Plasminogen activator inhibitor-1 4G/5G and 5,10-methylene-tetrahydrofolate reductase C677T polymorphisms in polycystic ovary syndrome

Pantelis G. Bagos

Department of Informatics with Applications in Biomedicine, University of Central Greece, Lamia 35100, Greece

1Correspondence address. E-mail: pbagos@ucg.gr

ABSTRACT: Polycystic ovary syndrome (PCOS) is a heterogeneous condition with unknown etiology and is considered to be the most common endocrine disorder in women of reproductive age. Two meta-analyses are presented here concerning the association of Plasminogen Activator Inhibitor 1 (PAI-1) 4G/5G polymorphism and the methylene-tetrahydrofolate reductase (MTHFR) C677T polymorphism with the risk of developing PCOS. Seven studies were included concerning PAI-1 (1538 cases, 710 controls) and six studies concerning MTHFR C677T (223 cases, 392 controls). Overall, a significant association was found for PAI-1, with the odds ratio (OR) for 4G carriers versus 5G homozygotes being equal to 1.600 (95% CI 1.052, 2.434) with strong evidence for dominant inheritance. There was however a large between-studies variability (I² = 67.3%). No evidence was found for association of MTHFR C677T polymorphism with PCOS (OR for the TT + CT versus CC comparison equal to 0.940 with 95% CI 0.561, 1.575). No evidence of publication bias was found in these meta-analyses. PAI-1 4G/5G polymorphism seems to be associated with the risk of developing PCOS. Further studies are needed in order to investigate the etiologic mechanism behind this association, as well as the interrelations with other components of the metabolic syndrome (hypertension, diabetes, etc.).

Key words: polycystic ovary syndrome / PAI-1 / MTHFR / meta-analysis / genetic epidemiology

Introduction

Polycystic ovary syndrome (PCOS) is considered to be the most common endocrine disorder in women of reproductive age (Azziz et al., 2004; Ehrmann, 2005; Pasquali and Gambineri, 2006). PCOS is a heterogeneous condition characterized by chronic anovulation, irregular menses, infertility, hyperandrogenism and insulin resistance (Ehrmann, 2005; Norman et al., 2007). The prevalence of PCOS is estimated to be ~7% in the general population in populations of Caucasian (Diamanti-Kandarakis et al., 1999; Asuncion et al., 2000; Azziz et al., 2004) or African–American origin (Knochenhauer et al., 1998), although in some Asian populations the prevalence is reported to be significantly lower (Chen et al., 2008). The fundamental defect that causes PCOS remains elusive and it is believed to be multifactorial in origin, where environmental factors are acting in a genetic background, resulting in a broad spectrum of reproductive, as well as metabolic, defects (Ehrmann, 2005; Diamanti-Kandarakis, 2008). Although infertility is the most common disorder associated with the syndrome (Hunter and Sterrett, 2000; Elghblawl, 2007), women with PCOS are also more likely to develop components of the metabolic syndrome such as diabetes, obesity, hypertension and dyslipidemia (Diamanti-Kandarakis and Economou, 2006; Essah et al., 2007), which in turn are major risk factors for cardiovascular disease (Dokras, 2008).

Various genetic markers have been implicated in predisposition to PCOS; however, no single variant has conclusively and repeatedly been associated with the syndrome (Diamanti-Kandarakis and Piperi, 2005; Luque-Ramirez et al., 2006; Urbanek, 2007). Type I Plasminogen Activator Inhibitor (PAI-1) is one of the primary regulators of the fibrinolytic system in vivo (Sprengers and Kluit, 1987). Over expression of PAI-1 may lead to impaired fibrinolytic system activity and increase the incidence of thrombotic events. A −675 4G/5G sequence polymorphism in the promoter of the PAI-1 gene has been correlated with levels of plasma PAI-1. It has been suggested that the 4G allele results in higher activity than the 5G allele, because in addition to the binding site for the transcriptional activator, the latter also contains a binding site for a transcriptional repressor. In the absence of bound
repressor, the basal level of PAI-1 transcription is increased (Eriksson et al., 1995). The 4G allele of PAI-1 has been recently linked to venous thromboembolism (Tsantes et al., 2007b), coronary disease (Ye et al., 2006) but not to ischemic stroke (Tsantes et al., 2007a, b). Methylenetetrahydrofolate reductase (MTHFR) is one of key the enzymes involved in folate metabolism and it catalyzes the irreversible conversion of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate. The latter constitutes the major circulating form of folate, which acts as a methyl donor for the remethylation of homocysteine to methionine. Two common polymorphisms in the MTHFR gene, C677T (A122Val) and A1298C (Glu429Ala), have been shown to decrease the enzymatic activity resulting in elevated plasma homocysteine levels (Frost et al., 1995), and these were subsequently linked to coronary and peripheral artery disease and further implicated in many diseases. A definite role has been suggested by large meta-analyses for hypertension (Qian et al., 2007) and stroke (Banerjee et al., 2007), whereas inconclusive evidence have been presented for coronary disease (Lewis et al., 2005), diabetic nephropathy (Zintzaras et al., 2007) and retinopathy (Zintzaras et al., 2005).

