Heritability of Polycystic Ovary Syndrome in a Dutch Twin-Family Study

J. M. Vink, S. Sadrzadeh, C. B. Lambalk and D. I. Boomsma


To subscribe to Journal of Clinical Endocrinology & Metabolism or any of the other journals published by The Endocrine Society please go to: http://jcem.endojournals.org//subscriptions/
Heritability of Polycystic Ovary Syndrome in a Dutch Twin-Family Study

J. M. Vink, S. Sadrzadeh, C. B. Lambalk, and D. I. Boomsma

Department of Biological Psychology (J.M.V., D.I.B.), Vrije Universiteit 1081 BT Amsterdam, The Netherlands; and Division of Reproductive Medicine (S.S., C.B.L.), Department of Obstetrics and Gynaecology, Vrije Universiteit University Medical Centre, 1081 HV Amsterdam, The Netherlands

Background: Polycystic ovary syndrome (PCOS) is one of the most common endocrine disorders among women of reproductive age. There is evidence for a genetic component in PCOS based on familial clustering of cases.

Objective: In the present study, the heritability of PCOS was estimated.

Design/Participants: Data from 1332 monozygotic twins (genetically identical) and 1873 dizygotic twins/singleton sisters of twins (who share on average 50% of their segregating genes) registered with The Netherlands Twin Register were used. PCOS was defined as less than nine menstrual cycles and acne or hirsutism in agreement with the 2003 Rotterdam consensus.

Poly cystic ovary syndrome (PCOS) is one of the most common endocrine disorders among women of reproductive age. The prevalence is estimated at 5–10% (1–3). In 2003 an international consensus on the definition of PCOS was published (4). PCOS is defined as at least two of the following symptoms: oligo- or anovulation, hyperandrogenism (defined as elevated androgen serum levels or hirsutism and/or acne), and polycystic ovaries on ultrasound (4). The finding of polycystic ovaries on ultrasonography was originally the hallmark for the diagnosis of the syndrome but represents a sign of a wide variety of disorders and appears to be a nonspecific finding in approximately 20% of the asymptomatic women (5). The symptoms of PCOS include the consequences of excessive androgen production, anovulation, and consequently infertility (6). Early identification of women at risk for PCOS can have profound implications on prevention of PCOS-associated endocrine disorders.

The pathogenesis of PCOS has not yet been elucidated, but familial clustering suggests genetic involvement. Studies in first-degree relatives of women affected by PCOS clearly indicate genetic influences, but no clear mode of inheritance has been identified (7, 8). A polygenic multifactorial model involving multiple genes is most likely (9). To identify genes playing a role in PCOS, linkage and association analyses were carried out. For example, a study with 37 candidate genes in the known pathways for PCOS showed linkage with the follistatin gene and suggestive linkage with CYP11A (10). Other studies failed to detect any consistent association between PCOS and follistatin (11) or CYP11A (12). Other candidate genes for PCOS are genes involved in the biosynthesis and metabolism of androgens, genes involved in folliculogenesis, and the secretion and action of insulin (9, 10, 13–15).

So far, no clear estimate of the impact of genes, the heritability of PCOS, is available. Twin-family studies are commonly used for this type of investigation. Dizygotic (DZ) twins, like ordinary siblings, on average share 50% of their segregating genes, whereas monozygotic (MZ) twins share all their genes. A higher association for PCOS in MZ twins, compared with DZ twins and siblings, indicates genetic influences. Twin data allow distinguishing between the influence of genetic and environmental factors on phenotypic variation (16). Genetic influences will lead to larger MZ than DZ/sister correlations. Environmental influences can be unique to individuals or can be shared by family members. Environmental influences shared by sisters growing up in the same family will lead to MZ, DZ, and sister-pair correlations of equal size. Unique environmental influences will not cause resemblance among sisters. Using statistical modeling techniques makes it possible to obtain a quantitative estimate of the genetic and environmental influence on PCOS. The aim of this study was to estimate the heritability of PCOS. First, a univariate model including genetic and environmental influences was fitted to data of Dutch twins and sisters. PCOS was defined as less than nine menstrual cycles a year plus acne or hirsutism. In addition, we investigated oligomenorrhea, acne, and hirsutism in a trivariate model (Fig. 1). This allowed us to study whether the three

