 153*(1), 115-121.

Tay, C. T., Teede, H. J., Hill, B., Loxton, D., & Joham, A. E. (2019). Increased prevalence of eating disorders, low self-esteem, and psychological distress in women with polycystic ovary syndrome: a community-based cohort study. *Fertility and Sterility*, *112*(2), 353-361.

Teede, H. J., Hutchison, S., Zoungas, S., & Meyer, C. (2006). Insulin resistance, the metabolic syndrome, diabetes, and cardiovascular disease risk in women with PCOS. *Endocrine*, *30*(1), 45-53.

Teede, H. J., Joham, A. E., Paul, E., Moran, L. J., Loxton, D., Jolley, D., & Lombard, C. (2013). Longitudinal weight gain in women identified with polycystic ovary syndrome: results of an observational study in young women. *Obesity*, *21*(8), 1526-1532.

Teede, H. J., Meyer, C., & Norman, R. J. (2005). Insulin-sensitisers in the treatment of polycystic ovary syndrome. *Expert Opinion on Pharmacotherapy*, *6*(14), 2419-2427.

Teede, H. J., Misso, M. L., Costello, M. F., Dokras, A., Laven, J., Moran, L., ... & Norman, R. J. (2018). Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *Human Reproduction*, *33*(9), 1602-1618.

Teede, H. J., Misso, M. L., Deeks, A. A., Moran, L. J., Stuckey, B. G., Wong, J. L., ... & Costello, M. F. (2011). Assessment and management of polycystic ovary syndrome: summary of an evidence-based guideline. *The Medical Journal of Australia*, *195*(6), S65.

Teede, H., Deeks, A., & Moran, L. (2010). 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., Misso, M., Costello, M., Dokras, A., Laven, J., Moran, L., Piltonen, T. and Norman, R. (2018) *International evidence-based guideline for the assessment and management of polycystic ovary syndrome*. Melbourne, Australia: Monash University.

Testa, M. A., & Simonson, D. C. (1996). Assessment of quality-of-life outcomes. *New England Journal of Medicine*, *334*(13), 835-840.

Thayer, R. E. (1989). *The Biopsychology of Mood and Arousal*. New York, USA: Oxford University Press.

Thomas, D., Elliott, E. J., & Naughton, G. A. (2006). Exercise for type 2 diabetes mellitus. *Cochrane Database of Systematic Reviews*, (3).

Thomson, R. L., Brinkworth, G. D., Noakes, M., Clifton, P. M., Norman, R. J., & Buckley, J. D. (2012). The effect of diet and exercise on markers of endothelial function in overweight and obese women with polycystic ovary syndrome. *Human Reproduction*, *27*(7), 2169-2176.

Thomson, R. L., Brinkworth, G. D., Noakes, M., Clifton, P. M., Norman, R. J., & Buckley, J. D. (2011). The effect of diet and exercise on vascular function in overweight and obese women with polycystic ovary syndrome. *Obesity Research & Clinical Practice*, (5), 15.

Thomson, R. L., Buckley, J. D., & Brinkworth, G. D. (2016). Perceived exercise barriers are reduced and benefits are improved with lifestyle modification in overweight and obese women with polycystic ovary syndrome: a randomised controlled trial. *BMC Women's Health*, *16*(1), 14.

Thomson, R. L., Buckley, J. D., Lim, S. S., Noakes, M., Clifton, P. M., Norman, R. J., & Brinkworth, G. D. (2010). Lifestyle management improves quality of life and depression in overweight and obese women with polycystic ovary syndrome. *Fertility and Sterility*, *94*(5), 1812-1816.

Thomson, R. L., Buckley, J. D., Moran, L. J., Noakes, M., Clifton, P. M., Norman, R. J., & Brinkworth, G. D. (2009). The effect of weight loss on anti-Müllerian hormone levels in overweight and obese women with polycystic ovary syndrome and reproductive impairment. *Human Reproduction*, *24*(8), 1976-1981.

Thomson, R. L., Buckley, J. D., Noakes, M., Clifton, P. M., Norman, R. J., & Brinkworth, G. D. (2008). The effect of a hypocaloric diet with and without exercise training on body composition, cardiometabolic risk profile, and reproductive function in overweight and obese women with polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, *93*(9), 3373-3380.

Tiggemann, M. (2005). Body dissatisfaction and adolescent self-esteem: Prospective findings. *Body Image*, *2*(2), 129-135.

Tremblay, M. S., Colley, R. C., Saunders, T. J., Healy, G. N., & Owen, N. (2010). Physiological and health implications of a sedentary lifestyle. *Applied Physiology, Nutrition, and Metabolism*, *35*(6), 725-740.

Trent, M., Austin, S. B., Rich, M., & Gordon, C. M. (2005). Overweight status of adolescent girls with polycystic ovary syndrome: body mass index as mediator of quality of life. *Ambulatory Pediatrics*, *5*(2), 107-111.

Turan, V., Mutlu, E. K., Solmaz, U., Ekin, A., Tosun, O., Tosun, G., ... & Malkoc, M. (2015). Benefits of short-term structured exercise in non-overweight women with polycystic ovary syndrome: a prospective randomized controlled study. *Journal of Physical Therapy Science*, *27*(7), 2293-2297.

Turner-McGrievy, G. M., Davidson, C. R., Wingard, E. E., & Billings, D. L. (2014). Low glycemic index vegan or low-calorie weight loss diets for women with polycystic ovary syndrome: a randomized controlled feasibility study. *Nutrition Research*, *34*(6), 552-558.

Turner-McGrievy, G., Davidson, C. R., & Billings, D. L. (2015). Dietary intake, eating behaviors, and quality of life in women with polycystic ovary syndrome who are trying to conceive. *Human Fertility*, *18*(1), 16-21.

Tworoger, S. S., Missmer, S. A., Eliassen, A. H., Barbieri, R. L., Dowsett, M., & Hankinson, S. E. (2007). Physical activity and inactivity in relation to sex hormone, prolactin, and insulin-like growth factor concentrations in premenopausal women. *Cancer Causes & Control*, *18*(7), 743-752.

Van Buuren, S. (2018). *Flexible Imputation of Missing Data* (2nd ed). Boca Raton, FL: Chapman & Hall/CRC.

van der Vange, N., Kloosterboer, H. J., & Haspels, A. A. (1987). Effect of seven low-dose combined oral contraceptive preparations on carbohydrate metabolism. *American Journal of Obstetrics and Gynecology*, *156*(4), 918-922.

Van Dyck, D., Cardon, G., Deforche, B., & De Bourdeaudhuij, I. (2015). IPAQ interview version: convergent validity with accelerometers and comparison of physical activity and sedentary time

levels with the self-administered version. *Journal of Sports Medicine and Physical Fitness, 55*(7-8), 776-786.

Van Zanten, J. J. V., Rouse, P. C., Hale, E. D., Ntoumanis, N., Metsios, G. S., Duda, J. L., & Kitas, G. D. (2015). Perceived barriers, facilitators and benefits for regular physical activity and exercise in patients with rheumatoid arthritis: a review of the literature. *Sports Medicine, 45*(10), 1401-1412.

Vanky, E., Kjøtrød, S., Salvesen, K. Å., Romundstad, P., Moen, M. H., & Carlsen, S. M. (2004). Clinical, biochemical and ultrasonographic characteristics of Scandinavian women with PCOS. *Acta Obstetricia et Gynecologica Scandinavica, 83*(5), 482-486.

Verweij, L. M., Terwee, C. B., Proper, K. I., Hulshof, C. T., & van Mechelen, W. (2013). Measurement error of waist circumference: gaps in knowledge. *Public Health Nutrition, 16*(2), 281-288.

Vigorito, C., Giallauria, F., Palomba, S., Cascella, T., Manguso, F., Lucci, R., ... & Orio, F. (2007). Beneficial effects of a three-month structured exercise training program on cardiopulmonary functional capacity in young women with polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, *92*(4), 1379-1384.

Villaseca, P., Hormaza, P., Cardenas, I., Oestreicher, E., & Arteaga, E. (2004). Ethinylestradiol/cyproterone acetate in polycystic ovary syndrome: lipid and carbohydrate changes. *The European Journal of Contraception & Reproductive Health Care*, *9*(3), 155-165.

Vink, J. M., Sadrzadeh, S., Lambalk, C. B., & Boomsma, D. I. (2006). Heritability of polycystic ovary syndrome in a Dutch twin-family study. *The Journal of Clinical Endocrinology & Metabolism*, *91*(6), 2100-2104.

Vizza, L., Smith, C. A., Swaraj, S., Agho, K., & Cheema, B. S. (2016). The feasibility of progressive resistance training in women with polycystic ovary syndrome: a pilot randomized controlled trial. *BMC Sports Science, Medicine and Rehabilitation*, *8*(1), 14.

von Elm, E., Altman. D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., Vandenbroucke, J. P. (2008). STROBE Initiative: The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies**.** *Journal of Clinical Epidemiology*, *61*(4), 344-9.

Vrbikova, J., Cifkova, R., Jirkovska, A., Lanska, V., Platilova, H., Zamrazil, V., & Starka, L. (2003). Cardiovascular risk factors in young Czech females with polycystic ovary syndrome. *Human Reproduction*, *18*(5), 980-984.

Wadhera, R. K., Steen, D. L., Khan, I., Giugliano, R. P., & Foody, J. M. (2016). A review of low-density lipoprotein cholesterol, treatment strategies, and its impact on cardiovascular disease morbidity and mortality. *Journal of Clinical Lipidology*, *10*(3), 472-489.

Wahid, A., Manek, N., Nichols, M., Kelly, P., Foster, C., Webster, P., ... & Roberts, N. (2016). Quantifying the association between physical activity and cardiovascular disease and diabetes: a systematic review and meta-analysis. *Journal of the American Heart Association*, *5*(9), e002495.

Wallach, E. E., McNeely, M. J., & Soules, M. R. (1988). The diagnosis of luteal phase deficiency: a critical review. *Fertility and Sterility*, *50*(1), 1-15.

Warburton, D. E., Nicol, C. W., & Bredin, S. S. (2006). Health benefits of physical activity: the evidence. *Canadian Medical Association Journal*, *174*(6), 801-809.

Ware, J. E., Jr., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36). Conceptual framework and item selection. *Medical Care, 30*(6), 473–483.

Ware, J. E., Kosinski, M., & Keller, S. D. (1998). SF-12. *How to score the SF-12 physical and mental health summary scales*, 1995.

Ware, J., Jr., Kosinski, M., & Keller, S. D. (1996). A 12-Item short-form health survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care, 34*(3), 220-233.

Warnecke, R. B., Johnson, T. P., Chavez, N., Sudman, S., O'rourke, D. P., Lacey, L., & Horm, J. (1997). Improving question wording in surveys of culturally diverse populations. *Annals of Epidemiology*, *7*(5), 334-342.

Warren, M. (2018). First analysis of 'pre-registered' studies shows sharp rise in null findings. *Nature*, 129-140.

Waters, L., Reeves, M., Fjeldsoe, B., & Eakin, E. (2012). Control group improvements in physical activity intervention trials and possible explanatory factors: a systematic review. *Journal of Physical Activity and Health*, *9*(6), 884-895.

Webb, M. A., Mani, H., Robertson, S. J., Waller, H. L., Webb, D. R., Edwardson, C. L., ... & Davies, M. J. (2018). Moderate increases in daily step count are associated with reduced IL6 and CRP in women with PCOS. *Endocrine Connections*, *7*(12), 1442.

