**Human Reproduction, Vol.25, No.2 pp. 544–551, 2010**
Advanced Access publication on November 12, 2009 doi:10.1093/humrep/dep399

human reproduction

ORIGINAL ARTICLE *Reproductive epidemiology*

# The prevalence of polycystic ovary syndrome in a community sample assessed under contrasting diagnostic criteria

**Wendy A. March[1,2], Vivienne M. Moore[2], Kristyn J. Willson[1], David I.W. Phillips[3], Robert J. Norman[1], and Michael J. Davies[1,4]**

[1]Discipline of Obstetrics/Gynaecology, Robinson Institute, The University of Adelaide, Adelaide, SA 5005, Australia [2]Discipline of Public Health, The University of Adelaide, Adelaide, SA 5005, Australia [3]Medical Research Council, Epidemiology Resource Centre and Developmental Origins of Health and Disease Division, University of Southampton, Southampton SO16 6YD, UK

[4]Correspondence address. E-mail: michael.davies@adelaide.edu.au

**BACKGROUND:** Polycystic ovary syndrome (PCOS) is considered to be the most common endocrine disorder in women of reproductive age, yet debate over appropriate diagnostic criteria and design limitations with sampling methodology have left some doubt as to the actual prevalence in the community. The objective of this study was to create a representative prevalence estimate of PCOS in the community under the National Institutes of Health (NIH) criteria and the more recent Rotterdam consensus criteria and Androgen Excess Society (AES) criteria.

**METHODS:** A retrospective birth cohort study was carried out in which 728 women born during 1973–1975 in a single maternity hospital were traced and interviewed in adulthood (age = 27–34 year; $n = 728$). Symptoms of PCOS (hyperandrogenism, menstrual dysfunction and polycystic ovaries) were identified by examination and the presence of polycystic ovaries in those that did not consent to the ultrasound were imputed.

**RESULTS:** The estimated prevalence of PCOS in this birth cohort using the NIH criteria was 8.7 $\pm$ 2.0% (with no need for imputation). Under the Rotterdam criteria, the prevalence was 11.9 $\pm$ 2.4% which increased to 17.8 $\pm$ 2.8% when imputed data were included. Under the AES recommendations, PCOS prevalence was 10.2 $\pm$ 2.2%, and 12.0 $\pm$ 2.4% with the imputed data. Of the women with PCOS, 68–69% did not have a pre-existing diagnosis.

**CONCLUSIONS:** The Rotterdam and AES prevalence estimates were up to twice that obtained with the NIH criteria in this, as well other prevalence studies. In addition, this study also draws attention to the issue of many women with PCOS in the community remaining undiagnosed.

**Key words:** endocrine disorder / polycystic ovary syndrome / hirsutism / menstrual irregularity / Rotterdam criteria

## Introduction

Polycystic ovary syndrome (PCOS) is a common endocrine disorder of women, characterized by a heterogeneous presentation of hyperandrogenism and ovulatory dysfunction. The aetiology is unknown but it has important long-term health implications, having been associated with type 2 diabetes, risk factors for cardiovascular disease (Amowitz and Sobel, 1999; Dokras, 2008) and endometrial carcinoma (Dahlgren *et al.*, 1991). As such, this disorder is a significant public health concern in society, which therefore indicates a need to accurately identify the proportion of women affected.

Despite PCOS being considered the most common endocrine disorder in women of reproductive age (Azziz *et al.*, 2004; Chang, 2004; Kauffman *et al.*, 2008), prevalence estimates are highly variable, ranging from 2.2% to as high as 26% (Knochenhauer *et al.*, 1998; Diamanti-Kandarakis *et al.*, 1999; Michelmore *et al.*, 1999; Asuncion *et al.*, 2000; Azziz *et al.*, 2004). This variability is due to several factors. Firstly, diagnosing the disorder is logistically difficult, with the necessity to carry out blood or ultrasound tests. This has resulted in prevalence studies being based on convenience samples and generally not exceeding 400 participants. For example, participants in commonly cited prevalence studies have been University employees (Knochenhauer *et al.*, 1998; Azziz *et al.*, 2004), or blood donors (Asuncion *et al.*, 2000) but there has been no indication of the representativeness of these subgroups in the studies. Secondly, considerable heterogeneity in the presentation of symptoms has contributed to a

lack of agreement over the diagnostic criteria used to define the condition. Recently new criteria have emerged but existing prevalence estimates have been based on prior National Institutes of Health (NIH) criteria (The Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group, 2004; Azziz *et al*., 2006). Consequently, there exists the necessity to provide an estimate of PCOS prevalence that is both representative and takes into account the differences in diagnosing the disorder.

The NIH diagnostic criteria were based on a consensus of experts who concluded that women have PCOS if they present with the combination of chronic oligo- or anovulation and clinical or biochemical signs of hyperandrogenism, with the exclusion of related disorders (Zawadski and Dunaif, 1992). A more recent workshop in Rotterdam, The Netherlands, gave rise to the Rotterdam criteria. This workshop suggested the addition of a third criteria—the presence of polycystic ovaries—as well as a statement that any two of the three criteria were sufficient for a positive diagnosis (The Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group, 2004). In contrast, the Androgen Excess Society (AES) maintains that androgen excess is a central feature of the disease and that PCOS should be defined by the presence of hyperandrogenism (clinical and/or biochemical) in combination with ovarian dysfunction (oligo-anovulation and/or polycystic ovaries), again with the exclusion of related disorders from other causes (Azziz *et al*., 2006).