Since the publication of the first works reporting the association of PAI-1 4G/5G and MTHFR C677T polymorphisms with PCOS (Glueck et al., 1999a, b), several attempts have been made to delineate these associations, but the results are controversial. Concerning PAI-1, some studies have provided strong evidence for the association of the 4G allele with the risk of developing the syndrome (Glueck et al., 1999b; Diamanti-Kandarakis et al., 2004), whereas others have provided no overall evidence for an association (Walch et al., 2005; Karadenez et al., 2007). Concerning MTHFR, several replication studies have failed to provide evidence for association (Sills et al., 2001; Tsanadis et al., 2002; Orio et al., 2003; Palep-Singh et al., 2007). In this work, a comprehensive literature search and a meta-analysis was performed in order to investigate the association of these polymorphisms with the development of PCOS.

Materials and Methods

Retrieval of published studies

A comprehensive electronic search of PubMed and Scopus was conducted until April 2008. The following keywords or combination of them were used: plasminogen activator inhibitor (PAI-1), methylene-tetrahydrofolate reductase (MTHFR), ‘polymorphism’, ‘allele’ etc. coupled with the term ‘Polycystic Ovary Syndrome’. The electronic investigation was supplemented by the assessment of the references of published studies and the manual search of abstracts from conference meetings.

Studies were included in the analysis if: (i) they examined the contribution of the either MTHFR C677T or the PAI-1 4G/5G polymorphisms to the development of PCOS and (ii) they provided adequate data to calculate an estimate of relative risk (RR) comparing cases and controls. To avoid selection bias, published manuscripts were considered for review without any language or quality restrictions (Stroup et al., 2000; Pan et al., 2005). Furthermore, to eliminate bias resulting from the ‘grey literature’ (Conn et al., 2003), it was decided that studies published in conference proceedings or as short abstracts (if any) should be included. Information regarding details of publications, study design, case definitions, patients’ characteristics, distribution of genotypes and genotyping procedures was recorded onto a structured form.

Statistical analysis

Odds ratio (OR) was the metric of choice in all contrasts assessed. The between-study heterogeneity was evaluated using the chi-square-based Cochran’s Q statistic (Petitti, 1994) and the inconsistency index (I²) (Higgins et al., 2003). Data were combined using and random-effects models (DerSimonian and Laird, 1986) with inverse-variance weighting. In case of heterogeneity, random-effects models are more appropriate as they estimate the between-study variance (ν²); however, when heterogeneity is absent, random- and fixed-effects methods coincide. Unless stated differently, random effects estimates are reported. A recently proposed methodology for the meta-analysis of gene–disease association studies that avoids the multiple comparisons and tests directly the genetic model was also implemented (Bagos and Nikolopoulos, 2007). Deviations from Hardy–Weinberg equilibrium (HWE) were calculated by the chi-squared method, and sensitivity analyses were performed for assessing the impact of studies that deviate significantly (Trikalinos et al., 2006).

Publication bias or other small studies-related bias was assessed by the Begg’s rank correlation method (Begg and Mazumdar, 1994), the Egger’s fixed-effects regression method (Egger et al., 1997) and its random-effects analogue (Thompson and Sharp, 1999). In an attempt to identify potential influential studies, we calculated the effects estimates by removing an individual study each time and then checked if the overall significance of the estimate or of the heterogeneity statistic was altered. Cumulative meta-analysis (Lau et al., 1992, 1995) was performed in order to identify a possible trend of the overall OR over years (Ioannidis and Trikalinos, 2005). For all analyses performed here, the software Stata 8 (StataCorp) was used. Statistically significant results were regarded as those with a P-value <0.05.