Weiner, C. L., Primeau, M., & Ehrmann, D. A. (2004). Androgens and mood dysfunction in women: comparison of women with polycystic ovarian syndrome to healthy controls. *Psychosomatic Medicine, 66*(3), 356-362.

Welt, C. K., Gudmundsson, J. A., Arason, G., Adams, J., Palsdottir, H., Gudlaugsdottir, G., ... & Crowley, W. F. (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.

Westin, L., Carlsson, R., Erhardt, L., Cantor-Graae, E., & McNeil, T. (1999). Differences in quality of life in men and women with ischemic heart disease: a prospective controlled study. *Scandinavian Cardiovascular Journal*, *33*(3), 160-165.

White, D., Leigh, A., Wilson, C., Donaldson, A., & Franks, S. (1995). Gonadotrophin and gonadal steroid response to a single dose of a long-acting agonist of gonadotrophin-releasing hormone in ovulatory and anovulatory women with polycystic ovary syndrome. *Clinical Endocrinology, 42*(5), 475-481.

Wild, S., Pierpoint, T., McKeigue, P., & 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, P. T., Stefanick, M. L., Vranizan, K. M., & Wood, P. D. (1994). The effects of weight loss by exercise or by dieting on plasma high-density lipoprotein (HDL) levels in men with low, intermediate, and normal-to-high HDL at baseline. *Metabolism*, *43*(7), 917-924.

Winter, E. M., & Fowler, N. (2009). Exercise defined and quantified according to the Système International d'Unites. *Journal of Sports Sciences*, *27*(5), 447-460.

World Health Organization. (2009). *Global Health Risks: Mortality and burden of disease attributable to selected major risks*. Geneva, Switzerland: WHO.

World Health Organization. (2010). *Global recommendations on physical activity for health*. Geneva, Switzerland: WHO.

World Health Organization. (2011). *Use of glycated haemoglobin (HbA1c) in diagnosis of diabetes mellitus: abbreviated report of a WHO consultation* (No. WHO/NMH/CHP/CPM/11.1). World Health Organization.

Wright, C. E., Zborowski, J. V., Talbott, E. O., McHugh-Pemu, K., & Youk, A. (2004). Dietary intake, physical activity, and obesity in women with polycystic ovary syndrome. *International Journal of Obesity*, *28*(8), 1026.

Wuensch, K. L. (2016). An introduction to path analysis. [online] http://core.ecu.edu/psyc/wuenschk/MV/SEM/Path.pdf. [Last accessed on 4th June 2020].

Yang, Z., Scott, C. A., Mao, C., Tang, J., & Farmer, A. J. (2014). Resistance exercise versus aerobic exercise for type 2 diabetes: a systematic review and meta-analysis. *Sports Medicine*, *44*(4), 487-499.

Yildiz, B. O., & Azziz, R. (2007). The adrenal and polycystic ovary syndrome. *Reviews in Endocrine and Metabolic Disorders*, *8*(4), 331-342.

Yildiz, B. O., Bozdag, G., Yapici, Z., Esinler, I., & 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., & Azziz, R. (2008). Impact of obesity on the risk for polycystic ovary syndrome. *The Journal of Clinical Endocrinology & Metabolism*, *93*(1), 162-168.

Young, D. R., Hivert, M. F., Alhassan, S., Camhi, S. M., Ferguson, J. F., Katzmarzyk, P. T., ... & Yong, C. M. (2016). Sedentary behavior and cardiovascular morbidity and mortality: a science advisory from the American Heart Association. *Circulation*, *134*(13), e262-e279.

Yuan, Y. C. (2000, April). Multiple imputation for missing data: Concepts and new development. In *Proceedings of the Twenty-Fifth Annual SAS Users Group International Conference* (Vol. 267).

Zawadski, J. K., and Dunaif, A. (1992). Diagnostic criteria for polycystic ovary syndrome; towards a rational approach. In, Dunaif, A., Givens, J. R. and Haseltine, F. (eds). *Polycystic Ovary Syndrome*. Boston: Blackwell Scientific. pp. 377-384.

Zhang, B., Wang, J., Shen, S., Liu, J., Sun, J., Gu, T., ... & Bi, Y. (2018). Association of androgen excess with glucose intolerance in women with polycystic ovary syndrome. *BioMed Research International*, *2018*.

Zhang, H. Y., Guo, C. X., Zhu, F. F., Qu, P. P., Lin, W. J., & Xiong, J. (2013). Clinical characteristics, metabolic features, and phenotype of Chinese women with polycystic ovary syndrome: a large-scale case–control study. *Archives of Gynecology and Obstetrics*, *287*(3), 525-531.

Zhang, H. Y., Zhu, F. F., Xiong, J., Shi, X. B., & Fu, S. X. (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., Liu, Y., Liu, X., Xu, L., Zhou, L., Tang, L., ... & Hu, R. (2015). High intake of energy and fat in Southwest Chinese women with PCOS: a population-based case-control study. *PloS One*, *10*(5), e0127094.

Zhang, Y., Alyass, A., Vanniyasingam, T., Sadeghirad, B., Flórez, I. D., Pichika, S. C., ... & Morgano, G. P. (2017). A systematic survey of the methods literature on the reporting quality and optimal methods of handling participants with missing outcome data for continuous outcomes in randomized controlled trials. *Journal of Clinical Epidemiology*, *88*, 67-80.

Chapter 7: Appendices

Appendix 7.1. Review of author's judgements about each risk of bias item for each included study. Support for judgement based upon evidence presented within each paper.

Trial	Bias Domain	Source of Bias	Author's judgement	Support for judgement
Almenning *et al.* 2015	Selection Bias	Random sequence generation	Low Risk	Women were stratified by BMI and allocated in a 1:1:1 manner to study arms. Computer number random generator developed and administered to randomise subjects.
		Allocation concealment	Low Risk	Baseline testing was done before randomisation
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (*i.e.* supervised exercise sessions)
	Detection Bias	Blinding of outcome assessment	High Risk	Follow-up testing was performed, and these measurements were done non-blinded to group assignment. An observer blinded for group allocation analysed the FMD.
	Attrition Bias	Incomplete outcome data	Low Risk	89 participants assessed for eligibility; 58 excluded and reasons provided. 31 randomised and allocated 10:11:10. 6 (19%) lost to follow up (reasons provided) and data analysis completed on those remaining. Consort flow diagram used.
	Reporting Bias	Selective reporting	Low Risk	Trial preregistered on ClinicalTrials.gov (NCT01919281) and all proposed outcomes reported in paper.
	Other bias	Group similarity at baseline	Low Risk	FMD% significantly lower in HIT group. No other significant differences at baseline.
		Adherence	Low Risk	87% for RT arm and 90% for HIT arm.
		Contamination	Unclear Risk	Physical activity in control group not reported
Brown *et al*. 2009	Selection Bias	Random sequence generation	Low Risk	Randomisation was accomplished by generating a random sequence of two variables (representing the two treatment

Trial	Bias Domain	Source of Bias	Author's judgement	Support for judgement
				groups) using the online program at http://graphpadcom/quickcalcs/randomize 2.cfm
		Allocation concealment	Low Risk	Each group assignment was placed in its own sequentially numbered envelope by an individual not involved in the study. Participants were assigned to a group based on these envelopes, and each participant had an equal chance of being randomised to either group.
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (i.e. supervised exercise sessions)
	Detection Bias	Blinding of outcome assessment	High Risk	Not reported
	Attrition Bias	Incomplete outcome data	High Risk	Attrition is reported in study but considerably greater in exercise group. Acknowledged as a limitation in study and reasons for attrition not clearly stated. Overall attrition reported as 43%.
	Reporting Bias	Selective reporting	Unclear Risk	Unable to locate prospectively published trial protocol
	Other bias	Group similarity at baseline	High Risk	Significant differences in age and lipid profiles. Also, although not statistically significant, exercisers tended to be heavier, less hyperandrogenic, less fit and more insulin resistant.
		Adherence	Low Risk	89.8% adherence to exercise reported
		Contamination	Unclear Risk	Physical activity in control group not reported
Bruner *et al.* 2006	Selection Bias	Random sequence generation	Unclear Risk	Not reported
		Allocation concealment	Low Risk	Researcher chose a sealed envelope for each participant indicating which treatment they would receive.
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (*i.e.* supervised exercise sessions).

Trial	Bias Domain	Source of Bias	Author's judgement	Support for judgement
	Detection Bias	Blinding of outcome assessment	High Risk	Not reported
	Attrition Bias	Incomplete outcome data	High Risk	Attrition not reported. There are data missing from results; LH:FSH - 2 women in EN group (lab error & pregnancy); FI - 1 from EN & 1 from N (lab error).
	Reporting Bias	Selective reporting	Unclear Risk	Unable to locate prospectively published trial protocol.
	Other bias	Group similarity at baseline	Low Risk	No significant difference between groups for all outcomes.
		Adherence	Unclear Risk	Not reported
		Contamination	Unclear Risk	Physical activity in control group not reported
Guzick *et al.* 1994	Selection Bias	Random sequence generation	Unclear Risk	Subjects were randomised method used, not reported
		Allocation concealment	Unclear Risk	Not reported
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (*i.e.* supervised exercise sessions).
	Detection Bias	Blinding of outcome assessment	High Risk	Not reported
	Attrition Bias	Incomplete outcome data	Low Risk	Reports those who were excluded during screening. 12 participants randomised; results for 12 presented in findings. No missing data.
	Reporting Bias	Selective reporting	Unclear Risk	Unable to locate prospectively published trial protocol.
	Other bias	Group similarity at baseline	Low Risk	No significant difference between treatment and control subjects for key outcomes of interest.
		Adherence	Unclear Risk	Not reported
		Contamination	Unclear Risk	Physical activity in control group not reported
Hoeger *et al.* 2004	Selection Bias	Random sequence generation	Low Risk	Randomisation schedule was computer generated in blocks by an independent pharmacy representative.

Trial	Bias Domain	Source of Bias	Author's judgement	Support for judgement
		Allocation concealment	Unclear Risk	The block schedule was blinded to the investigators. Methods not reported.
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (*i.e.* supervised exercise sessions). However, participants and investigators double blinded to placebo or metformin by independent pharmacist.
	Detection Bias	Blinding of outcome assessment	High Risk	Not reported
	Attrition Bias	Incomplete outcome data	High Risk	Detailed analysis of attrition and adherence throughout. Balanced attrition across groups and explanations given for drop out. However, attrition is high in trial (39%).
	Reporting Bias	Selective reporting	Unclear Risk	Unable to locate prospectively published trial protocol.
	Other bias	Group similarity at baseline	Low Risk	No significant differences between groups for all outcomes.
		Adherence	Unclear Risk	Not reported
		Contamination	Unclear Risk	Physical activity in control group not reported
Konopka *et al*. 2015	Selection Bias	Random sequence generation	Unclear Risk	Women were randomised but unclear what method was used to do this.
		Allocation concealment	Unclear Risk	Women were assessed before and after the intervention. Unclear when randomisation took place and whether investigators were blinded.
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (*i.e.* supervised exercise sessions).
	Detection Bias	Blinding of outcome assessment	High Risk	Not reported
	Attrition Bias	Incomplete outcome data	High Risk	No attrition reported. However, hyperinsulinemic-euglycemic clamp only completed in a subset of obese women.