No dedicated prevalence studies have used the Rotterdam or AES criteria to assess PCOS prevalence in a Caucasian population, although two have used the Rotterdam criteria in Asian populations (Chen *et al*., 2008; Kumarapeli *et al*., 2008). However, it is unlikely that these prevalences are applicable to Caucasians as the presentation of PCOS symptoms appears to vary considerably between Asian and Caucasian populations (Balen and Michelmore, 2002). Several small studies have applied the Rotterdam criteria to a Caucasian population and although they have not been dedicated prevalence studies, they suggest that PCOS prevalence could effectively double under the Rotterdam criteria (Lowe *et al*., 2005; Broekmans *et al*., 2006). Consequently, the PCOS prevalence commonly referenced in literature is not consistent with current opinion on appropriate diagnostic criteria and therefore a gap exists in the literature to provide a PCOS prevalence estimate using the three sets of criteria.

To address the shortfalls discussed, this study reports on PCOS prevalence under all three diagnostic criteria—the NIH, Rotterdam and AES. Furthermore, by carrying out a large community-based, highly inclusive study, this affords the utmost opportunity to provide a representative value of PCOS prevalence in the community. To achieve this aim, we interviewed an unselected cohort of predominately Caucasian women aged 27–34 years, born in Adelaide, South Australia. In this group of women, we assessed the presence of relevant PCOS symptoms to determine the prevalence of PCOS under the NIH, Rotterdam and AES diagnostic criteria.

# Materials and Methods

## Participants

The survey was based on a cohort born in Adelaide, South Australia, in which all consecutive female babies born during January 1973–December 1975 in the Queen Elizabeth Hospital (QEH), who survived to discharge, were traced when they were around 30 years of age. This sampling frame was chosen because it was the most logistically feasible, that maximized the representativeness in presenting prevalence and enabled further research on PCOS. Births at the QEH were broadly representative of those in Adelaide, as it was one of two major maternity hospitals, delivering approximately 12% of births in the city, with a geographically large catchment area and sociodemographic profile broadly representative of all women giving birth in South Australia. Daughters were usually traced through their mothers, whose details were in the hospital records. Where this was not possible alternative methods were used—for example, we searched phone records if there were a limited number with the surname, or historical birth announcements in the newspaper to identify the first name of the daughter, who was contacted directly using electoral roll information. From 2199 birth records, 2046 (93.0%) daughters were traced and 62 were deceased or disabled (3.0%), leaving 1984 (90.2%) women who were invited to participate in the study. Women living outside the Adelaide metropolitan area or interstate were excluded from the present study because they would not be able to attend the medical facility for an ultrasound scan ($n = 609$). Of the 1375 (69.3% of 1984) living in Adelaide, the response rate of those agreeing to be interviewed was 52.9% ($n = 728$; Fig. 1). Participating women were representative of all female babies born in the same time period at the QEH in terms of birthweight, multiple births ($n = 15$ twins), birth order and country of origin of mother (data not shown). The study contained a significantly greater proportion of mothers of a higher socioeconomic status [Socio-Economic Indexes for Areas (SEIFA) Index of Relative Disadvantage (McLennan, 1998)] but this difference was less than 5% [high SEIFA group: all QEH births $n = 92$ (5.5%); interviewed daughters $n = 64$ (8.8%); $P = 0.0002$].

## Study protocol

Interviews were conducted by trained research nurses, usually in the home of the participant. A medical history was obtained, gathering detailed anthropometry and focusing on the symptoms of PCOS—specifically menstrual irregularities and clinical hyperandrogenism, as well as other questions relevant to the diagnosis, such as gynaecological history and whether the participant had previously been diagnosed with PCOS. Anthropometry measurements were taken in duplicate and the mean was reported. Weight and percent body fat were obtained using a bio-impedance body analyser (Tanita TBF 5383611). Menstrual irregularity was assessed as the presence of chronic amenorrhea, or a usual cycle length of less than 21 days or more than 35 days, or greater than a 4-day variation between cycles (Polson *et al*., 1988; Cresswell *et al*., 1997). Where there was a factor that may affect their menstrual cycle [i.e. currently using the oral contraceptive pill (OCP) ($n = 275$) or other hormonal contraceptive measures ($n = 9$), breastfeeding or pregnancy ($n = 14$), or a relevant medical procedure such as a hysterectomy ($n = 4$)], the women were asked to describe their former menstrual cycle. However, in some cases this was not possible ($n = 31$); for example, they may have been taking the OCP almost continuously since menarche. Clinical hyperandrogenism was assessed as the self-reported degree of hirsutism using the modified Ferriman-Gallwey (mF-G) scoring method (Hatch *et al*., 1981). The women compared the amount of body hair they had before hair removal with a chart displaying degree of hair growth in nine regions. If women reported menstrual irregularity and/or an mF-G score $\geq 8$ they were invited to a clinical examination. The remaining women were considered to not have PCOS and therefore did not participate further.