Results: Results point to a strong contribution of familial factors to PCOS. The resemblance in monozygotic twin sisters (tetrachoric correlation 0.71) for PCOS was about twice as large as in dizygotic twin and other sisters (tetrachoric correlation 0.38). Univariate analyses point to strong contributions of genetic factors to the variance in PCOS. Next, a trivariate genetic analysis of oligomenorrhea, acne, and hirsutism was carried out. This analysis confirmed that the familial component in PCOS is due to genetic factors.

Conclusions: This study demonstrated a large influence of genetic factors to the pathogenesis of PCOS, justifying the search for susceptibility genes. (J Clin Endocrinol Metab 91: 2100–2104, 2006)

Polycystic ovary syndrome (PCOS) is one of the most common endocrine disorders among women of reproductive age. The prevalence is estimated at 5–10% (1–3). In 2003 an international consensus on the definition of PCOS was published (4). PCOS is defined as at least two of the following symptoms: oligo- or anovulation, hyperandrogenism (defined as elevated androgen serum levels or hirsutism and/or acne), and polycystic ovaries on ultrasound (4). The finding of polycystic ovaries on ultrasonography was originally the hallmark for the diagnosis of the syndrome but represents a sign of a wide variety of disorders and appears to be a nonspecific finding in approximately 20% of the asymptomatic women (5). The symptoms of PCOS include the consequences of excessive androgen production, anovulation, and consequently infertility (6). Early identification of women at risk for PCOS can have profound implications on prevention of PCOS-associated endocrine disorders.

The pathogenesis of PCOS has not yet been elucidated, but familial clustering suggests genetic involvement. Studies in first-degree relatives of women affected by PCOS clearly indicate genetic influences, but no clear mode of inheritance has been identified (7, 8). A polygenic multifactorial model involving multiple genes is most likely (9). To identify genes playing a role in PCOS, linkage and association analyses were carried out. For example, a study with 37 candidate genes in the known pathways for PCOS showed linkage with the follistatin gene and suggestive linkage with CYP11A (10). Other studies failed to detect any consistent association between PCOS and follistatin (11) or CYP11A (12). Other candidate genes for PCOS are genes involved in the biosynthesis and metabolism of androgens, genes involved in folliculogenesis, and the secretion and action of insulin (9, 10, 13–15).

So far, no clear estimate of the impact of genes, the heritability of PCOS, is available. Twin-family studies are commonly used for this type of investigation. Dizygotic (DZ) twins, like ordinary siblings, on average share 50% of their segregating genes, whereas monozygotic (MZ) twins share all their genes. A higher association for PCOS in MZ twins, compared with DZ twins and siblings, indicates genetic influences. Twin data allow distinguishing between the influence of genetic and environmental factors on phenotypic variation (16). Genetic influences will lead to larger MZ than DZ/sister correlations. Environmental influences can be unique to individuals or can be shared by family members. Environmental influences shared by sisters growing up in the same family will lead to MZ, DZ, and sister-pair correlations of equal size. Unique environmental influences will not cause resemblance among sisters. Using statistical modeling techniques makes it possible to obtain a quantitative estimate of the genetic and environmental influence on PCOS. The aim of this study was to estimate the heritability of PCOS. First, a univariate model including genetic and environmental influences was fitted to data of Dutch twins and sisters. PCOS was defined as less than nine menstrual cycles a year plus acne or hirsutism. In addition, we investigated oligomenorrhea, acne, and hirsutism in a trivariate model (Fig. 1). This allowed us to study whether the three.
variables are indicators of a single latent unobserved trait (PCOS).

**Subjects and Methods**

**Subjects**

This study is part of an ongoing twin family study on health-related behavior in participants of The Netherlands Twin Register (17). With an interval of 2–3 yr, twins and family members receive mailed surveys. For the purpose of this study, data from the 2000 survey were used (18). Zygosity was based on (longitudinal) questionnaire data or, when available, DNA typing (n = 572 females). Agreement between zygosity based on questionnaire data and zygosity based on DNA results is 96%.