Trial	Bias Domain	Source of Bias	Author's judgement	Support for judgement
	Reporting Bias	Selective reporting	Low Risk	Trial preregistered on ClinicalTrials.gov (NCT02105428 and NCT01477164).
	Other bias	Group similarity at baseline	Low Risk	No significant differences between groups for all outcomes.
		Adherence	Unclear Risk	Not reported.
		Contamination	Unclear Risk	Physical activity in control group not reported.
Nasrekani *et al.* 2016	Selection Bias	Random sequence generation	Unclear Risk	Following eligibility screening and informed consent participants were randomised. Method of randomisation is not reported.
		Allocation concealment	Unclear Risk	Not reported whether assessors were blinded to allocation. Randomisation occurred before baseline measurements.
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (*i.e.* supervised exercise sessions).
	Detection Bias	Blinding of outcome assessment	High Risk	Not reported.
	Attrition Bias	Incomplete outcome data	Low Risk	20 participants randomised and all data reported. No use of Consort flow diagram.
	Reporting Bias	Selective reporting	Unclear Risk	Unable to locate prospectively published trial protocol.
	Other bias	Group similarity at baseline	Low Risk	No significant differences between groups for all outcomes.
		Adherence	Unclear Risk	Not reported.
		Contamination	Unclear Risk	Physical activity in control group not reported
Nybacka *et al.* 2011	Selection Bias	Random sequence generation	Low Risk	The randomisation was carried out with the permuted-block randomization method with ten blocks and a block size of 6.
		Allocation concealment	Unclear Risk	Not reported.
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (*i.e.* supervised exercise sessions).

Trial	Bias Domain	Source of Bias	Author's judgement	Support for judgement
	Detection Bias	Blinding of outcome assessment	High Risk	Same investigators completed outcome assessments but unclear whether they were blinded to allocation. Blinding unlikely.
	Attrition Bias	Incomplete outcome data	High Risk	Attrition is reported for each arm. Higher in 2 groups. But overall 25%. Reasons stated as personal or medical grounds.
	Reporting Bias	Selective reporting	Unclear Risk	Unable to locate prospectively published trial protocol.
	Other bias	Group similarity at baseline	Low Risk	Baseline characteristics were comparable regarding age, BMI, body composition, and endocrine, metabolic and gynaecological outcomes.
		Adherence	Unclear Risk	Not reported.
		Contamination	Unclear Risk	Physical activity in control group not reported
Petranyi *et al.* 2011	Selection Bias	Random sequence generation	Unclear Risk	Participants were age matched between groups. Method of sequence generation not reported.
		Allocation concealment	Unclear Risk	Not reported.
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (*i.e.* supervised exercise sessions).
	Detection Bias	Blinding of outcome assessment	High Risk	Not reported.
	Attrition Bias	Incomplete outcome data	Low Risk	Attrition not reported. 56 participants randomised and data present for all. No use of Consort (or similar) flow diagram.
	Reporting Bias	Selective reporting	Unclear Risk	Unable to locate prospectively published trial protocol.
	Other bias	Group similarity at baseline	Unclear Risk	Significance of baseline differences not reported. There appears to be some variation across outcomes.
		Adherence	Unclear Risk	Not reported.
		Contamination	Unclear Risk	Physical activity in control group not reported

Trial	Bias Domain	Source of Bias	Author's judgement	Support for judgement
Roessler *et al.* 2013	Selection Bias	Random sequence generation	Unclear Risk	Participants were randomised but the method used to generate sequence is not reported.
		Allocation concealment	Unclear Risk	Not reported.
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (*i.e.* supervised exercise sessions).
	Detection Bias	Blinding of outcome assessment	High Risk	Not reported.
	Attrition Bias	Incomplete outcome data	Low Risk	Three participants did not complete – injury (not study related) and time concerns stated. Baseline data presented for all participants and separately for completers.
	Reporting Bias	Selective reporting	Unclear Risk	Unable to locate prospectively published trial protocol.
	Other bias	Group similarity at baseline	Low Risk	No significant difference between groups for all outcomes reported.
		Adherence	High Risk	Aerobic exercise adherence was 67%
		Contamination	High Risk	This was a crossover design. Control group received group counselling sessions that explored motivation for, and barriers to PA. Exercise not stated but likely that behaviour may have been influenced.
Sa *et al.* 2015	Selection Bias	Random sequence generation	High Risk	Randomisation sequence computer generated but allocation to exercise group was partially based on ability to attend the 16 weeks of training, which was limited for some participants (n = 5) due to their remote geographical location.
		Allocation concealment	High Risk	Five participants allocated to control group as they were unable to attend all sessions.
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (*i.e.* supervised exercise sessions).

Trial	Bias Domain	Source of Bias	Author's judgement	Support for judgement
	Detection Bias	Blinding of outcome assessment	High Risk	Not reported.
	Attrition Bias	Incomplete outcome data	Low Risk	Consort flow diagram used; 30 randomised and baseline data presented for all in initial study. Post-intervention data presented for completers
	Reporting Bias	Selective reporting	High Risk	Unable to locate prospectively published trial protocol. A range of baseline outcomes reported, but no post-intervention analysis completed.
	Other bias	Group similarity at baseline	Low Risk	No significant differences between groups at baseline.
		Adherence	Unclear Risk	Not reported.
		Contamination	High Risk	Five participants in the control group did not receive the allocated intervention due to living in a remote geographical location
Saremi *et al.* 2013	Selection Bias	Random sequence generation	Unclear Risk	Quasi-randomisation. Methods used for sequence generation not reported.
		Allocation concealment	Unclear Risk	Not reported.
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (*i.e.* supervised exercise sessions).
	Detection Bias	Blinding of outcome assessment	High Risk	Not reported.
	Attrition Bias	Incomplete outcome data	Low Risk	22 randomised and all post-intervention data present. No evidence of Consort flow diagram.
	Reporting Bias	Selective reporting	Unclear Risk	Unable to locate prospectively published trial protocol.
	Other bias	Group similarity at baseline	Unclear Risk	Significant differences not reported between groups. Some variability in data (HOMA-IR and lipid profile).
		Adherence	Unclear Risk	Not reported.
		Contamination	Unclear Risk	Physical activity in control group not reported

Trial	Bias Domain	Source of Bias	Author's judgement	Support for judgement
Saremi *et al.* 2016	Selection Bias	Random sequence generation	Unclear Risk	Method of randomisation and sequence generation not reported.
		Allocation concealment	Low Risk	Investigators were blinded to group allocation prior to baseline testing.
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (*i.e.* supervised exercise sessions). However, allocation of placebo and calcium supplement was blinded to participant.
	Detection Bias	Blinding of outcome assessment	High Risk	Not reported.
	Attrition Bias	Incomplete outcome data	Unclear Risk	Attrition not reported. Consort flow diagram not presented. Number of participants randomised unclear.
	Reporting Bias	Selective reporting	Unclear Risk	Unable to locate prospectively published trial protocol.
	Other bias	Group similarity at baseline	Unclear Risk	Significant differences not reported between groups. Some variability in data (fasting insulin, blood glucose and lipid profile).
		Adherence	Unclear Risk	Not reported.
		Contamination	Unclear Risk	Physical activity in control group not reported
Stener-Victorin *et al.* 2009	Selection Bias	Random sequence generation	Low Risk	Randomly allocated in a 2:2:1 ratio to low-frequency EA, physical exercise, or no active intervention. To ensure equal proportions of age and BMI in each study arm, randomisation was stratified by those variables. Computer-generated randomisation within each stratum was conducted using permuted blocks of five.
		Allocation concealment	Unclear Risk	Allocation was concealed until interventions were assigned. Methods used, not reported.

Trial	Bias Domain	Source of Bias	Author's judgement	Support for judgement
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (*i.e.* supervised exercise sessions and EA).
	Detection Bias	Blinding of outcome assessment	High Risk	Not reported
	Attrition Bias	Incomplete outcome data	High Risk	Attrition data reported throughout each stage of the study. Comparable dropout in each arm of study and appropriate reasons provided. However, 29% attrition from randomisation to post-intervention and 40% to follow-up.
	Reporting Bias	Selective reporting	Low Risk	Trial preregistered on ClinicalTrials.gov (NCT00484705). 23 participants were recruited for microneurography; no criteria given for inclusion or detail on method of selection.
	Other bias	Group similarity at baseline	Low Risk	No significant differences between groups for all outcomes.
		Adherence	Low Risk	Mean number of weekly sessions reported; ~3 per week in PA group.
		Contamination	High Risk	There were no differences between the groups (PA, EA and control) in self-reports of PA frequency.
Thomson *et al.* 2008	Selection Bias	Random sequence generation	Low Risk	A parallel study design where subjects were randomly assigned by computer generation into three 20-wk lifestyle interventions.
		Allocation concealment	Unclear Risk	Not reported
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (*i.e.* supervised exercise sessions).
	Detection Bias	Blinding of outcome assessment	High Risk	Not reported.

Trial	Bias Domain	Source of Bias	Author's judgement	Support for judgement
	Attrition Bias	Incomplete outcome data	High Risk	Overview of reasons for dropout provided in study flow diagram but high rates reported – 49%
	Reporting Bias	Selective reporting	Unclear Risk	Unable to locate prospectively published trial protocol.
	Other bias	Group similarity at baseline	Low Risk	No significant differences between groups at baseline.
		Adherence	Unclear Risk	Not reported.
		Contamination	Unclear Risk	Physical activity in control group not reported
Turan *et al.* 2015	Selection Bias	Random sequence generation	Low Risk	A computer generated random number table was used to generate sequence for allocation.
		Allocation concealment	Low Risk	Randomised following baseline testing. Allocation concealed using pre-labelled, sealed envelopes.
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (i.e. supervised exercise sessions).
	Detection Bias	Blinding of outcome assessment	High Risk	Not reported.
	Attrition Bias	Incomplete outcome data	Low Risk	Small attrition (n = 2) from exercise group due to non-attendance of exercise sessions.
	Reporting Bias	Selective reporting	Unclear Risk	Unable to locate prospectively published trial protocol.
	Other bias	Group similarity at baseline	Low Risk	There were no significant differences between groups at baseline.
		Adherence	Low Risk	Two participants' data removed from analysis as adherence was < 75%
		Contamination	Unclear Risk	Physical activity in control group not reported
Vigorito *et al.* 2007	Selection Bias	Random sequence generation	Unclear Risk	Women were randomly subdivided into groups. Methods not reported.
		Allocation concealment	Unclear Risk	Not reported.
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (i.e. supervised exercise sessions).