Of the 277 women that met the clinical examination conditions, 108 (39.0%) consented to a vaginal ultrasound of the ovaries and blood test (Fig. 1). Serum TSH, prolactin and 17-hydroxyprogesterone levels were

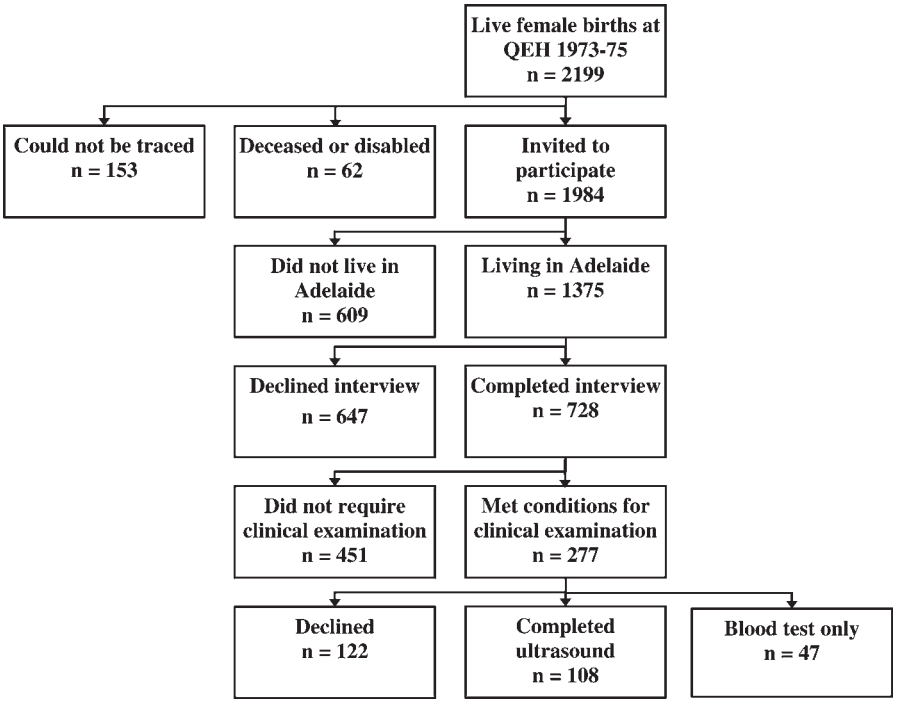


**Figure 1** Flow chart of women involved in the study; from tracing through to the interview and clinical examination.

measured to exclude, respectively, hypothyroidism, hyperprolactinemia and congential adrenal hyperplasia as confounding causes for menstrual dysfunction. An additional 47 women had a blood test but no ultrasound. The study was approved by the QEH and the University of Adelaide ethics committees and all participants gave written consent.

## Defining PCOS

Under the NIH criteria, PCOS was defined as the combination of menstrual disorders (as indicated previously) together with clinical and/or biochemical hyperandrogenism. Thus, it was not necessary for women to have the clinical examination to be assessed as having PCOS under the NIH criteria. For the Rotterdam criteria, PCOS was defined by the presence of two or more of the following; clinical and/or biochemical hyperandrogenism, menstrual disorders and polycystic ovaries. The AES definition was analogous to the Rotterdam criteria but excluded women with only menstrual dysfunction and polycystic ovaries. Clinical hyperandrogenism was defined by a mF-G score $\geq$8, derived from the 95th percentile of a population and chosen because it is commonly accepted as representing abnormal hair growth in a Caucasian population (Hatch *et al*., 1981; DeUgarte *et al*., 2006). Free testosterone (free T) was used to assess hyperandrogenemia because it was considered one of the more sensitive methods (The Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group, 2004); and elevated levels were delineated by the upper 95th percentile (34.2 pmol/l) in 100 women with no signs of PCOS with regard to menstrual history, hirsutism and polycystic ovaries—a method commonly used in previous PCOS prevalence studies to establish the normal free T range. Polycystic ovaries were identified by vaginal ultrasound, conducted in the follicular phase or when hormonal assessment showed no follicular activity. Ultrasounds were performed using a General Electric Logiq 5 scanner with an E8C 6.5 MHz micro-convex vaginal probe. A positive finding of polycystic ovaries required either 12 or more follicles measuring 2–9 mm in diameter, or increased ovarian volume (>10 cm) in at least one of the ovaries (Balen *et al*., 2003).

## Hormonal analysis

Free T was calculated from the total testosterone and SHBG levels as described by Vermeulen *et al*. (1999). SHBG activity was measured by immunoradiometric assay (IRMA) (Spectria SHBG IRMA, Orion Diagnostica, Finland) with intra- and inter-assay variations of 5.5 and 6.9%, respectively. Total testosterone was measured by radioimmunoassay (RIA) using the DSL-4100 kit (Diagnostic systems laboratories, Inc., Webster, TX, USA), with 8.1 and 10.5% intra- and inter-assay variations, respectively.

## Statistical analysis

Descriptive statistics were generated to enable comparisons between groups. Continuous variables were checked for normality and means were presented with standard deviations, or medians and interquartile ranges, as appropriate. Distributions were compared using Student's *t*-test or Mann–Whitney *U* as appropriate. Categorical variables were compared using Pearson's $\chi^2$ test. A *P*-value of $<0.05$ was considered significant. All statistical analyses were carried out in SPSS 15.0.0 for Windows (SPSS Inc., Chicago, IL, USA).

Since not all women who met the conditions for a clinical examination actually had an ultrasound, the additional number of women likely to have polycystic ovaries was calculated. This imputation was undertaken with the aforementioned women subdivided into PCOS phenotypes as per the Rotterdam criteria. For each phenotype subset, the proportion imputed to have polycystic ovaries was calculated by multiplying the proportion with polycystic ovaries in the group completing an ultrasound, by the number not undertaking an ultrasound.