In total, 4236 females participated in the 2000 survey. Spouses of male twins (n = 265) and half-siblings (n = 17) were excluded. When data on menstrual cycle were missing (n = 726), it was not possible to classify the subject for PCOS. Data on zygosity were missing for 15 twins. The remaining data set for the univariate analyses of PCOS contained 3205 females: 1332 MZ twins, 680 DZ (same sex) twins, 474 females from dizygotic opposite sex pairs, and 719 (nontwin) sisters.

**Phenotype definition**

PCOS was defined based on questions about the number of menstrual cycles per year, when not using contraception (with answer categories more than eight, less than nine, less than six, two or less), about suffering from acne/pimples (yes or no) and suffering from hirsutism (yes or no). PCOS was defined as less than nine natural menstrual cycles a year combined with hirsutism or acne. In addition, the survey provided information on date of birth, age at menarche, birth weight, current height and weight, having children, and smoking habits. Characteristics of participants are listed in Table 1.

**Statistical analyses**

Modeling of twin data allows discrimination between phenotypic variance due to genetic factors and environmental factors (19). The additive genetic effects of contributing gene loci are expressed in the additive genetic variance reflecting the narrow-sense heritability of PCOS. Another source of genetic variation is dominance; this is the extent to which the effects of alleles at a locus do not simply add up but reflect nonadditive gene action. Variance caused by shared environmental effects is reflected in common environmental variance. Environmental effects that are not shared between family members result in unique environmental variance ($e^2$). This later estimate also includes measurement error. Therefore, the unique environmental variance is always specified in the model. Phenotypic similarities in MZ twins can be due to common environmental and genetic influences. Unique environmental influences contribute to the differences between MZ twins. DZ twins, like other siblings, share approximately 50% of their genetic makeup. The correlation between their additive genetic values is 0.5, and the correlation between their nonadditive genetic values is 0.25. Common environmental effects contribute similarly to similarities between DZ and MZ twins. Adding singleton sisters of twins and females from non-twin sisters.

![Diagram of the common pathway model. Oligomenorrhea (yes/no); Ac, acne (yes/no); Hi, hirsutism (yes/no); A, additive genetic influences; C, common environment; E, nonshared environment; $A_s$, specific additive genetic influences; $D_s$, specific dominant genetic influences; $E_s$, specific unique environment; parentheses, percentage of the total variance. The total variance for each variable is constrained at 1. For example, the total variance of acne = variance explained by specific additive genetic factors ($72\%$) + specific unique environmental factors ($21\%$) + factor loading of the latent PCOS construct ($0.07$) + additive genetic influences on PCOS ($79\%$) + factor loading of the latent PCOS construct ($0.07$) + unique environmental influences on PCOS ($21\%$) = $72\% + 21\% + 0.07\times 79\% + 0.7\times 21\% = 100\%$.](image)

**Table 1. Characteristics of the study population**

<table>
<thead>
<tr>
<th></th>
<th>Nine or more menstrual cycles (n = 2947)</th>
<th>Less than 9 menstrual cycles (n = 258)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MZ twins</td>
<td>1213</td>
<td>119 (8.9%)</td>
<td></td>
</tr>
<tr>
<td>DZ twins</td>
<td>1073</td>
<td>81 (7.0%)</td>
<td></td>
</tr>
<tr>
<td>Sisters</td>
<td>661</td>
<td>58 (8.1%)</td>
<td>0.540</td>
</tr>
<tr>
<td>Acne</td>
<td>30.0%</td>
<td>27.1%</td>
<td>0.326</td>
</tr>
<tr>
<td>Hirsutism</td>
<td>8.5%</td>
<td>14.5%</td>
<td>0.001</td>
</tr>
<tr>
<td>Age (yr)</td>
<td>29.4</td>
<td>30.1</td>
<td>0.332</td>
</tr>
<tr>
<td>Age at first menarche (yr)</td>
<td>13.1</td>
<td>13.4</td>
<td>0.002</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>169.9</td>
<td>168.4</td>
<td>0.000</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>65.5</td>
<td>65.8</td>
<td>0.681</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>22.7</td>
<td>23.2</td>
<td>0.023</td>
</tr>
<tr>
<td>Birth weight (g)</td>
<td>2657</td>
<td>2697</td>
<td>0.545</td>
</tr>
<tr>
<td>Having children</td>
<td>33.4%</td>
<td>37.2%</td>
<td>0.213</td>
</tr>
<tr>
<td>Current smoker</td>
<td>23.0%</td>
<td>23.0%</td>
<td>0.987</td>
</tr>
</tbody>
</table>
dizygotic opposite sex twin pairs to the study population enhances the statistical power for the estimation of the contribution of genetic and environmental influences (20).