Trial	Bias Domain	Source of Bias	Author's judgement	Support for judgement
	Detection Bias	Blinding of outcome assessment	Low Risk	All clinical assessments were performed by the same physician who was blinded to the patient allocation into the study protocol.
	Attrition Bias	Incomplete outcome data	Low Risk	All subjects completed the study protocol.
	Reporting Bias	Selective reporting	Unclear Risk	Unable to locate prospectively published trial protocol.
	Other bias	Group similarity at baseline	Unclear Risk	Significant differences at baseline not reported. Patients share similar characteristics across groups.
		Adherence	Low Risk	All participants completed the study protocol. Attendance was 100% in exercise group.
		Contamination	Low Risk	Small decrease (-0.1 MET hr/wk) in LTPA for control group.
Vizza *et al.* 2016	Selection Bias	Random sequence generation	Low Risk	Randomisation assignments were generated via an online randomisation programme.
		Allocation concealment	Low Risk	Randomisation done by an investigator not involved in the data collection and given to participants in sealed envelopes upon completion of baseline testing.
	Performance Bias	Blinding of participants and personnel	High Risk	Impossible to blind participants and personnel due to nature of the trial (*i.e.* supervised exercise sessions).
	Detection Bias	Blinding of outcome assessment	High Risk	Clinical assessment completed by lead investigator - lead investigator also completed weekly status check with participants to monitor adverse events. Suggests no blinding.
	Attrition Bias	Incomplete outcome data	High Risk	15 participants randomised, 13% attrition across trial. Attrition detailed in results section and baseline data from non-completers used in results. Baseline data carried forward for two participants in the PRT and three in the control group that did not complete follow-up testing.

Trial	Bias Domain	Source of Bias	Author's judgement	Support for judgement
	Reporting Bias	Selective reporting	Unclear Risk	Unable to locate prospectively published trial protocol.
	Other bias	Group similarity at baseline	High Risk	Paper reports no significant difference between baseline characteristics of groups but does note trends in waist and hip circumference being higher in exercise group. Looking at the descriptive characteristics, mean body weight and BMI are considerably greater in the exercise group albeit with large standard deviations.
		Adherence	High Risk	Supervised training sessions have very good attendance (95%) but home-based component was only 51%.
		Contamination	Unclear Risk	Physical activity in control group not reported

Appendix 7.2. Overview of previous systematic reviews assessing the effectiveness of exercise, or lifestyle in the management of polycystic ovary syndrome.

Study	Comparisons	Included studies (participants)	Methods	Outcomes
Harrison (2011)	1. EX vs. control	5 RCT and 3 cohort studies (*n* = 421)	*Databases*: MEDLINE, PsycINFO EMB Reviews, EMBASE, CINAHL *Quality assessment*: Standard forms adapted from the Cochrane Handbook of Systematic Reviews[1] and Quality of Reporting of Meta-Analyses checklist[3] *Analysis*: Qualitative synthesis only	BW▲, IR▲, Lipids, BP, reproductive function▲
Moran (2011)	1. LSM (EX, diet or EX and diet combined) vs. control	6 RCT (*n* = 164)	*Databases*: CENTRAL, MEDLINE, CINAHL, EMBASE, PsycINFO, AMED *Quality assessment*: Cochrane risk of bias tool, sensitivity analyses, I^2 statistic *Analysis*: Qualitative synthesis, meta-analysis (RE) assessment of MD post-intervention values.	LB, pregnancy, miscarriage, OC, OR, T*, SHBG, FT, FG*, BW*, BMI, WC*, WHR, OGTT, FBG, TC, HDL-C, LDL-C, TG, FI*, QOL
Domecq (2013)	1. LSM (EX, diet or EX and diet combined) vs. control 2. LSM vs. metformin	9 RCT (*n* = 610)	*Databases*: MEDLINE, EMBASE, CENTRAL, Web of Science, Scopus, PsycINFO, CINAHL *Quality assessment*: Cochrane risk of bias tool, sensitivity analyses, I^2 statistic *Analysis*: Qualitative synthesis, meta-analysis (FE and RE) MD of change from baseline to post-intervention.	FBG*, FI*, FG*, fertility, amenorrhea, acne
Haqq (2015)	1. EX vs. control 2. LSM vs. control	12 RCT (*n* = 668)	*Databases*: PubMed, CINAHL, CENTRAL	BMI*, BW*, WC*, WHR*, BF%*, FI, FBG, HOMA, TG,

Study	Comparisons	Included studies (participants)	Methods	Outcomes
			Quality assessment: Modified PEDro[4], sensitivity analyses, Cochrane Q, Egger plots *Analysis*: Qualitative synthesis, meta-analysis (FE and RE) MD of change from baseline to post-intervention	TC, LDL-C, HDL-C, CRP*, RHR*, VO_2 peak*
Benham (2018)	1. LSM (EX, diet or EX and diet combined) vs. control	7 RCT, 1 non-RCT and 6 UCT (n = 617)	*Databases*: MEDLINE, EMBASE *Quality assessment*: Newcastle-Ottawa Scale (NRCT), Cochrane risk of bias tool (RCT), I^2 statistic *Analysis*: Qualitative synthesis, semi-quantitative analysis of reproductive outcomes, meta-analysis (FE and RE) MD of change from baseline to post-intervention	LB, Pregnancy, OR, OC, MF, MCL, BMI, WC*, BF%, VO_2 max, SBP*, DBP, FBG, FI*, TC*, LDL-C*, HDL-C, TG*, HbA1c, HRQoL

Key: LSM: lifestyle modification; EX: exercise; RCT: randomised controlled trial; UCT: uncontrolled trial; CINAHL: Cumulative Index to Nursing and Allied Health Literature; EMBASE: Excerpta Medica dataBASE; CENTRAL: Cochrane Central Register of Controlled Trials; AMED: Allied and Complementary Medicine Database; FBG: fasting blood glucose; FI: fasting insulin; BMI: body mass index; BW: body weight; WC: waist circumference; WHR: waist-to-hip ratio; BF%: body fat percentage; HOMA: homeostatic model assessment; TG: triglycerides; TC: total cholesterol; LDL-C: low-density lipoprotein cholesterol; HDL-C: high-density lipoprotein cholesterol; CRP: C-reactive protein; RHR: resting heart rate; SBP: systolic blood pressure; DBP: diastolic blood pressure; LB: live births; OR: ovulation rate; OC: ovulatory cycle; MF: menstrual frequency; MCL: menstrual cycle length; FG: Ferriman-Gallwey score; T: testosterone; SHBG: sex hormone binding globulin; FT: free testosterone; OGTT: oral glucose tolerance test; QOL: quality of life; HRQoL: health-related quality of life; IR: insulin resistance; HbA1c: glycated haemoglobin; MD: mean difference; FE: fixed effects; RE: random effects; ▲: qualitative improvements consistently reported within the review; *: statistically significant (P <0.05) improvements in exercise groups compared to control. [1]Higgins and Green (2009); [3]Moher *et al.* (1999); [4]Maher *et al.* (1993).

Appendix 7.3. Details of excluded studies and reasons for exclusion.

Study	Reason for Exclusion
Almenning *et al.* 2015	Conference abstract; full study included in analysis
Asante *et al.* 2014	Conference abstract; full study included in analysis
Bachani *et al*. 2012	Comparison ineligible. Compares lifestyle modification with pharmacological intervention.
Barr *et al*. 2011	Ineligible study design. No intervention applied
Bongaard 2010	Comparison ineligible. Compares a low-GI diet with a moderate- to high-GI diet
Brown *et al.* 2005	Conference abstract; full study included in analysis
Chizen *et al*. 2014	Comparison ineligible. Compares a pulse diet with National Cholesterol Education Program Therapeutic Lifestyle Changes (TLC) Diet.
Christiansen *et al*. 2014	Ineligible patient population. Pregnant women.
Crosignani *et al*. 2003	Ineligible study design. Not a randomised controlled trial.
Curi *et al.* 2012	Comparison ineligible. Compares lifestyle modification with pharmacological intervention.
Fux Otta *et al*. 2010	Comparison ineligible. Compared lifestyle and pharmacological intervention with lifestyle and placebo.
Galletly *et al*. 2007	Comparison ineligible. Compares a low-protein high-carbohydrate diet with a high-protein low-carbohydrate diet.
Gambineri *et al*. 2006	Comparison ineligible. Compares a hypocaloric diet with placebo to a hypocaloric diet with pharmacological interventions.
Giallauria *et al*. 2008	Ineligible study design. Patients not randomised.
Harris-Glocker *et a*l. 2010	Ineligible comparison. Compares lifestyle modification and placebo with lifestyle modification and pharmacological intervention.
Hoeger *et al*. 2008	Ineligible patient population. Adolescent patients were used in the trial.
Jaffe *et al*. 2006	Comparison ineligible. Compares high-carbohydrate low-fat diet and placebo with high-carbohydrate low-fat diet and pharmacological intervention.
Jedel *et al*. 2010	Conference abstract; full study included in analysis
Johansson *et al*. 2013	Ineligible comparison. Compares acupuncture to physical therapy.
Karimzadeh *et al*. 2010	Comparison ineligible. Compares lifestyle to pharmacological interventions.
Kumar *et al*. 2010	Comparison ineligible. Compares lifestyle to lifestyle and pharmacological interventions.
Ladson *et al*. 2011	Comparison ineligible. Compares lifestyle to lifestyle and pharmacological interventions.
Ladson *et al*. 2011	Comparison ineligible. Compares lifestyle and placebo to lifestyle and pharmacological interventions.
Le Donne *et al*. 2012	Comparison ineligible. Compares diet with diet and pharmacological interventions.
Legro *et al*. 2014	Comparison ineligible. Compares lifestyle with lifestyle and pharmacological interventions.
Legro *et al*. 2015	Comparison ineligible. Compares lifestyle with lifestyle and pharmacological interventions.

Study	Reason for Exclusion
Liao *et al*. 2008	Ineligible study design. Observational design.
Lindholm *et al*. 2008	Comparison ineligible. Compares lifestyle and placebo to lifestyle and pharmacological interventions.
Ma *et al*. 2007	Ineligible comparison. Compares weight loss treatment to weight loss treatment and pharmacological intervention.
Machlitt *et al*. 2007	Ineligible comparison. Compares lifestyle and placebo with lifestyle and pharmacological intervention.
Marzouk *et al*. 2015	Ineligible intervention. Intervention use dietary advice and caloric restriction.
McBreairty *et al*. 2015	Ineligible comparison. Compares pulse based diet and exercise with National Cholesterol Education Program therapeutic lifestyle changes (TLC) diet and exercise.
Mehrabani *et al*. 2012	Ineligible intervention. Two hypocaloric diets are used for the intervention.
Moran *et al.* 2002	Ineligible intervention. Low-protein and a high-protein hypocaloric diets are used as the intervention.
Moran et *al*. 2003	Ineligible intervention. Meal replacement programme utilised; no comparison made.
Moran *et al*. 2006	Ineligible intervention. Two diets are utilised in the intervention.
Nidhi *et al*. 2012	Ineligible patient population. Adolescent patients were used in the trial. Also compares two exercise modalities.
Nidhi *et al*. 2013	Ineligible patient population. Adolescent patients were used in the trial. Also compares two exercise modalities.
Nidhi *et al*. 2013	Ineligible patient population. Adolescent patients were used in the trial.
Nybacka *et al*. 2014	Conference abstract; full study included in analysis
Nybacka *et al*. 2012	Conference abstract; full study included in analysis
Omar *et al*. 2013	Ineligible patient population. Compares women with PCOS to healthy controls.
Orio *et al.* 2016	Ineligible comparison. Compares lifestyle to pharmacological interventions.
Ornstein *et al.* 2011	Ineligible patient population. Adolescent patients were used in the trial.
Palomba *et al*. 2007	Ineligible comparison. Compares lifestyle to lifestyle and pharmacological interventions.
Palomba *et al*. 2010	Ineligible study design. Non-randomised controlled trial.
Papakonstantinou *et a*l. 2016	Ineligible intervention. Compares two dietary interventions in a cross-over design.
Pasquali *et al.* 1986	Ineligible comparison. Compares hypocaloric diet with diet and pharmacological intervention.
Pasquali *et al.* 2000	Ineligible patient population. Compares hypocaloric diet with diet and pharmacological intervention.
Popova *et al*. 2010	Ineligible comparison. Compares lifestyle with lifestyle and pharmacological interventions.
Randeva *et al*. 2002	Ineligible study design. Non-randomised controlled trial.
Redman *et al.* 2011	Ineligible patient population. Women with PCOS are compared to healthy controls.
Roessler *et al.* 2011	Conference abstract; full study included in analysis
Silva Dantas *et al.* 2015	Ineligible patient population. Compares women with PCOS to healthy controls.
Sorensen *et al*. 2011	Ineligible intervention. Either a high- or standard-protein diet are used for the intervention.
Tang *et al.* 2005	Ineligible comparison. Compares lifestyle and placebo with lifestyle and pharmacological intervention.
Thomson *et al*. 2009	Conference abstract; full study included in analysis
Thomson *et al*. 2011	Ineligible study design. No control group or comparison made.