# Results

## The study population

Table I characterizes all women interviewed ($n = 728$), and of these, the women who met the conditions for clinical examination ($n = 277$), and those for whom an ultrasound was completed ($n = 108$). The participants reported that they were primarily of European ancestry (94% of 725), with the remainder being Pacific Islander or Aboriginal (3%), Asian (1%), Middle-eastern (1%) or South American (1%). Almost a third of women of European decent (29% of 686) reported Mediterranean ethnicity (Greek, Italian or Spanish). Some 203 (28%) of the study participants were obese ($\geq$30 kg/m$^2$).

Of the 728 women interviewed, 154 (21.2%) were classified as hirsute (mF-G score $\geq$8) and 173 (23.8%) had a history of menstrual dysfunction. There were 41 women (5.6%) who reported a prior PCOS diagnosis. At interview, 275 women (38%) were currently using the OCP. Women currently taking the OCP (per Table I) did not differ significantly in the proportion with PCOS or any of the PCOS criteria to those not taking the OCP [e.g. 25.1 versus 23.0% ($P = 0.53$), 20.7 versus 21.4% ($P = 0.85$); taking versus not taking the OCP, for menstrual dysfunction and hirsutism, respectively]. Therefore these two groups (taking and not taking the OCP) were not assessed separately.

Not all women who met the conditions for a clinical examination had an ultrasound, as 122 declined, 33 were not referred due to clerical errors, seven were interviewed after the clinical examination period closed, and we were unable to trace seven women. However, the ethnicity, anthropometry, parity, degree of hirsutism and history of menstrual dysfunction for women who completed an ultrasound was not significantly different from those who did not do so (data not shown). Among the 108 women who completed the ovarian ultrasound, 41 (38.0%) had polycystic ovaries. None of these women exhibited signs of other disorders that should be excluded before a diagnosis of PCOS could be made.

## PCOS prevalence

Table II compares the estimated prevalence of PCOS using the three different criteria. The combination of hirsutism and/or high free T (above the 95th percentile) and menstrual dysfunction occurred in 63 women. This group was therefore classified as having PCOS under the NIH criteria, giving an overall prevalence of 8.7% (63/728; 95% confidence interval: 6.6, 10.7). A total of 87 women fulfilled the Rotterdam criteria giving a prevalence of 11.9% (95% confidence interval: 9.6, 14.3). This comprised the 63 women with hirsutism and/or high free T and menstrual dysfunction and a further 24 with polycystic ovaries in combination with either hirsutism, high free T

**Table I Age, anthropometric and gynaecological characteristics in the study population, the women who met the conditions for the clinical examination and those undertaking the ultrasound**

| | $n_{avail}$ | All women interviewed | $n_{avail}$ | Met conditions for clinical examination | $n_{avail}$ | Had ultrasound |
|---|---|---|---|---|---|---|
| *n* | | 728 | | 277 | | 108 |
| Age (years) | 728 | 30.2 (29.9–30.9) | 277 | 30.2 (29.8–30.9) | 108 | 30.2 (29.1–30.4) |
| Height (m) | 728 | 163.7 ± 6.8 | 277 | 163.9 ± 6.6 | 108 | 163.7 ± 6.6 |
| Weight (kg) | 726 | 69.0 (60.0–83.2) | 277 | 71.3 (60.2–88.7) | 108 | 72.1 (59.6–95.7) |
| Body mass index (kg/m$^2$) | 726 | 25.7 (22.5–30.9) | 277 | 26.5 (22.9–32.8) | 108 | 27.2 (23.1–34.6) |
| Waist circumference (cm) | 723 | 79.6 (72.6–91.8) | 276 | 82.1 (74.0–96.9) | 108 | 82.4 (74.0–104.2) |
| Body fat (%) | 711 | 35.5 (28.9–43.0) | 273 | 36.5 (30.0–44.5) | 107 | 37.8 ± 10.2 |
| Free T (pmol/l) | | | 155 | 26.0 ± 18.6 | 108 | 24.5 ± 23.9 |
| mF-G score | 728 | 3.0 (1.0–7.0) | 277 | 8.0 (3.0–11.0) | 108 | 8.0 (4.0–11.8) |
| Hirsutism (mF-G score $\geq$8) | 728 | 21.2% (154) | 277 | 55.6% (154) | 108 | 58.3% (63) |
| Menstrual irregularity | 728 | 23.8% (173) | 277 | 62.5% (173) | 108 | 66.7% (72) |
| Amenorrhea | | 0.8% (6) | | 2.2% (6) | | 3.7% (4) |
| <21 days | | 2.2% (16) | | 5.8% (16) | | 2.8% (3) |
| >35 days | | 13.5% (98) | | 35.4% (98) | | 42.6% (46) |
| >4-day variation | | 7.3% (53) | | 19.1% (53) | | 18.5% (20) |
| Polycystic ovaries | | | | | 108 | 38.0% (41) |
| Currently taking an oral contraceptive | 728 | 37.8% (275) | 277 | 38.6% (107) | 108 | 37.0% (40) |
| Caucasian ethnicity | 727 | 94% (686) | | 94% (261) | | 90% (97) |
| Parity | 728 | | 277 | | 108 | |
| 0 births | | 42.9% (312) | | 46.2 (128) | | 54.6% (59) |
| 1–2 births | | 46.8% (341) | | 43.6 (121) | | 35.2% (38) |
| 3+ births | | 10.3% (75) | | 10.1 (28) | | 10.2% (11) |

Data are mean ± SD, median (IQR), or percentage (*n*). $n_{avail}$= the number of women available for each characteristic measured.