Results

PAI-1 4G/5G polymorphism

The literature search concerning the association of PAI-1 4G/5G polymorphism with PCOS resulted in 13 studies in total. After examination of the full-text articles, seven eligible studies remained (Table I). One additional study (Glueck et al., 1999a) was excluded since it contained data for the same control population reported in a larger study (Glueck et al., 1999b). Some other studies that were initially identified during the search, did not contain case–control genotype data in order to be included in the meta-analysis, or were case-reports and reviews (data not shown). In total, the seven eligible studies provided data for 1538 cases and 710 controls. The pooled allele frequency for the 4G allele was 54.09% for cases and 45.76% for controls. In all

Results

PAI-1 4G/5G polymorphism

The literature search concerning the association of PAI-1 4G/5G polymorphism with PCOS resulted in 13 studies in total. After examination of the full-text articles, seven eligible studies remained (Table I). One additional study (Glueck et al., 1999a) was excluded since it contained data for the same control population reported in a larger study (Glueck et al., 1999b). Some other studies that were initially identified during the search, did not contain case–control genotype data in order to be included in the meta-analysis, or were case-reports and reviews (data not shown). In total, the seven eligible studies provided data for 1538 cases and 710 controls. The pooled allele frequency for the 4G allele was 54.09% for cases and 45.76% for controls. In all studies except from one (Diamanti-Kandarakis et al., 2004), the control populations were found to be under HWE.

The application of the logistic regression-based model for meta-analysis of case–control association studies (Bagos and Nikolopoulos, 2007), revealed an OR of 1.598 (95% CI: 1.248, 2.046) for the 4G4G versus 5G5G genotypes and an OR of 1.678 (95% CI 1.129, 2.495) for the 4G4G versus 5G5G genotypes. These two ORs were not significantly different (P = 0.779) and thus a dominant mode of action was proposed by the method. Similar results were also suggested by the genetic model-free approach (Minelli et al., 2005) as well as by using the general multivariate framework that was proposed recently (Bagos, 2008). In these analyses, the ratio of the two log ORs (denoted by λ) took values ranging from 0.81 to 0.90,
suggesting clearly a dominant model of inheritance. The overall $I^2$ from the logistic regression model was found equal to 62.4%, suggesting a considerable between-studies heterogeneity. The genetic model was found to be marginally inhomogeneous across studies ($P = 0.057$).

Since a dominant genetic model was suggested, a rational choice would be to pool the 4G4G and the 4G5G genotypes and compare them against 5G5G. The combined OR using the method of DerSimonian and Laird (1986) was equal to 1.600 (95% CI 1.052, 2.434) with no between-studies variability. The logistic regression model was found equal to 62.4%, suggesting a considerable between-studies heterogeneity. The genetic model was found to be marginally inhomogeneous across studies ($P = 0.057$).

Since a dominant genetic model was suggested, a rational choice would be to pool the 4G4G and the 4G5G genotypes and compare them against 5G5G. The combined OR using the method of DerSimonian and Laird (1986) was equal to 1.600 (95% CI 1.052, 2.434) with a between-studies variance equal to 0.205 and an $I^2$ equal to 67.3%. The results are presented in Fig. 1.

Sensitivity analysis did not reveal a single influential study, since in three out of the seven cases (after removing a study from the analysis) the magnitude of the estimate remained nearly the same although the significance changed. However, this is something expected since the overall 95% CI (1.052, 2.434) was approaching unity. One study that deviated from HWE was among the studies that if removed the overall significance changed. However, the limited number of studies does not allow us to draw conclusions for such a bias. All the three methods for detecting publication or other small study-related bias failed to provide evidence for its presence ($P$-values >0.6 in all cases). The cumulative meta-analysis did not reveal any significant trend of the point estimates over time (data not shown).