Study	Reason for Exclusion
Turner-McGrievy *et al*. 2014	Ineligible study design. No lifestyle intervention
Turner-McGrievy *et al*. 2015	Ineligible comparison. Compares vegan diet to low-calorie diet.

Appendix 7.4. STROBE Statement—Checklist of items that should be included in reports of ***case-control studies***


Aston University
Illustration removed for copyright restrictions

Appendix 7.5. 12-item Short Form (SF-12) Health Survey


Aston University
Illustration removed for copyright restrictions


Aston University
Illustration removed for copyright restrictions

Appendix 7.6. - Polycystic Ovary Syndrome Questionnaire (PCOS-Q) - Self-Administered.


Aston University
Illustration removed for copyright restrictions


Aston University
Illustration removed for copyright restrictions


Aston University
Illustration removed for copyright restrictions

Appendix 7.7. International Physical Activity Questionnaire (IPAQ)


Aston University
Illustration removed for copyright restrictions


Aston University
Illustration removed for copyright restrictions


Aston University
Illustration removed for copyright restrictions


Aston University
Illustration removed for copyright restrictions


Aston University
Illustration removed for copyright restrictions

Appendix 7.8. Demographic questionnaire

Aston University

Understanding physical activity behaviours in women with polycystic ovary syndrome (PCOS)

Thank you for agreeing to participate in this study. The first thing we need from you is some information about you and your background. Please try to be as accurate as possible when providing your answers and please answer all questions. We remind you at this point that all information provided will be held in strictest confidence and at no point will your identity be revealed to anyone beyond the research team.

Name of Student Researcher: Mr. Chris Kite

Principal Investigator: Dr. James Brown

Aston University Study Number: 1442

Participant Identification Number for this study:

1. Please enter your date of birth (DD/MM/YYYY):

   ☐☐/☐☐/☐☐☐☐

2. Have you ever been diagnosed with PCOS? Yes ☐ No ☐

   (If you have answered NO to this question please skip ahead to Question 5)

3. For those who answered YES to Question 2. How long has it been since you were diagnosed with PCOS?

   _____ Years _____ Months

4. There are different varieties (phenotypes) of PCOS. Has your GP/consultant advised you, or are you aware which phenotype you have?

   Polycystic ovaries, menstrual disruption and excess androgens ☐

   Polycystic ovaries and menstrual disruption ☐

   Polycystic ovaries and excess androgens ☐

   Menstrual disruption and excess androgens ☐

   Not sure ☐

*Note: polycystic ovaries will usually be identified by ultrasound. Menstrual disruption refers to irregular or absent periods. Excess androgens mean you have elevated sex hormones (typically testosterone) causing symptoms like acne or excess hair growth.

5. What is your height?

   _____ Feet _____ inches or _____metres _____ cm

6. What is your weight?

   _____ stone _____ lbs or _____ kg

7. What is your waist measurement (circumference)?

_____ Inches or _____ cm

8. How would you define your ethnicity?

White ☐
Gypsy or Traveller ☐
Black or Black British: Caribbean ☐
Black or Black British: African ☐
Other Black background ☐
Asian or Asian British: Indian ☐
Asian or Asian British: Pakistani ☐
Asian or Asian British: Bangladeshi ☐
Chinese ☐
Other Asian Background ☐
Mixed: White & Black Caribbean ☐
Mixed: White & Black African ☐
Mixed: White & Asian ☐
Other Mixed background ☐
Arab ☐
Other Ethnic background ☐
Not known ☐
Decline to indicate ☐

9. Please indicate your marital status:

Married ☐ Single ☐
Divorced ☐ Widowed ☐

Civil-partnership ☐ Other (please specify) _____

10. Please tell us your occupational status:

Student	☐	Full-time employed	☐
Part-time employed	☐	Unemployed	☐
Retired	☐	House person	☐

Other (please specify) _____

11. Please tell us your highest level of education:

Secondary (O-level, GCSE, etc.) ☐ College (A-level, BTEC, etc.) ☐

Undergraduate (BSc, BA, etc.) ☐ Postgraduate (MA, MSc, PG-Cert, etc) ☐

Doctorate (PhD, MD, etc.) ☐ Other (please specify) ☐

12. Do you have children?

Yes ☐ No ☐

If yes, how many children do you have: Under 5 years _____
6-11 years _____
11-16 years _____

13. How many individual's live in your house?

Adults_____ Children_____

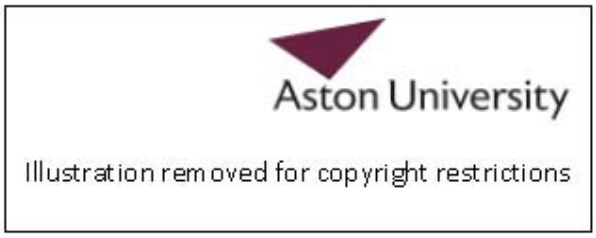
14. Please tell us your current annual household income:

Less than £39,999	☐	£40,000 to £79,999	☐
More than £80,000	☐		

Please ensure that you have answered all questions as accurately as possible. Should you have any queries about any item(s) on this form, please do not hesitate to contact a member of the research team (Chris Kite: [redacted]).

Thank you for taking the time to complete this questionnaire

Appendix 7.9. Rosenberg Self-esteem (RSE) Scale



Appendix 7.10. Self-efficacy for Exercise Scale

Appendix 7.11. Exercise Benefits/Barriers Scale (EBBS)


Aston University
Illustration removed for copyright restrictions

Appendix 7.12. Original project application to UK Biobank for access to database

Project Title (200 characters):
Clustering of cardiometabolic risk factors and their association with physical activity and sedentary time in women with and without polycystic ovary syndrome: a principal component analysis.

Research Question and aims (up to 5000 characters or 200 words):

- Do cardiometabolic risk factors cluster together in women with polycystic ovary syndrome (PCOS), and is the clustering pattern different compared to healthy controls?
- Are low physical activity (PA) levels and high sedentary time risk factors in the development of cardiometabolic conditions, do differences exist between women with PCOS and a healthy control group?

Our aims are:

- To describe the characteristics of female participants with and without PCOS according to their PA levels and sitting profiles: high PA-low sit (low cardiometabolic risk, reference group), high PA-high sit and low PA-low sit (intermediate risk), or low PA-high sit (high cardiometabolic risk) and the associations between groups and prevalence of cardiovascular disease (CVD), type 2 diabetes and metabolic syndrome.
- To compare age- and BMI-matched women with and without PCOS to identify whether CVD and metabolic risk factors are unique to women with PCOS, and to see if clustering of risk factors is different for women with PCOS to those who are obese only.
- To investigate whether other factors (such as socioeconomic status and mental health) cluster within subjects at increased risk of cardiometabolic conditions.

Does your project require biological samples?
No

Does your project require UK Biobank to re-contact participants?
No

Please provide information on each of the following:

A3. The background and scientific rationale of the proposed research project in general (up to 5000 characters or 300 words):

Polycystic ovary syndrome (PCOS) is a common endocrine disorder, affecting up to 21% of reproductive-aged women. PCOS is characterised by hyperandrogenism and ovulatory disruption, and typically manifests in a range of undesirable symptoms such as acne, hirsutism and infertility. Women with PCOS are also at an increased risk of cardiometabolic disturbances, such as obesity, increased waist circumference, and elevated cholesterol. Typically, management of PCOS focuses on lifestyle changes, incorporating increased physical activity (PA), aiming to alleviate symptoms, and lower the associated risk of type 2 diabetes and cardiovascular disease (CVD).

Although increasing PA is a well-established method for improving CVD risk factors, the effect of reducing sedentary behaviours has become a more recent focus. The proportion of sedentary time an individual engages in has been strongly linked to cardiometabolic risk, independent of PA.

If a diagnosis of PCOS increases cardiometabolic risk, then those who are not physically active and also spend more time sitting are further increasing the risk of disease.

If health professionals are able to better understand the characteristics of women with PCOS (and their counterparts without PCOS), and risk factors that are modifiable (e.g., PA levels), they may be able to more effectively target interventions for these risk factors, improving health-related outcomes in these women. Furthermore, this may allow earlier identification of individuals at risk of metabolic syndrome as a co-morbidity, and thus enable intervention that may reduce risk, decrease morbidity, improve patient quality of life, and lower subsequent health care cost.

A4. A brief description of the method(s) to be used (up to 5000 characters or 300 words):

Data from participants with PCOS plus age and age- and BMI-matched control groups of women without PCOS will be analysed. Outcomes selected according to diagnostic features of metabolic syndrome (waist circumference, blood pressure, blood glucose, and lipid profile). The physical (weight, BMI, cardiorespiratory fitness, inflammatory and androgenic biomarkers) and psychosocial (anxiety, depression, mood, self-esteem, and quality of life) characteristics associated with PCOS are also included. Incorporation of self-reported PA and sedentary behaviours, in addition to accelerometer measures where available.

Shapiro-Wilk tests of normality will be performed. Appropriate descriptive statistics (mean ± SD or median and IQR) of data will be presented. Depending upon distribution either independent samples T-tests or Mann-Whitney U will be performed to assess between population differences. Outliers (>3 standard deviations from the mean) will be identified and removed.

We will use principal component analysis (PCA) to assess whether one component can explain the variation within the data of each population. PCA is a statistical technique for identifying patterns within groups of correlated variables. For each population, we will assess associations between outcomes using Pearson's correlation coefficients; variables that are weakly correlated to others, will be excluded from this analysis.

PCA groups together linear combinations of the correlated variables into principle components. To allow interpretation, the identified principle components will be modified using orthogonal varimax rotation. This transforms the identified components so that each one is distinct (uncorrelated) from the others. The variables within each component are then more highly correlated than they are with variables in other components. The total variance explained by each component (eigenvalues) will be calculated by summing the squared correlation coefficients (r^2) between independent variables and calculated components. The cumulative sum of eigenvalues will then be calculated allowing an explanation of observed variance and a direct comparison of women with and without PCOS.

A5. The type and size of dataset required (e.g., case-control subset, men only, imaging data only, whole cohort, etc.) (up to 5000 characters or 100 words):

All women with PCOS (defined as ICD10 E28.2) from the Biobank (n = 331) and both an age- and age- and BMI-matched control group of women without PCOS.