Because the phenotype was a dichotomous variable, a threshold model was used (21). A categorical characteristic such as PCOS is assumed to have an underlying liability, which is continuous and normally distributed in the population. The liability to PCOS is divided into two categories, yes and no, separated by a single threshold. The threshold is obtained from the observed proportions in the two categories. Individuals falling below the threshold do not have PCOS, and those exceeding the threshold do suffer from PCOS.

Information about twin resemblance in liability is given by tetrachoric correlations.

Genetic models were applied to raw ordinal data using the Mx statistical program (22). First, a full model with genetic, common environmental, and unique environmental influences was fitted. Next, the full model was reduced by excluding the genetic or common environment component. The reduced models were compared with the full model by hierarchic χ² tests. The χ² statistic was calculated by subtracting the −2log likelihood of the goodness of fit of a reduced model from the full model. If the reduced model does not describe the data significantly worse than the full model, the reduced model can be considered as the best model. Last, a multivariate model was fitted to the data on oligomenorrhea, hirsutism, and acne. To investigate whether the three variables define a single construct of PCOS, a common pathway model was applied. This model assumes that all three variables are indicators of a single latent unobserved trait (PCOS). The relative importance of the latent trait on the observed variables (oligomenorrhea, hirsutism, and acne) is indicated by the value of the loadings from the latent factor to the observed traits. The variance of the latent factors is decomposed into genetic and environmental components. The variance of the observed traits that is not attributable to the PCOS factor is also composed into genetic and environmental components (Fig. 1) (23).

Results

Characteristics of the study population are listed in Table 1. The prevalence for oligomenorrhea varied from 7.0% in singleton sisters to 8.1% in MZ twins and 8.9% in DZ twins. The prevalences were not significantly different between MZ twins, DZ twins, and singleton sisters (P = 0.540). In total, 8% of the females included in this study reported to have less than nine menstrual cycles in a year. The percentage of females suffering from acne was not different for females with more or less than nine menstrual cycles. In contrast, the females with less than nine menstrual cycles report more often to suffer from hirsutism than the females with nine or more menstrual cycles a year. The two groups also differed significantly in age at menarche and height and body mass index but not in current weight and birth weight. There were 92 women (2.9%) classified as having PCOS. The prevalence was not significantly different for MZ twins, DZ twins, or singleton sisters (P = 0.836).

Table 2 shows the tetrachoric correlations for MZ twin pairs and DZ twin and sister pairs. The MZ correlations are higher than the DZ and sister correlations, suggesting a large genetic influence on all variables. The pattern of correlations for acne, hirsutism, and PCOS suggest additive genetic influences (MZ correlation twice as high as the DZ correlation). For oligomenorrhea, the MZ correlation exceeds more than twice the DZ correlation, suggesting genetic dominance or epistasis.