**Table II Number of individuals with PCOS under the NIH, Rotterdam and AES criteria among 728 unselected women**

| PCOS criteria | Total known PCOS *n* (% ± CI) | Total + imputed polycystic ovaries[a] |
|---|---|---|
| 1. NIH | 63 (8.7 ± 2.0) | |
| (Menstrual dysfunction + hirsutism and/or high free T) | | |
| 2. Rotterdam | 87 (11.9 ± 2.4) | 129.5 (17.8 ± 2.8) |
| Phenotypes | | |
| a. Menstrual dysfunction + hirsutism and/or high free T + polycystic ovaries | 17 (2.3 ± 1.1) | 27.4 (3.8 ± 1.4) |
| b. Menstrual dysfunction + hirsutism and/or high free T only (no polycystic ovaries) | 21 (2.9 ± 1.2) | 35.6 (4.9 ± 1.6) |
| a. or b. Menstrual dysfunction + hirsutism and/or high free T+? polycystic ovaries (unknown if polycystic ovaries present as did not have an ultrasound) | 25 (3.4 ± 1.3) | values added to phenotype a. or b. |
| c. Hirsutism and/or high free T + polycystic ovaries | 11 (1.5 ± 0.9) | 24.5 (3.4 ± 1.3) |
| d. Menstrual dysfunction + polycystic ovaries | 13 (1.8 ± 1.0) | 42.1 (5.8 ± 1.7) |
| 3. Androgen Excess Society (a., b. or c.) | 74 (10.2 ± 2.2) | 87.5 (12.0 ± 2.4) |

CI = 95 percent confidence interval.
[a]Includes imputed values of the number of women with polycystic ovaries in the group that did not have an ultrasound. This was calculated within each phenotype by multiplying the proportion with polycystic ovaries in the group completing an ultrasound, by the number not undertaking an ultrasound [e.g. for phenotype d. (13/34) × 76 = 29.1; where *n* = 76 did not have an ultrasound in this phenotype; and 29.1 + 13 = 42.1 (the imputed value)].

or menstrual dysfunction. However, by assuming that the women who did not have an ultrasound had a similar prevalence of polycystic ovaries and including the imputed data, this prevalence estimate increased to 17.8% (95% confidence interval: 15.0, 20.6). On the basis of the AES recommendations, 74 were classified as having PCOS (comprising the 63 women with menstrual dysfunction and hirsutism and 11 with hirsutism and/or high free T and polycystic ovaries). This equated to a prevalence of 10.2% (95% confidence interval: 8.0, 12.4). This prevalence increased to 12.0% (95% confidence interval: 9.7, 14.4) upon inclusion of the imputed data of women who did not have an ultrasound examination.

# Discussion

This study found the prevalence of PCOS under the NIH criteria to be 8.7 ± 2.0%. This value was slightly higher than the 6.5–6.8% obtained in three other prevalence studies that also used the NIH criteria (Diamanti-Kandarakis *et al*., 1999; Asuncion *et al*., 2000; Azziz *et al*., 2004). With the Rotterdam criteria, the prevalence was 11.9 ± 2.4% but increased to 17.8 ± 2.8% when imputed data were included. There have been no dedicated prevalence studies using the Rotterdam criteria but the un-imputed prevalence obtained in this study was equivalent to that in a self reported questionnaire on PCOS in Caucasian women which also used these criteria (Lowe *et al*., 2005). Under the AES recommendations, PCOS prevalence was 10.2 ± 2.2%, and 12.0 ± 2.4% with the imputed data. There are no published prevalence studies using the AES criteria for comparative purposes. This study therefore found that the imputed prevalence of PCOS under the Rotterdam criteria was over twice the NIH prevalence obtained in this study, and even the un-imputed prevalence, under both the Rotterdam and AES criteria were approximately twice that of values in previous NIH criteria prevalence studies.

The methodological strengths of this study include it being the largest and only community-based prevalence study of PCOS to be carried out on a primarily ethnically homogeneous Caucasian population. Previous PCOS prevalence studies have relied upon convenience samples, for example, recruiting women through publicity campaigns (Diamanti-Kandarakis *et al*., 1999), university employee medical examinations (Knochenhauer *et al*., 1998; Azziz *et al*., 2004) or blood donors (Asuncion *et al*., 2000). These studies were not fully randomized and are unlikely to be representative of the general population, as university employees are likely to be of higher socioeconomic status and blood donors are likely to be healthier than the general population. Unfortunately, no statistical evidence was supplied in these prevalence studies to determine the representativeness of participants. In contrast, our study was inclusive of all women across a whole birth cohort, which would be less likely to discriminate based on socioeconomic and health status and includes statistical comparisons with non-respondents. The study has the additional advantage of being highly inclusive, as women who were pregnant, on the OCP or had hysterectomies were questioned on PCOS symptoms retrospectively. Furthermore, the representation of prevalence under the current diagnostic criteria, as well as each of the criteria separately permits values obtained in this study to be compared with a wider range of other studies. The study population is also comparable to other western populations (Razak *et al*., 2005) in terms of levels of obesity and waist circumference, although the levels reported are lower than those reported in the USA (Flegal *et al*., 2002; Zhu *et al*., 2002). As a consequence, this is one of the most representative studies of PCOS prevalence in Caucasian women to date.