A6. The expected value of the research (taking into account the public interest requirement) (up to 5000 characters or 100 words):

PCOS affects up to 20% of reproductive-aged women; typically, these women report undesirable symptoms including lower quality of life and increased risk of cardiometabolic conditions. Whilst studies have shown that exercise interventions may have beneficial effects on health outcomes in

PCOS, there is limited information about the health-effect of PA, or of excessive sitting/sedentary behaviours as part of daily living. Therefore principle component analysis will be used to identify whether risk factors for cardiometabolic conditions (including activity levels) cluster within women with PCOS, and if there is any variation when compared to healthy controls.

A7. Please provide up to 6 keywords which best summarise your proposed research project:

Polycystic ovary syndrome, physical activity, sitting time, cardiometabolic risk, principal component analysis, clustering

A8. Please provide a lay summary of your research project in plain English, stating the aims, scientific rationale, project duration and public health impact (up to 5000 characters or 400 words):

Polycystic ovary syndrome (PCOS) affects many women worldwide and causes a range of symptoms such as acne, excess hair, irregular periods, and sometimes, infertility. Women with PCOS are also more likely to be obese, have high blood pressure and high cholesterol which can lead to diabetes and heart disease.

Exercise is often recommended to manage the symptoms of PCOS. However, less is known about the effects of day-to-day activities. Day-to-day activities are the things you do for work, walking from place to place, household chores or sports you play for fun. We are also interested in the time you spend sitting. Those who sit for long periods of time have a higher risk of disease, and when activity levels are low, the risk is even greater. To our knowledge, this has not been investigated in women with PCOS.

Participants in the UK Biobank have reported their physical activity and sitting behaviours. We will assess these reported values to find out whether being more physically active, and sitting less reduces the risk of developing further conditions in PCOS. We will also compare this data to a group of women without PCOS to find out whether it is more (or less) important if you have been diagnosed with PCOS.

Gaining a better understanding about the role of being active, and high sitting time in PCOS will help to develop treatment guidelines and shape the information provided by health professionals. It will also help to educate women with PCOS about the physical and mental benefits of becoming more active.

A9. Will the research project result in the generation of any new data fields derived from existing complex datasets, such as imaging, accelerometry, electrocardiographic, linked healthcare data, etc, which might be of significant utility to other researchers:

No

A10. What is the estimated duration of your project, in months? If you consider (because for example the project is one involving the generation of hypotheses) that it would be difficult to set a fixed end point, we are prepared to consider a rolling 3-year period (during which annual updates are required):

12 months

Appendix 7.13. Descriptive statistics for outcomes included in cluster analysis; includes interquartile range (IQR), calculation of outlier cut points and trimmed values.

Outcome	PCOS				Age-matched				Age + BMI			
	IQR	IQR*1.5	5th	95th	IQR	IQR*1.5	5th	95th	IQR	IQR*1.5	5th	95th
SHBG (nmol/L)	39.42	59.13	16.59	135.63	40.77	61.16	20.45	139.85	38.55	57.83	21.47	128.13
T (nmol/L)	0.73	1.10	0.49	2.72	0.68	1.01	0.45	2.15	0.72	1.09	0.51	2.34
HDL-C (mmol/L)	0.46	0.68	0.92	2.06	0.52	0.79	1.00	2.23	0.50	0.75	0.94	2.09
HbA1c	6.20	9.30	28.50	53.53	4.80	7.20	27.54	40.26	5.20	7.80	27.90	44.31
IGF-1 (nmol/L)	9.40	14.10	11.44	33.52	7.49	11.24	13.78	34.78	8.56	12.85	11.96	32.78
LDL-C (mmol/L)	1.04	1.56	2.13	4.81	1.03	1.55	2.14	4.70	1.08	1.62	2.21	4.78
TG (mmol/L)	1.19	1.78	0.65	4.47	0.83	1.24	0.58	3.16	0.88	1.32	0.60	3.11
CRP (mg/L)	4.16	6.24	0.27	15.96	1.96	2.93	0.21	9.59	4.48	6.72	0.25	16.53
TC (mmol/L)	1.36	2.04	3.85	7.37	1.29	1.93	3.97	7.20	1.42	2.14	3.90	7.14
SBP (mmHg)	21.00	31.50	103.00	159.00	21.00	31.50	104.80	159.00	22.00	33.00	107.00	162.00
DBP (mmHg)	15.25	22.88	66.00	102.00	14.00	21.00	64.00	98.00	15.00	22.50	65.00	101.00
BF (%)	11.90	17.85	24.38	52.22	11.20	16.80	23.22	48.30	11.80	17.70	24.61	51.99
BMI (kg/m^2)	11.31	16.97	21.03	47.41	7.16	10.75	19.96	37.81	11.31	16.97	21.04	47.29
WC (cm)	26.50	39.75	67.00	125.15	18.00	27.00	66.00	108.50	26.00	39.00	69.00	128.00

Key: IQR: interquartile range; IQR*1.5: cut point for identification of outliers above 3rd quartile, or below 1st quartile; 5th: 5th centile; 95th: 95th centile; SHBG: sex hormone binding globulin; T: testosterone; HDL-C: high-density lipoprotein cholesterol; HbA1c: glycated haemoglobin; IGF-1: insulin-like growth factor-1; LDL-C: low-density lipoprotein cholesterol; TG: triglycerides; CRP: C-reactive protein; TC: total cholesterol; SBP: systolic blood pressure; DBP: diastolic blood pressure; BF%: body fat percentage; BMI: body mass index; WC: waist circumference.

For the PCOS, age-, and BMI + age-matched groups, the following number of cases were transformed respectively: SHBG: 10, 17 and 15; testosterone: 7, 7 and 9; HDL-C: 5, 7 and 7; HbA1c: 14, 11 and 19; IGF-1: 2, 8, 4; LDL-C: 2, 4 and 6; triglycerides: 10, 13 and 17; CRP: 9, 17 and 17; total cholesterol: 3, 7 and 7; SBP: 2, 6 and 11; DBP: 1, 5 and 6; body fat: 2, 1 and 5; BMI: 5, 11 and 8; waist circumference: 2, 6 and 4

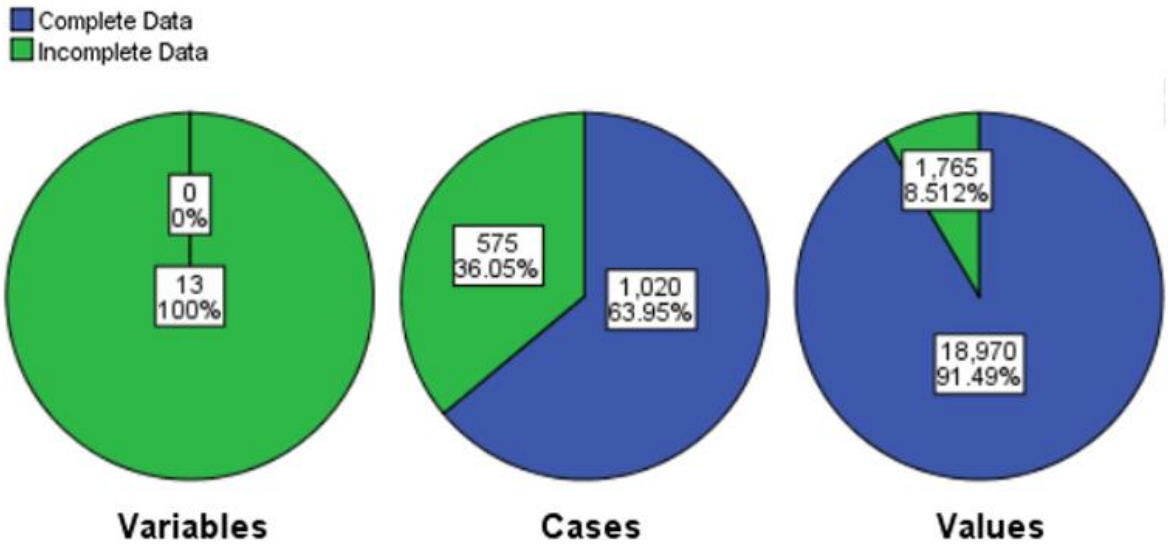
.

Appendix 7.14. Summary of missing data

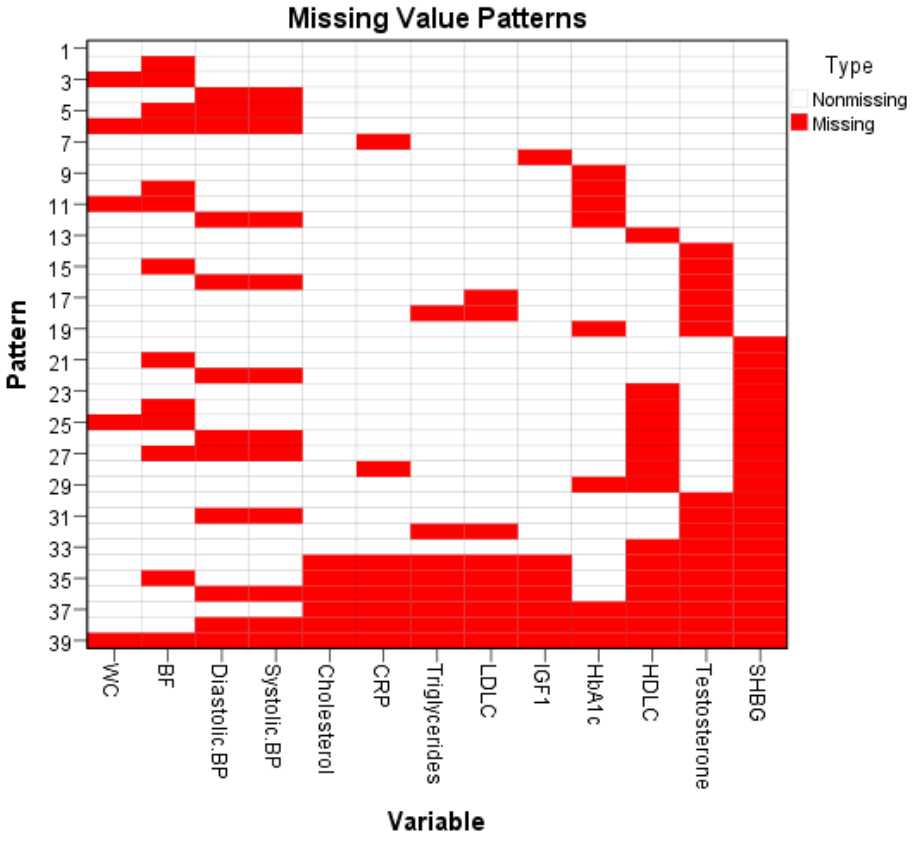
Outcome	Missing		Valid *N*	Mean	SD
	N	Percent			
SHBG (nmol/L)	279	17.5	1316	60.18	36.15
Testosterone (nmol/L)	264	16.6	1331	1.26	.67
Glucose (mmol/L)	258	16.2	1337	5.02	1.24
HDL-C (mmol/L)	256	16.1	1339	1.47	.38
HbA1c (mmol/mol)	127	8.0	1468	34.97	7.41
IGF-1 (nmol/L)	119	7.5	1476	22.75	6.36
LDL-C (mmol/L)	116	7.3	1479	3.38	.80
Triglycerides (mmol/L)	115	7.3	1480	1.50	.95
CRP (mg/L)	115	7.3	1480	3.57	5.59
Cholesterol (mmol/L)	113	7.1	1482	5.46	1.03
Systolic BP (mmHg)	101	6.4	1494	129.55	17.22
Diastolic BP (mmHg)	101	6.4	1494	81.03	10.83
Body Fat (%)	48	3.0	1548	37.98	8.46
WC (cm)	12	0.8	1583	89.37	16.87