### MTHFR C677T polymorphism

The literature search concerning the association of MTHFR C677T polymorphism with PCOS yielded in total 10 studies. Five of the studies that were initially identified during the search, did not contain case–control genotype data in order to be included in the meta-analysis, or were case-reports and reviews (Carlsen et al., 2007). One study (Palep-Singh et al., 2007), provided data for different subgroups (Caucasian, Southern Asian) and thus these were counted as separate studies. The data from the six eligible studies arising from five published works are presented in Table II. In total, the six studies provided data for 223 cases and 393 controls, which are figures rather small for a meta-analysis of six studies. The pooled allele frequency for the 677T allele was 38.24% for cases and 33.02% for controls. All studies were found to have the control population under HWE. The results are presented in Fig. 2.

None of the examined contrasts of genotypes yielded a significant finding. By pooling the T allele carriers together (TT + CT versus CC) the combined OR using the method of Dersimonian and Laird was equal to 0.940 (95% CI 0.561, 1.575) with no between-studies variance and an $I^2$ equal to 0%. Similarly, the comparison of TT versus CT + CC yielded an OR of 0.886 (95% CI 0.506, 1.551), once again with no between-studies variability. The logistic regression model provided similar results (data not shown). The cumulative meta-analysis showed that the first published study that investigated solely this polymorphism (Sills et al., 2001) was the only one suggesting

---

**Table I** The studies identified concerning the association the PAI-1 4G/5G polymorphism with PCOS

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Country</th>
<th>Racial</th>
<th>4G4G genotype</th>
<th>4G5G genotype</th>
<th>5G5G genotype</th>
<th>4G allele frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>descent</td>
<td>Cases</td>
<td>Controls</td>
<td>Cases</td>
<td>Controls</td>
</tr>
<tr>
<td>Glueck et al. (1999b)</td>
<td>1999</td>
<td>USA</td>
<td>Caucasian</td>
<td>38</td>
<td>47</td>
<td>78</td>
<td>102</td>
</tr>
<tr>
<td>San Millan et al. (2004)</td>
<td>2004</td>
<td>Spain</td>
<td>Caucasian</td>
<td>18</td>
<td>8</td>
<td>37</td>
<td>17</td>
</tr>
<tr>
<td>Diamanti-Kandarakis et al. (2004)</td>
<td>2004</td>
<td>Greece</td>
<td>Caucasian</td>
<td>39</td>
<td>13</td>
<td>39</td>
<td>18</td>
</tr>
<tr>
<td>Walch et al. (2005)</td>
<td>2005</td>
<td>Austria</td>
<td>Caucasian</td>
<td>27</td>
<td>32</td>
<td>50</td>
<td>46</td>
</tr>
<tr>
<td>Zhao et al. (2005)</td>
<td>2005</td>
<td>China</td>
<td>Asian</td>
<td>58</td>
<td>16</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td>Glueck et al. (2006)</td>
<td>2006</td>
<td>USA</td>
<td>Caucasian</td>
<td>258</td>
<td>29</td>
<td>460</td>
<td>58</td>
</tr>
<tr>
<td>Karadeniz et al. (2007)</td>
<td>2007</td>
<td>Turkey</td>
<td>Caucasian</td>
<td>22</td>
<td>38</td>
<td>48</td>
<td>41</td>
</tr>
</tbody>
</table>

The detailed genotypes per study are listed along with the general characteristics of the study (year, ethnicity of participants and the reference).

---

**Figure 1** The forest plot describing the meta-analysis for the association of plasminogen activator inhibitor (PAI-1) 4G/5G polymorphism with PCOS. The ORs corresponding to the dominant model are presented (4GSG+4G4G versus 5G5G). Each study is depicted with size inversely proportional to the variance logOR, accompanied by the respective 95% confidence intervals. The combined overall estimate and the weight given to each study were calculated with the method of DerSimonian and Laird. The dotted vertical line indicates the overall estimate, whereas the solid one indicates the null effect (OR = 1).
MTHFR C677T polymorphism, no evidence was found for a link with PCOS, and it is highly unlikely that this association exists. As discussed in the introduction, this polymorphism has routinely been implicated in various diseases, and perhaps the first published study that investigated the role of both the MTHFR and PAI-1 polymorphisms (Glueck et al., 1999b) played an influential for the subsequent research. Nevertheless, there is a possibility that future large studies maybe needed to further investigate this issue.