The sampling frame does, however, restrict generalizations to women born in Australia who are of Caucasian extraction and in the 27–34 year age group. The restricted age range was a design feature to compress variability in disease experience due to age, but as a consequence prevalence estimates should be generated for

further age ranges. The restriction of our study to Adelaide residents was unlikely to have biased results greatly as a majority of women born in the study period were still resident in Adelaide (69.3%). The response rate of those agreeing to be interviewed of 52.9% was not unexpected from such an intrusive study but it is possible that this may have affected the representativeness. However, statistical analyses showed no significant difference in demographic characteristics, apart from slightly higher (by 3.3%) socio-economic status, between respondents and non-respondents. The inclusion of OCP users increased the representativeness of the study but this may have also increased the error associated with estimates as the presence of menstrual irregularity and hirsutism relied upon recall and the OCP can reduce hyperandrogenism (Mulders *et al*., 2005), or ovarian volume and follicle number (Somunkiran *et al*., 2007). However, we found no significant difference in the prevalence of PCOS symptoms between OCP and non-OCP users.

Ultrasounds and blood tests were not carried out on women who did not present with clinical symptoms of PCOS, as was also the case in an Asian community study of PCOS prevalence (Kumarapeli *et al*., 2008). There may have been some women in this group who had hyperandrogenemia and polycystic ovaries but literature suggests this is likely to be less than 1% of those with PCOS (Kumarapeli *et al*., 2008). The inability to perform ultrasounds on all participants with one or more PCOS criteria was accounted for by assuming the prevalence of polycystic ovaries was the same in women who did not undergo an ultrasound (Azziz *et al*., 2004). This increased the error associated with estimating the number with PCOS but it is unlikely this would have been great, as anthropometry and PCOS symptomatology among those who did and did not undergo an ultrasound were not significantly different. Despite these limitations, it is argued that this strategy is preferable to the convenience sampling used in previous PCOS prevalence studies of Caucasian populations which introduced bias that could not be quantified or adjusted for in the analysis.

This study suggests that PCOS prevalence under the NIH criteria, in a predominately Caucasian community, is higher than previously believed. However, because the definition and prevalence of PCOS symptoms, particularly hyperandrogenism and ovulatory dysfunction, differ widely between previous studies, the capacity to reconcile prevalence estimates with our study is limited. For example, in defining menstrual dysfunction, oligo-menorrhoea has been variously defined as the number of cycles per year (Knochenhauer *et al*., 1998; Azziz *et al*., 2004) and, or the number of days between cycles (Diamanti-Kandarakis *et al*., 1999; Asuncion *et al*., 2000; Azziz *et al*., 2004). The prevalence of women reporting menstrual dysfunction in our study, defined by current status (or usual status if taking OCP, pregnant etc.), was high (23.8%) and may be due to our definition of menstrual irregularity including those with polymenorrhea or cycles with a variation of 4-days or more (Polson *et al*., 1988; Cresswell *et al*., 1997). However, this value was in a similar range to the study of Azziz *et al*. (2004) who also included polymenorrhea. Furthermore, if those with polymenorrhea had been excluded, the un-imputed prevalence under all three PCOS diagnostic criteria would have only deceased by 0.3%. The identification of clinical hyperandrogenism has also been broad, including the presence of acne or alopecia as criteria in one study (Asuncion *et al*., 2000) and the cut off values used in the mF-G scale to assess hirsutism have varied between six (Knochenhauer *et al*., 1998; Diamanti-Kandarakis *et al*., 1999; Azziz *et al*., 2004) and eight (Asuncion *et al*., 2000). Our study used a mF-G score cut-off of $\geq$8 but if it had been six, a further 13 women would have been identified as having PCOS, increasing the prevalence under the NIH criteria to 10.4% ($n = 76$). As a consequence, it is difficult to compare PCOS prevalence between studies and as Asuncion *et al*. (2000) pointed out, if Diamanti-Kandarakis *et al*. (1999) had considered women with clinical hyperandrogenism, normal androgen levels and oligomenorrhea to have PCOS, as was the case in our study, the prevalence would have been 10.4%. Likewise, if Asuncion *et al*. (2000) had included women with polymenorrhea and the mF-G cut-off had been six instead of eight, the prevalence would have been higher but conversely, it may also have been lower if acne or alopecia had been excluded.

The prevalence of hirsutism in our study was higher (21.2%) than generally reported in the population (DeUgarte *et al*., 2006), as well as three other PCOS prevalence studies (Knochenhauer *et al*., 1998; Asuncion *et al*., 2000; Azziz *et al*., 2004), which may have elevated our PCOS prevalence. However, it was comparable to a Greek prevalence study of 29% (Diamanti-Kandarakis *et al*., 1999) and as our study population had a high proportion of women with Mediterranean ethnicity (29%), this could have contributed to the high prevalence. Alternatively, the lower hirsutism prevalence in other studies may be an artefact of alternate methodologies. The mF-G scores were self-reported in our study to enable the highest participation rate and thereby the broadest representation of the community. If hirsutism were assessed by a third party, stigmatizing symptoms may be underreported, or hair growth prior to hair removal may not be assessed as this can only be self reported. Furthermore, the possibility that self-reporting may have led to higher mF-G scores is not supported by a study that found large discrepancies in mF-G scoring by the patient, physician and research nurse but no evidence of a trend towards any one assessor being consistently higher than another (Wild *et al*., 2005).