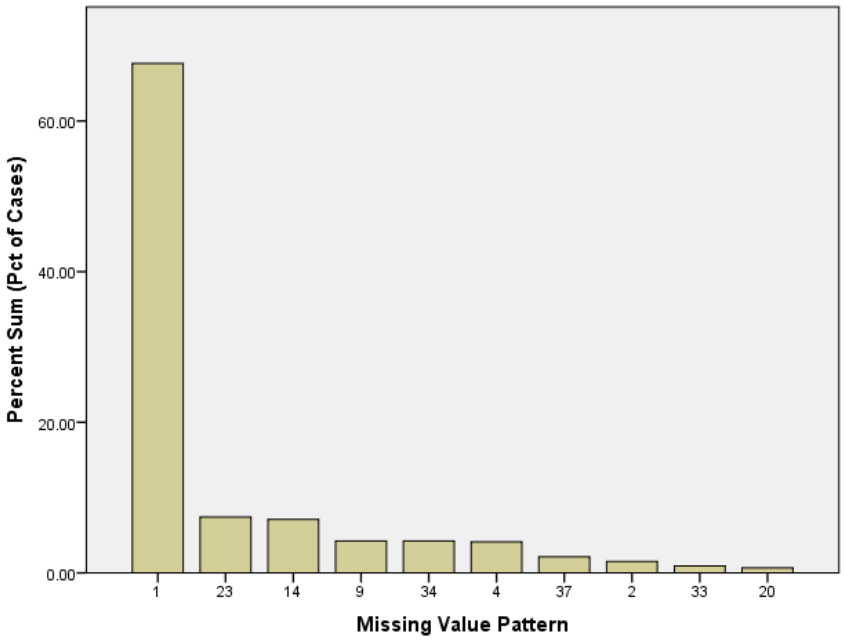
Key: SHBG: sex hormone binding globulin; HDL-C: high-density lipoprotein cholesterol; HbA1c: glycated haemoglobin; IGF-1: insulin-like growth factor-1; LDL-C: low-density lipoprotein cholesterol; BP: blood pressure; BMI: body mass index; WC: waist circumference; *N*: number of missing values; Percent: percentage of missing values from all cases.

Appendix 7.15. Summary of missing data analysis



Appendix 7.16. Patterns of missing variables and ten most common missing data patterns.





Appendix 7.17. Missing data, assumption checks (distribution, skewness and kurtosis) for each outcome by comparator group

	PCOS				Age-matched				Age + BMI-matched			
	Missing (%)	**Skewness**	**Kurtosis**	***P***	**Missing (%)**	**Skewness**	**Kurtosis**	***P***	**Missing (%)**	**Skewness**	**Kurtosis**	***P***
Stature (cm)	3 (0.9)	-0.0493	0.507	.207	3 (0.5)	-0.0697	-0.0543	.207	6 (0.9)	0.0601	0.428	.030
Body weight (kg)	3 (0.9)	0.836	0.623	<.001	4 (0.6)	1.24	2.56	<.001	6 (0.9)	0.903	0.813	.194
BMI (kg/m^2)	3 (0.9)	0.833	0.593	<.001	4 (0.6)	1.39	2.93	<.001	6 (0.9)	0.856	0.739	<.001
Waist Circumference (cm)	3 (0.9)	0.508	0.00614	<.001	4 (0.6)	0.924	0.818	<.001	5 (0.8)	0.535	-0.0813	<.001
Hip Circumference (cm)	3 (0.9)	0.855	0.535	<.001	4 (0.6)	0.956	1.33	<.001	5 (0.8)	0.849	0.674	.028
WHR	3 (0.9)	0.614	0.345	<.001	4 (0.6)	0.526	0.543	<.001	5 (0.8)	0.118	-0.0656	<.001
WSR	3 (0.9)	0.506	-0.187	<.001	3 (0.5)	0.562	2.50	<.001	6 (0.9)	0.553	0.0401	<.001
Body Fat (%)	12 (3.8)	-0.491	-0.143	<.001	17 (2.7)	-0.0275	-0.259	.110	18 (2.8)	-0.447	-0.0643	<.001
Systolic BP (mmHg)	21 (6.6)	0.320	0.0518	.056	43 (6.7)	0.781	1.10	<.001	37 (5.8)	0.984	2.24	<.001
Diastolic BP (mmHg)	21 (6.6)	0.226	0.0167	.174	43 (6.7)	0.330	0.260	.002	37 (5.8)	0.407	0.725	<.001
Total Cholesterol (mmol/L)	21 (6.6)	0.352	0.490	.052	50 (7.8)	0.478	0.547	<.001	42 (6.6)	0.537	1.27	<.001
HDL-C (mmol/L)	47 (14.7)	0.958	0.833	<.001	110 (17.2)	0.601	0.591	<.001	99 (15.5)	0.831	1.26	<.001
LDL-C (mmol/L)	21 (6.6)	0.246	0.222	.289	52 (8.2)	0.439	0.542	<.001	43 (6.7)	0.619	1.08	<.001
Triglycerides (mmol/L)	21 (6.6)	2.32	7.69	<.001	51 (8.0)	2.70	12.8	<.001	43 (6.7)	2.59	11.6	<.001
HbA1c (mmol/mol)	31 (9.7)	3.43	15.9	<.001	55 (8.6)	5.86	63.5	<.001	41 (6.4)	3.89	23.3	<.001
Glucose (mmol/L)	47 (14.7)	5.08	39.7	<.001	111 (17.4)	7.61	97.6	<.001	100 (15.7)	5.02	37.2	<.001
IGF-1 (nmol/L)	22 (6.9)	0.204	-0.0701	.126	51 (8.0)	0.457	0.462	<.001	46 (7.2)	0.251	0.106	.025
Testosterone (nmol/L)	56 (17.6)	2.79	15.9	<.001	112 (17.6)	1.39	4.03	<.001	96 (15.0)	4.58	45.2	<.001
SHBG (nmol/L)	51 (16.0)	1.80	4.42	<.001	125 (19.6)	1.51	3.01	<.001	103 (16.1)	1.48	2.86	<.001
Oestradiol (pmol/L)	131 (41.1)	4.21	25.5	<.001	314 (49.2)	2.31	6.60	<.001	304 (47.6)	2.98	12.6	<.001
CRP (mg/L)	21 (6.6)	2.69	9.84	<.001	52 (8.2)	5.11	33.6	<.001	42 (6.6)	4.03	29.2	<.001
MET-mins/wk	66 (20.6)	2.32	6.89	<.001	113 (17.7)	1.85	4.14	<.001	26 (4.1)	1.79	4.06	<.001
Summed PA (mins/wk)	66 (20.6)	1.45	2.09	<.001	113 (17.7)	1.32	1.90	<.001	26 (4.1)	1.45	2.49	<.001
MET-mins/wk VPA	66 (20.6)	4.18	22.3	<.001	113 (17.7)	2.91	13.1	<.001	26 (4.1)	3.25	16.9	<.001
MET-mins/wk MPA	66 (20.6)	2.13	4.06	<.001	113 (17.7)	2.32	5.70	<.001	26 (4.1)	2.38	5.85	<.001
Screen time (hrs/day)	-	1.83	6.68	<.001	-	1.23	2.26	<.001	-	2.37	9.70	<.001
Sedentary time (hrs/day)	-	1.64	5.53	<.001	-	1.05	1.84	<.001	-	2.18	8.38	<.001

Appendix 7.18. Frequencies of whether PA recommendations are being achieved with and without walking included and χ^2 Tests of frequencies for both scenarios.

Meeting PA recommendations		Study Group		
		PCOS	**Age-matched**	**BMI + Age -matched**
		n = 253	*n* = 525	*n* = 612
No walking included	Yes	112 (44.3)	278 (53.0)	291 (47.5)
	No	141 (55.7)	247 (47.0)	321 (52.5)
Includes walking	Yes	186 (73.5)	412 (78.5)	476 (77.8)
	No	67 (26.5)	113 (21.5)	136 (22.2)
		Value	**df**	***P***
No walking included	χ^2	6.06	2	.048
	n	1390		
Includes walking	χ^2	2.55	2	.279
	n	1390		

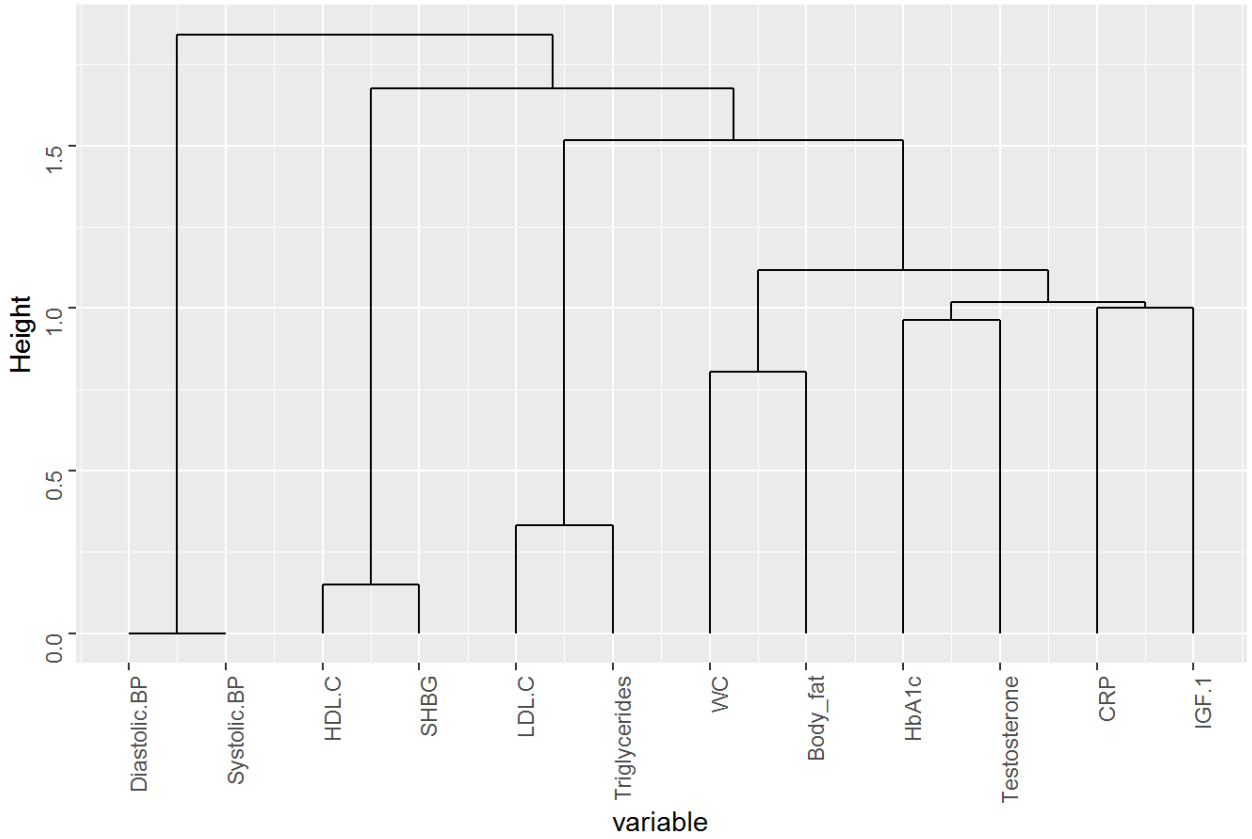
Key: IPAQ: International Physical Activity Questionnaire; PCOS: polycystic ovary syndrome; age-matched: control group age-matched only to case; BMI: body mass index; BMI + age-matched: control group both BMI + age-matched to case; χ^2: chi-square statistic; *n*: total participants in analysis; data are presented as number of participants in each IPAQ activity group (percentage of study group); χ^2: chi-squared statistic; df: degrees of freedom; *P*: statistical significance.