The exact biological mechanism that could explain the proposed association between PAI-1 and PCOS is currently unknown, but there is evidence suggesting that plasma levels of PAI-1 could be involved. Several studies have already provided evidence that plasma levels of PAI-1 are significantly higher in PCOS patients compared with healthy controls (Velazquez et al., 1997; Atiomo et al., 1998; Orio et al., 2004; Tarkun et al., 2004). These results are also in agreement with a recent study in transgenic mice, where polycystic structures and significantly elevated plasma testosterone levels are observed following an abnormal elevation of plasma PAI-1 levels (Devlin et al., 2007). Two of the studies included in this meta-analysis (Diamanti-Kandarakis et al., 2004; Glueck et al., 2006), provided additionally data for PAI-1 plasma levels for the different genotypes. In agreement with the work of Eriksson et al. (1995), they found that 4G carriers have significantly higher PAI-1 plasma levels compared with 5G homoyzogotes, and thus a rational explanation for the proposed dominant genetic model of inheritance is suggested. If more data were available, these results would have been incorporated in the meta-analysis. In such a case, we could have followed the so-called Mendelian randomization approach (Minelli et al., 2004) and try to model in a multivariate analysis the pairwise association of genotype–phenotype (4G/5G polymorphism with PAI-1 levels) along with the genotype–disease association (4G/5G with PCOS). It is thus clear that future studies addressing these issues should be conducted.

Meta-analysis is an acceptable methodology suitable for the dealing with genetic-association data, since in the majority of the cases, the risk associated with a particular variant ranged between 1.1 and 1.5. In such cases, the individual studies are usually small and underpowered and thus, unable to provide a definite answer even in the case where a true association exists. Thus, meta-analysis can effectively combine data from several studies increasing the statistical power (lower type II error rate). An alternative would be the design of large genetic association studies possessing the available statistical

### Table II The studies identified concerning the association the MTHFR C677T polymorphism with PCOS

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Country</th>
<th>Racial descent</th>
<th>TT genotype</th>
<th>CT genotype</th>
<th>CC genotype</th>
<th>T allele frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sills et al.</td>
<td>2001</td>
<td>USA</td>
<td>Caucasian</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Glueck et al.</td>
<td>1999</td>
<td>USA</td>
<td>Caucasian</td>
<td>14</td>
<td>119</td>
<td>23</td>
<td>89</td>
</tr>
<tr>
<td>Tsanadis et al.</td>
<td>2002</td>
<td>Greece</td>
<td>Caucasian</td>
<td>12</td>
<td>20</td>
<td>14</td>
<td>19</td>
</tr>
<tr>
<td>Orio et al.</td>
<td>2003</td>
<td>Italy</td>
<td>Caucasian</td>
<td>16</td>
<td>17</td>
<td>41</td>
<td>38</td>
</tr>
<tr>
<td>Palep-Singh et al.</td>
<td>2007</td>
<td>UK</td>
<td>Asian</td>
<td>14</td>
<td>9</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Palep-Singh et al.</td>
<td>2007</td>
<td>UK</td>
<td>Caucasian</td>
<td>11</td>
<td>10</td>
<td>12</td>
<td>5</td>
</tr>
</tbody>
</table>

The detailed genotypes per study are listed along with the general characteristics of the study (year, ethnicity of participants and the reference).

### Discussion

To the best of the author’s knowledge, this is one of the first (if not the only) meta-analysis performed in order to investigate the genetic component of PCOS. Two polymorphisms were analyzed, the PAI-1 4G/5G polymorphism and the MTHFR C677T polymorphism. Evidence was provided that the PAI-1 4G/5G polymorphism is significantly associated with the risk of developing PCOS, since carriers of the 4G allele have an increased risk (OR = 1.6) of developing the syndrome compared with the 5G homozygotes. Concerning the MTHFR C677T polymorphism, no evidence was found for a link with PCOS, and it is highly unlikely that this association exists. As discussed in the introduction, this polymorphism has routinely been implicated in various diseases, and perhaps the first published study that investigated the role of both the MTHFR and PAI-1 polymorphisms (Glueck et al., 1999b) played an influential for the subsequent research. Nevertheless, there is a possibility that future large studies maybe needed to further investigate this issue.