This study, as well as others investigating PCOS prevalence has drawn attention to the issue of many people in the community with PCOS remaining undiagnosed. In our study, under the NIH criteria, 43 (68% of 63) women did not have an existing PCOS diagnosis and under the Rotterdam criteria this figure was 60 (69% of 87). Consequently, the opportunity to manage symptoms, as well as instigate proactive health regimes that may minimize long-term illness is not possible in the large number of women with undiagnosed PCOS. Conversely, there were 21 (51% of 41) women with a prior PCOS diagnosis that were not classified as having PCOS in this study under the NIH criteria, and 14 (34% of 41) under the Rotterdam criteria. Therefore, either PCOS was assessed differently when they were previously diagnosed, or the symptoms may have resolved, as can occur particularly following weight loss (Crosignani *et al*., 2003). If this were the case, the prevalence would be higher than figures quoted here. For example, if the 14 women with only a prior PCOS diagnosis were added to the Rotterdam criteria figures in this study, the prevalence would increase to 13.9%; and if imputed values were also included the prevalence would reach levels equivalent to those of polycystic ovary prevalence on ultrasound alone (Polson *et al*., 1988; Cresswell *et al*., 1997).

In conclusion, this study is the largest and arguably the most inclusive and representative study conducted to date, as well as the first to compare all protocols commonly used internationally for the diagnosis of PCOS. The prevalence of PCOS in this study using the NIH criteria was slightly higher than values obtained in other prevalence studies,

and based on the more current Rotterdam or AES criteria, prevalence was far in excess of previously quoted figures using the NIH criteria. However, variability in the application of diagnostic criteria and their cut-offs, together with uncertain sampling frames limits our ability to compare and reconcile the range of existing prevalence estimates with our study. To overcome this, we recommend additional refinements specifying more exact measures within each of the diagnostic criteria for future planned studies. The prevalence results must also be viewed with respect to the current PCOS diagnostic criteria being a classification system for a disorder with unknown aetiology and as such the prevalence will change as more is understood of the disorder and the criteria are subsequently refined. In the interim, to permit temporal comparisons of PCOS prevalence, future prevalence studies should include the prevalence of the individual diagnostic criteria, as well as how each of the criteria was defined. In addition, studies should include a measure of representativeness, such as response rates and statistical comparisons with a standard population.

## Authors' Roles

V.M. and M.D. were the principal investigators of the study and formulated the research question. V.M, M.D. and R.N. contributed to the study design and W.M., V.M., M.D. and D.P. contributed to drafting the article and interpretation of data. K.W. and W.M conducted the analyses. All authors contributed to the critical revision of important intellectual content and final approval of the manuscript.

## Acknowledgements

We gratefully acknowledge the contribution from the women who participated in the study. We are also grateful to Kendall Smith for the study co-ordination and management of the interviews and Nanette Kretschmer for her involvement in tracing the women, and Fred Amato and staff at Repromed for their assistance with blood assays. Also, thank you to the many other staff members involved in the interviews, data base construction and data entry.

## Funding

This project was funded by a grant from the National Health and Medical Research Council of Australia (NHMRC) Strategic Award No. 465455.

## References

Amowitz LL, Sobel BE. Cardiovascular consequences of polycystic ovary syndrome. *Endocrinol Metab Clin North Am* 1999;**28**:439–458.

Asuncion M, Calvo RM, San Millan JL, Sancho J, Avila S, Escobar-Morreale HF. A prospective study of the prevalence of the polycystic ovary syndrome in unselected Caucasian women from Spain. *J Clin Endocrinol Metab* 2000;**85**:2434–2438.

Azziz R, Woods KS, Reyna R, Key TJ, Knochenhauer ES, Yildiz BO. The prevalence and features of the polycystic ovary syndrome in an unselected population. *J Clin Endocrinol Metab* 2004;**89**:2745–2749.

Azziz R, Carmina E, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Futterweit W, Janssen OE, Legro RS, Norman RJ, Taylor AE *et al*. Criteria for defining polycystic ovary syndrome as a predominantly hyperandrogenic syndrome: an androgen excess society guideline. *J Clin Endocrinol Metab* 2006;**91**:4237–4245.

Balen A, Michelmore K. What is polycystic ovary syndrome? Are national views important? *Hum Reprod* 2002;**17**:2219–2227.

Balen AH, Laven JSE, Tan SL, Dewailly D. Ultrasound assessment of the polycystic ovary: international consensus definitions. *Hum Reprod Update* 2003;**9**:505–514.

Broekmans FJ, Knauff EAH, Valkenburg O, Laven JS, Eijkemans MJ, Fauser B. PCOS according to the Rotterdam consensus criteria: change in prevalence among WHO-II anovulation and association with metabolic factors. *BJOG* 2006;**113**:1210–1217.

Chang RJ. A practical approach to the diagnosis of polycystic ovary syndrome. *Am J Obstet Gynecol* 2004;**191**:713–717.

Chen XL, Yang DZ, Mo YQ, Li L, Chen YX, Huang YH. Prevalence of polycystic ovary syndrome in unselected women from southern China. *Eur J Obstet Gynecol Reprod Biol* 2008;**139**:59–64.

Cresswell JL, Barker DJ, Osmond C, Egger P, Phillips DI, Fraser RB. Fetal growth, length of gestation, and polycystic ovaries in adult life. *Lancet* 1997;**350**:1131–1135.