Appendix 7.19. Descriptive statistics for number of comorbidities, and only those related to metabolic health split by study cohort

PA Outcome	Group	Mean ± SD	Median (IQR)	Maximum	Skewness	Kurtosis	Shapiro-Wilk *P*
Total Comorbidities	PCOS	1.33 ± 1.72	1 (2)	9	1.82	3.91	<.001
	Age	0.47 ± 1.06	0 (1)	10	3.59	17.9	<.001
	BMI + Age	0.63 ± 1.28	0 (1)	10	3.44	15.7	<.001
Metabolic Comorbidities	PCOS	0.88 ± 1.36	0 (1)	10	2.05	5.03	<.001
	Age	0.25 ± 0.80	0 (0)	10	4.82	29.2	<.001
	BMI + Age	0.39 ± 1.01	0 (0)	10	4.02	20.3	<.001

Key: PA: Physical activity; Group: study group; Mean: average number of morbidities; SD: standard deviation; Median: median number of morbidities; IQR: interquartile range; Maximum: maximum number of morbidities reported within study group; *P*: significance from Shapiro-Wilk test; PCOS: women with PCOS group; Age: age-matched control group; BMI + age: BMI and age-matched group.

Appendix 7.20. Cluster plot of missing data and summary of missing data



Often when DBP is missing, so too is SBP. In addition, HDL-C and SHBG tend to be missing together and so to do LDL-C and triglycerides.

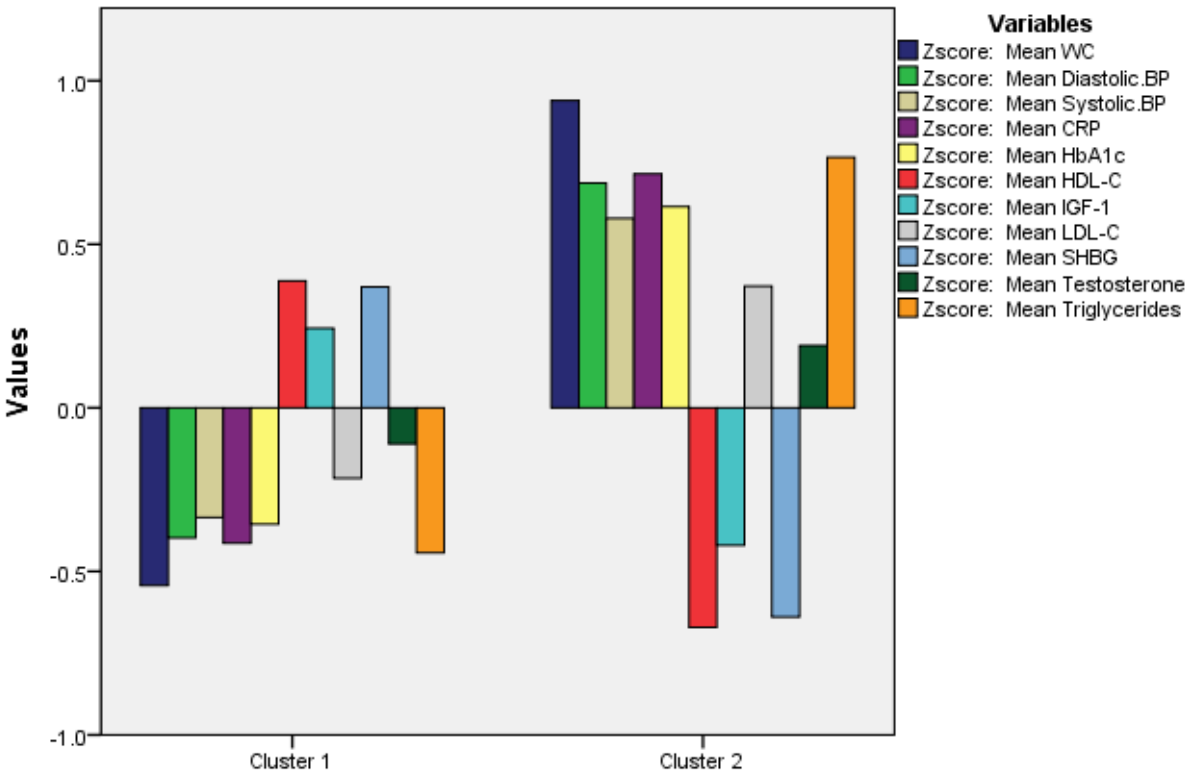
Although missing less frequently, three additional pairs also tend to be missing together. They are waist circumference and body fat, HbA1c and testosterone, and also CRP and IGF-1.

Appendix 7.21. Overview of imputed values and complete data set following imputation aggregation.

Outcomes	Pre-imputed Values (mean ± SD)	Complete data after imputation (mean ± SD)
Waist circumference (cm)	91.55 ± 15.52	89.26 ± 16.57
Diastolic BP (mmHg)	81.89 ± 11.26	80.96 ± 10.34
Systolic BP (mmHg)	131.13 ± 17.04	129.36 ± 15.89
C-reactive protein (mg/L)	3.97 ± 4.36	3.29 ± 4.13
HbA1c (mmol/mol)	34.49 ± 4.84	34.42 ± 4.64
HDL-C (mmol/L)	1.44 ± 0.37	1.46 ± 0.34
IGF-1 (nmol/L)	22.37 ± 6.56	22.66 ± 6.03
LDL-C (mmol/L)	3.36 ± 0.78	3.37 ± 0.75
SHBG (nmol/L)	59.36 ± 34.31	59.02 ± 31.07
Testosterone (nmol/L)	1.22 ± 0.54	1.23 ± 0.50
Triglycerides (mmol/L)	1.49 ± 0.85	1.46 ± 0.79

Key: SD: standard deviation; BP: blood pressure; HbA1c: glycated haemoglobin; HDL-C: high-density lipoprotein cholesterol; IGF-1: insulin-like growth factor-1; LDL-C: low-density lipoprotein cholesterol; SHBG: sex hormone binding globulin; mean: average values; SD: standard deviation.

Appendix 7.22. Graphical overview of final cluster centres for standardised variables.



Appendix 7.22. Change in cluster centres following 12 iterations - convergence achieved due to no change in cluster centres

	Initial Cluster Centres		Final Cluster Centres	
	Cluster One	**Cluster Two**	**Cluster One**	**Cluster Two**
Z score: WC	-1.22260	2.03621	-.54332	.93939
Z score: DBP	-2.02728	2.71140	-.39755	.68735
Z score: SBP	-2.22446	1.42455	-.33502	.57925
Z score: CRP	-.75960	3.00911	-.41367	.71524
Z score: HbA1c	-.67252	.03798	-.35559	.61481
Z score: HDL-C	.67177	-.97177	.38815	-.67110
Z score: IGF-1	1.23418	-.64303	.24309	-.42030
Z score: LDL-C	-.38437	1.31129	-.21520	.37208
Z score: SHBG	-.84106	.12650	.36989	-.63954
Z score: Testosterone	2.16762	-1.44133	-.10998	.19015
Z score: Triglycerides	-.94356	5.42349	-.44312	.76614

Key: WC: waist circumference; BP: blood pressure; CRP: C-reactive protein; HbA1c: glycated haemoglobin; HDL-C: high-density lipoprotein cholesterol; IGF-1: insulin-like growth factor-1; LDL-C: low-density lipoprotein cholesterol; SHBG: sex hormone-binding globulin.

Appendix 7.23. ANOVA between cluster one and cluster two for standardised final cluster centres.

	Cluster		Error			
	Mean Square	**df**	**Mean Square**	**df**	**F**	***P***
Z score: WC	812.025	1	.490	1589	1658.545	<.001
Z score: Diastolic BP	434.749	1	.727	1589	597.980	<.001
Z score: Systolic BP	308.747	1	.806	1589	382.906	<.001
Z score: CRP	470.734	1	.704	1589	668.292	<.001
Z score: HbA1c	347.828	1	.782	1589	444.946	<.001
Z score: HDL-C	414.437	1	.740	1589	560.191	<.001
Z score: IGF-1	162.552	1	.898	1589	180.949	<.001
Z score: LDL-C	127.393	1	.920	1589	138.401	<.001
Z score: SHBG	376.371	1	.764	1589	492.782	<.001
Z score: Testosterone	33.270	1	.980	1589	33.960	<.001
Z score: Triglycerides	540.128	1	.661	1589	817.493	<.001

Key: WC: waist circumference; BP: blood pressure; CRP: C-reactive protein; HbA1c: glycated haemoglobin; HDL-C: high-density lipoprotein cholesterol; IGF-1: insulin-like growth factor-1; LDL-C: low-density lipoprotein cholesterol; SHBG: sex hormone-binding globulin; *P*: significance value; df: degrees of freedom; F: test statistic.

Appendix 7.24. Descriptive statistics for missing cases based upon physical activity and sedentary behaviour risk.

	Mean ± SD	**Median (IQR)**	**Shapiro-Wilk *P***
Waist Circumference (cm)	91.1 ± 17.4	88.0 (27.0)	<.001
Diastolic BP (mmHg)	82.3 ± 10.7	81.9 (14.0)	.040
Systolic BP (mmHg)	131.6 ± 15.1	130.0 (21.0)	.003
C-reactive protein (mg/L)	3.8 ± 4.5	1.8 (4.63)	<.001
HbA1c (mmol/mol)	35.7 ± 6.0	34.3 (5.9)	<.001
HDL-C (mmol/L)	1.4 ± 0.3	1.3 (0.5)	<.001
IGF-1 (nmol/L)	21.7 ± 6.5	21.6 (8.8)	.696
LDL-C (mmol/L)	3.4 ± 0.8	3.4 (1.0)	.084
SHBG (nmol/L)	56.3 ± 32.5	51.2 (43.0)	<.001
Testosterone (nmol/L)	1.2 ± 0.5	1.2 (0.6)	<.001
Triglycerides (mmol/L)	1.7 ± 0.9	1.45 (1.2)	<.001
MET-mins/wk●	518.7 ± 595.3	297.0 (-)	.373
Sitting time (hrs/day)▲	4.7 ± 2.3	4.0 (3.0)	<.001

Key: ●: 204 cases missing; ▲: 8 cases missing; BP: blood pressure; HbA1c: glycated haemoglobin; HDL-C: high-density lipoprotein cholesterol; IGF-1: insulin-like growth factor-1; LDL-C: low-density lipoprotein cholesterol; SHBG: sex hormone binding globulin; MET-mins/wk: Metabolic equivalent of task-minutes per week; SD: standard deviation; IQR: interquartile range; *P*: statistical significance.