The exact biological mechanism that could explain the proposed association between PAI-1 and PCOS is currently unknown, but there is evidence suggesting that plasma levels of PAI-1 could be involved. Several studies have already provided evidence that plasma levels of PAI-1 are significantly higher in PCOS patients compared with healthy controls (Velazquez et al., 1997; Atiomo et al., 1998; Orio et al., 2004; Tarkun et al., 2004). These results are also in agreement with a recent study in transgenic mice, where polycystic structures and significantly elevated plasma testosterone levels are observed following an abnormal elevation of plasma PAI-1 levels (Devlin et al., 2007). Two of the studies included in this meta-analysis (Diamanti-Kandarakis et al., 2004; Glueck et al., 2006), provided additionally data for PAI-1 plasma levels for the different genotypes. In agreement with the work of Eriksson et al. (1995), they found that 4G carriers have significantly higher PAI-1 plasma levels compared with 5G homoyzogotes, and thus a rational explanation for the proposed dominant genetic model of inheritance is suggested. If more data were available, these results would have been incorporated in the meta-analysis. In such a case, we could have followed the so-called Mendelian randomization approach (Minelli et al., 2004) and try to model in a multivariate analysis the pairwise association of genotype–phenotype (4G/5G polymorphism with PAI-1 levels) along with the genotype–disease association (4G/5G with PCOS). It is thus clear that future studies addressing these issues should be conducted.

Meta-analysis is an acceptable methodology suitable for the dealing with genetic-association data, since in the majority of the cases, the risk associated with a particular variant ranged between 1.1 and 1.5. In such cases, the individual studies are usually small and underpowered and thus, unable to provide a definite answer even in the case where a true association exists. Thus, meta-analysis can effectively combine data from several studies increasing the statistical power (lower type II error rate). An alternative would be the design of large genetic association studies possessing the available statistical

### Figure 2 The forest plot describing the meta-analysis for the association of methylene-tetra hydrofolate reductase (MTHFR) C677T polymorphism with PCOS. The ORs corresponding to the dominant model are presented (TT + CT versus CC). Each study is depicted with size inversely proportional to the variance of log OR, accompanied by the respective 95% confidence intervals. The combined overall estimate and the weight given to each study were calculated with the method of DerSimonian and Laird. The dotted vertical line indicates the overall estimate, whereas the solid one indicates the null effect (OR = 1).
power to detect a probable association (Ioannidis et al., 2003). In the case of the PAI-1 4G/5G polymorphism, the result of this meta-analysis suggest a rather moderate risk associated with the variant, and thus, the meta-analysis presented here was able to detect it, even if the confidence interval’s lower bound is approaching unity. The results of the sensitivity analysis, where excluding each one of the studies resulted in non-significant estimates in three out of the seven cases, should be interpreted as a justification of the meta-analytic strategy. If we had available for instance, only six or fewer studies, the most likely outcome would be to have non-significant estimates. Consequently, this is the benefit of meta-analysis: the ability to increase the statistical power and provide a definite answer where individual studies fail.

A counter-argument to the above, is the bias that may be introduced in such a meta-analysis suggesting an association that does not exist (type I error). It should be mentioned at this point that every effort has been performed to conduct appropriately the meta-analysis and avoid any possible source of bias. Quality scoring has not been performed since it is considered subjective (Greenland, 1998); non-English articles were identified, retrieved and included in the analysis in order to avoid the local literature bias (Pan et al., 2005); deviations from the HWE were assessed (Trikalinos et al., 2006) and every appropriate test for detecting publication bias or other small study-related bias were performed (Egger et al., 1997; Sterne et al., 2000). Finally, the problem of the early extreme estimates appearing in the meta-analysis of genetic association studies (Ioannidis and Trikalinos, 2005) that correlates with the replication validity of studies in genetic epidemiology (Ioannidis et al., 2001) was evaluated. The last two forms of bias could severely bias the results of a meta-analysis resulting in a false association; however, no such evidence was observed in this work, further strengthening further the validity of the results presented here.

Population-based genetic association studies are themselves subject to limitations such as the limited size of subjects sample, the existence of heterogeneity in PCOS definition, the small size of the effect estimates, the multiple testing and the inappropriate selection of controls (Bloom et al., 2006). Meta-analysis cannot correct all the biases of individual studies but it generates a statistical conclusion with larger power and precision and addresses the heterogeneity of the results (Petiti, 1994). Thus, a significant association could in principle be due to population stratification, or to genotyping errors. Furthermore, a true relationship between an allele and a disease may be attributed to the physically close location of the associated allele to the actual disease susceptibility locus that is currently unknown (Diamanti-Kandarakis and Piperi, 2005). Meta-analysis cannot correct for all the above-mentioned problems, but if it is conducted properly, it minimizes the chances of them influencing the results.