First, univariate genetic models were fitted to the PCOS data (Table 3). In the full model, 66% of the variance is explained by genetic factors (additive genetic variance), 12% by shared environment factors (common environmental variance), and 29% by unique environmental factors (e²). These results suggested that the role of shared environmental factors is small or nonsignificant. The full model could be reduced to a model including only genetic factors [a² = 72%, 95% confidence interval (CI) 46–88%] and unique environmental factors (e² = 28%, 95% CI 12–54%). Next, we explored whether the shared environmental influence could be dropped (model 2) and whether both the shared environmental influence and the genetic influence could be dropped (model 3). The fit of the submodels was compared with the fit of the full model (χ² = −2LLfull model − −2LLsubmodel) and Δdf = dfull model − dfsub-model) to check whether the submodels fit the data significantly worse. The statistical power of the analyses did not allow distinguishing between model 2 and model 3 (Table 3) because both models did not significantly worsen the fit compared to model 1 (P > 0.05). The best model is a model with the smallest number of parameters necessary to explain the data adequately. The Akaike’s Information Criterion (AIC = χ² − 2df) is a measure of the parsimony of the model and a lower value of AIC indicates a more parsimonious model. The AIC (in Table 3) indicated that the model including additive genetic influences and unique environmental factors was the most parsimonious model. In addition, a common pathway model was fitted to the data. In this model, three variables (oligomenorrhea, hirsutism, and acne) are hypothesized to be indicators of a single latent phenotype (PCOS). The variation in the latent PCOS construct is decomposed into additive genetic, common, and unique environmental sources. The model also allows genetic and environmental parameters for the unique variance for each variable. Figure 1 displays the model and the estimates of the parameters. The total variance of the latent and the observed variables (oligomenorrhea, acne, and hirsutism) is constrained to be 1. The parameter estimates (from the best model) are depicted in Fig. 1. The percentage of the total variance is given between parentheses. This model, which has a larger statistical power to distinguish between genetic and common environmental influences on familial resemblance than the univariate model, clearly shows the importance of genetic factors to PCOS. Heritability is estimated at 79%, and there is no influence of common family environment.

Discussion

To our knowledge this is the first quantitative estimation of the genetic influence on the pathogenesis of PCOS using twin data. The prevalence of PCOS was the same for MZ

<table>
<thead>
<tr>
<th>Variable</th>
<th>MZF</th>
<th>DZF/sisters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oligomenorrhea</td>
<td>0.67 (0.49 to 0.80)</td>
<td>0.07 (−0.19 to 0.34)</td>
</tr>
<tr>
<td>Acne</td>
<td>0.78 (0.69 to 0.84)</td>
<td>0.44 (0.30 to 0.56)</td>
</tr>
<tr>
<td>Hirsutism</td>
<td>0.86 (0.75 to 0.92)</td>
<td>0.28 (0.05 to 0.50)</td>
</tr>
<tr>
<td>PCOS</td>
<td>0.71 (0.43 to 0.88)</td>
<td>0.38 (0.00 to 0.66)</td>
</tr>
</tbody>
</table>

MZF, Monozygotic females; DZF/sisters, dizygotic females twins and non-twin sisters.

TABLE 2. Twin and sibling correlations (95% CI) for oligomenorrhea (less than nine menstrual cycles in a year), acne, hirsutism, and PCOS (defined as less than nine menstrual cycles per year and acne or hirsutism)
twins, DZ twins, and singleton sisters. However, the prevalence in our sample was slightly lower than reported in other population studies (1–3). It is possible that the prevalence of PCOS is lower in families that produced twins. However, the prevalence in female spouses of male twins (in general not coming from twin families) is 2.3% (n = 5 of 215 subjects), which suggests that the low prevalence in our sample is not the result of selecting twin families. The prevalence is in line with a Dutch cross-sectional population-based study in adolescents (being more than 3 yr after menarche), which indicated a prevalence of about 3% of oligomenorrhea combined with signs of hyperandrogenism (24). Whether the low prevalences are due to lifestyle differences in The Netherlands, compared with other countries, or to other factors, is unknown.

We did not include ultrasound criteria in the definition of PCOS but relied on self-reported items on oligomenorrhea, hirsutism, and acne. This is very much analogous to the link between oligomenorrhea and cardiovascular diseases form the Nurse’s Health Study (25). Epidemiological studies have shown that oligomenorrhea, hirsutism, and acne are very good indicators of PCOS in the general population (14).

Our study points to a strong contribution of genetic factors to PCOS and indicates that a model including additive genetic factors and unique environmental factors is the most parsimonious. In this model the variance in the pathogenesis of PCOS is for 72% due to genetic influences. The high heritability was confirmed by the results from the multivariate genetic analysis, which has a larger power to detect genetic influences (26).