Crosignani PG, Colombo M, Vegetti W, Somigliana E, Gessati A, Ragni G. Overweight and obese anovulatory patients with polycystic ovaries: parallel improvements in anthropometric indices, ovarian physiology and fertility rate induced by diet. *Hum Reprod* 2003;**18**:1928–1932.

Dahlgren E, Friberg LG, Johansson S, Lindstrom B, Oden A, Samsioe G, Janson PO. Endometrial carcinoma; ovarian dysfunction—a risk factor in young women. *Eur J Obstet Gynecol Reprod Biol* 1991;**41**:143–150.

DeUgarte CM, Woods KS, Bartolucci AA, Azziz R. Degree of facial and body terminal hair growth in unselected black and white women: toward a populational definition of hirsutism. *J Clin Endocrinol Metab* 2006;**91**:1345–1350.

Diamanti-Kandarakis E, Kouli CR, Bergiele AT, Filandra FA, Tsianateli TC, Spina GG, Zapanti ED, Bartzis MI. A survey of the polycystic ovary syndrome in the Greek island of Lesbos: hormonal and metabolic profile. *J Clin Endocrinol Metab* 1999;**84**:4006–4011.

Dokras A. Cardiovascular disease risk factors in polycystic ovary syndrome. *Semin Reprod Med* 2008;**26**:39–44.

Flegal KM, Carroll MD, Ogden CL, Johnson CL. Prevalence and trends in obesity among US adults, 1999–2000. *J Am Med Assoc* 2002; **288**:1723–1727.

Hatch R, Rosenfield RL, Kim MH, Tredway D. Hirsutism-implications, etiology, and management. *Am J Obstet Gynecol* 1981;**140**:815–830.

Kauffman RP, Baker TE, Baker VM, DiMarino P, Castracane D. Endocrine and metabolic differences among phenotypic expressions of polycystic ovary syndrome according to the 2003 Rotterdam consensus criteria. *Am J Obstet Gynecol* 2008;**198**:670.e1–670.e10.

Knochenhauer ES, Key TJ, Kahsar-Miller M, Waggoner W, Boots LR, Azziz R. Prevalence of the polycystic ovary syndrome in unselected black and white women of the southeastern United States: a prospective study. *J Clin Endocrinol Metab* 1998;**83**:3078–3082.

Kumarapeli V, Seneviratne RD, Wijeyaratne CN, Yapa R, Dodampahala SH. A simple screening approach for assessing community prevalence and phenotype of polycystic ovary syndrome in a semiurban population in Sri Lanka. *Am J Epidemiol* 2008;**168**:321–328.

Lowe P, Kovacs G, Howlett D. Incidence of polycystic ovaries and polycystic ovary syndrome amongst women in Melbourne, Australia. *Aust N Z J Obstet Gynaecol* 2005;**45**:17–19.

McLennan W. *Census of Population and Housing: Socioeconomic Indexes for Areas (Catalogue No. 2039.0)*. Canberra, Australia: Australian Bureau of Statistics, 1998.

Michelmore KF, Balen AH, Dunger DB, Vessey MP. Polycystic ovaries and associated clinical and biochemical features in young women. *Clin Endocrinol (Oxf)* 1999;**51**:779–786.

Mulders A, Kate-Booij MT, Pal R, De Kruif M, Nekrui L, Oostra BA, Fauser BCJM, Laven JSE. Influence of oral contraceptive pills on phenotype expression in women with polycystic ovary syndrome. *Reprod Biomed Online* 2005;**11**:690–695.

Polson DW, Adams J, Wadsworth J, Franks S. Polycystic ovaries–a common finding in normal women. *Lancet* 1988;**1**:870–872.

Razak F, Anand S, Vuksan V, Davis B, Jacobs R, Teo KK, Yusuf S. Ethnic differences in the relationships between obesity and glucose-metabolic abnormalities: a cross-sectional population-based study. *Int J Obes (Lond)* 2005;**29**:656–667.

Somunkiran A, Yavuz T, Yucel O, Ozdemir I. Anti-Müllerian hormone levels during hormonal contraception in women with polycystic ovary syndrome. *Eur J Obstet Gynecol Reprod Biol* 2007;**134**:196–201.

The Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertil Steril* 2004;**81**:19–25.

Vermeulen A, Verdonck L, Kaufman JM. A critical evaluation of simple methods for the estimation of free testosterone in serum. *J Clin Endocrinol Metab* 1999;**84**:3666–3672.

Wild RA, Vesely S, Beebe L, Whitsett T, Owen W. Ferriman Gallwey self-scoring I: performance assessment in women with polycystic ovary syndrome. *J Clin Endocrinol Metab* 2005;**90**:4112–4114.

Zawadski JK, Dunaif A. Diagnostic criteria for polycystic ovary syndrome: towards a rational approach. In Dunaif AGJ, Haseltine F (eds). *Polycystic Ovary Syndrome*. Boston: Blackwell Scientific, 1992, 377–384.

Zhu S, Wang Z, Heshka S, Heo M, Faith MS, Heymsfield SB. Waist circumference and obesity-associated risk factors among whites in the third National Health and Nutrition Examination Survey: clinical action thresholds. *Am J Clin Nutr* 2002;**76**:743–749.

*Submitted on August 17, 2009; resubmitted on October 13, 2009; accepted on October 15, 2009*