The only unresolved problem in the meta-analysis presented here is that of the large between-studies heterogeneity that was observed. Even if in principle, it does not cancel the results of the meta-analysis (since random effects models were used providing more conservative estimates) but its existence demands an explanation. Heterogeneity is present in the majority of the meta-analyses of genetic association studies (Ioannidis et al., 2003), and usually for explaining it, adjustment using meta-regression for various study-level covariates that may have influenced the results is required. Such covariates in the case of PCOS would be mean age, the mean BMI (body mass index) and other phenotypic characteristics of cases and controls such as mean cholesterol levels, triglycerides or the particular phenotypic expressions of the syndrome (defining the different subtypes) such as the hirsutism score, testosterone and other androgen levels, fasting glucose or insulin levels. Unfortunately, such aggregate study-level covariates could not be extracted from all the eligible studies due to different reporting strategies in each study. Even if they were available, it would be risky to consider them all simultaneously, due to the small number of included studies (Thompson and Higgins, 2002). This issue could be resolved by performing a so-called pooled meta-analysis, where data from individual participants would be combined and the influence of various patient-level covariates could be assessed. Individual patients’ data meta-analysis can resolve many of the problems encountered in a traditional summary-data meta-analysis such as the standardization of case definitions, outcomes, covariates, the inclusion of updated information, better control of confounding, evaluation of alternative genetic models, consistent treatment of subpopulations, assessment of sampling bias and so on. On the contrary, a major drawback is that a much greater commitment of time and resources is required to collect primary data and to coordinate a large collaborative international project requiring an increased overall budget (Steinberg et al., 1997; Ioannidis et al., 2002).

However, a few factors that are highly likely to be sources of heterogeneity can be considered. A first candidate is the different definitions of PCOS used in different studies, since a consensus on PCOS and its subtypes has not been reached yet, or simply because older studies used different standards. For instance, one of the studies was performed on women with ultrasound evidence for polycystic ovaries but with not a definite diagnosis for PCOS (Sills et al., 2001). Another candidate is the inappropriate selection of controls used in the case–control studies. Corollary to these, there is also the problem of interaction with other outcomes related to the metabolic syndrome. Since PCOS is now being considered as another phenotypic expression related to the metabolic syndrome such as diabetes, dislipidemia, obesity and hypertension, including or excluding patients or controls suffering from the above-mentioned conditions would produce heterogeneity. Interestingly, PAI-1 4G/5G polymorphism has been also implicated for predisposition to diabetes (McCormack et al., 1996; Lopes et al., 2003), obesity (Hoffstedt et al., 2002; Berberoglu et al., 2006) and other symptoms related to PCOS (Lopez-Bermejo et al., 2007), thus providing further support for the results of this meta-analysis. Further studies that address the interaction issues simultaneously are needed in order to determine the etiologic relationship of the proposed association. For instance, case–control studies performed on non-diabetic or lean PCOS patients would provide evidence for the direction of the association, i.e. whether PAI-1 4G/5G polymorphism is an independent risk factor for PCOS or acts through increasing the risk of the metabolic syndrome. Thus, another purpose of the meta-analysis, i.e. to provide directions for future research, is fulfilled within this work.

Acknowledgements

The author would like to thank the three anonymous reviewers and the associate editor for their valuable comments which helped in improving the quality of the manuscript.
References


Bloom MS, Schisterman EF, Hediger ML. Selecting controls is not selecting “normals”: design and analysis issues for studying the etiology of polycystic ovarian syndrome. Fertil Steril 2006;86:1–12.

Carlsten SM, Kjortod S, Vanky E, Romundstad P. Homocysteine levels are unaffected by metformin treatment in both nonpregnant and pregnant women with polycystic ovarian syndrome. Acta Obstet Gynecol Scand 2007;86:145–150.


Glueck CJ, Wang P, Fontaine R, Tracy T, Sieve-Smith L. Metformin-induced resumption of normal menses in 39 of 43 (91%) previously amenorrheic women with the polycystic ovary syndrome. Metabolism 1999a;48:511–519.


Submitted on September 20, 2008; resubmitted on November 16, 2008; accepted on November 18, 2008