There are few data regarding PCOS in twins. Jahanfar et al. (27) studied 34 twin pairs (19 MZ and 15 DZ) from an original group of 500 twin pairs that volunteered for PCOS-related ultrasonographic, clinical, and biochemical evaluation. Of the 68 individuals, 33 had PCO ovaries on ultrasound scan, 19 had acne, 12 were hirsute, and seven had oligo/amenorrhea. Eleven pairs were discordant for PCO ovaries on ultrasound scan (five were MZ and six were DZ twins). From this small study they concluded that PCOS is not the result of a single autosomal genetic defect but that PCOS may be the result of polygenic factors or that environmental factors are involved in the pathogenesis of this disorder (27).

Several other studies showed that there appears to be evidence for a genetic component in PCOS based on familial clustering of cases (7, 8, 15). In most family studies, the number of participants is small and PCOS was defined in different ways. In accordance with the Rotterdam consensus (4), we defined PCOS as less than nine natural menstrual cycles a year combined with hirsutism or acne. No ultrasound data were available.

The twin correlations for oligomenorrhea, hirsutism, and acne showed that these variables are largely influenced by genetic factors because the MZ twin correlation was (more than) twice the DZ/sister correlation. Noteworthy, the DZ correlation for oligomenorrhea is less than twice the MZ correlation, indicating that nonadditive genetic influences play an important role. The results are in accordance with other studies. For example, in 2002 Bataille et al. (28) showed that 81% of the variance in acne was attributable to additive genetic effects, whereas the remaining 19% were attributed to unique environmental factors. Previous studies also showed that androgen levels and androgen production rates in humans are under genetic control (29).

In a further step, we modeled the three variables, oligomenorrhea, acne, and hirsutism, in an independent pathway model. This model confirmed our finding with the univariate analyses: the latent variable (PCOS) was highly influenced by genetic variance (79%) and unique environmental influence (21%). Shared environmental influence did not contribute to the latent variable. The results point to the importance of genetic involvement in PCOS and justify the continuous effort to trace the responsible genes. This has been relatively unsuccessful so far, probably due to a number of reasons. One major problem remains: the definition of the phenotype. For example, the most recently developed guideline allows for no less than four possibly distinct phenotypes (3). Another problem is the underlying complexity of the disorder. The metabolic nature of PCOS with combined dysregulation of carbohydrate and fat metabolism and abnormal steroid hormone secretion (9) has led to the suggestion of numerous candidate genes. For a detailed update on this, we referred to several recently published overviews (10–12). Additional difficulties are the lack of a clear male phenotype and PCOS being a major cause of female infertility (9). Finally, environmental factors are also of importance. Environmental factors, i.e., weight gain, may trigger the development of PCOS in predisposed women (10–12). The environmental factors may vary between populations and may actually themselves include a genetic component (30).

Currently a prevailing view is that genetic compounds account for disturbed regulation of ovarian androgenic activity and that environmental circumstances that are of influence on glucose/insulin household predominantly act by aggravating the syndrome through hyperinsulinism and insulin resistance (10–12).

In summary, a number of studies have shown familial
aggregation for PCOS, but less was known about the magnitude of a genetic effect, and the putative PCOS genes have not yet been identified. The present study points to a strong contribution of genetic factors to the pathogenesis of PCOS, justifying further search for these susceptibility genes.

Acknowledgments

We thank Dr. R. Homburg for his contribution to the manuscript.

Received July 6, 2005. Accepted October 3, 2005.

Address all correspondence and requests for reprints to: J. M. Vink, Department of Biological Psychology, Vrije Universiteit, van de Boe-echorststraat 1, 1081 BT Amsterdam, The Netherlands. E-mail: jm.vink@psv.vu.nl.

This work was supported by The Netherlands Organization for Scientific Research (NWO 985-10-002) and the Vrije Universiteit Medical Centre (VUMc Grant 3500-724675).

References


JCEM is published monthly by The Endocrine Society (http://www.endo-society.org), the foremost professional society serving the endocrine community.