PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Diagnosis and management of polycystic ovary syndrome in the United Kingdom (2004-2014), a retrospective cohort study
AUTHORS	Ding, Tao; Baio, Gianluca; Hardiman, Paul; Petersen, Irene; Sammon, Cormac

VERSION 1 - REVIEW

REVIEWER	Renato Pasquali University of Bologna, Italy
REVIEW RETURNED	18-May-2016

GENERAL COMMENTS	The authors planned this study to estimate the incidence and prevalence of PCOS in UK primary care and investigate prescribing patterns before and after a PCOS diagnosis. They conclude that, compared to rates estimated in community samples, the incidence of women presenting in primary care with PCOS diagnoses and features is low. Among the women that do present, only 50% were observed to have a recorded PCOS diagnosis. The data may be of importance, but there are a few methodological aspects that should be considered. Specific comments: • Background: please note that hyperandrogenemia is a fundamental criterion (and biomarker) to define hyperandrogenism in PCOS and that an androgen profile rather than testosterone by itself should be considered, according to the recent literature • Please add full name for AHD • Severe insulin resistant states should also be considered as an exclusion criteria for PCOS, according to the consensus (i.e. the AEPCOS consensus) • The definition of “diagnosed cases” vs that of “probable cases” clearly depends on the Code list. If I understand correctly, “probable cases” are much likely to be PCOS. This is probably due to the fact that ≥ 2 criteria are considered in defining PCOS. Table 1 clearly shows that this was not considered (only 8.3% cases) in defining PCOS in the group of sc “Diagnosed cases”. I wonder whether this represents a bias in the definition of PCOS. In addition, this approach could clearly imply a diagnostic bias. The authors should explain why they decided to use this approach. In fact, if the data related to “diagnosed PCOS” are not considered, all findings including prevalence of PCOS and treatment proportion do not change the final message. • Definitely, the number of 597 in the “diagnosed cases” compared to the 7057 in the “probable cases” needs a specific comment. In particular, it should be clarified why, by using the diagnostic criteria (≥ 2), such a large difference may exist. Does it depend on the code used (see supplement table SI? (for example acne is not so frequent in adult PCOS women, whereas hirsutism is much more common).

	• It should be noted that the prevalence rates are rather low compared to those expected on the basis of epidemiological surveys performed in Europe (>6%) • In the Discussion section you mention that "…the inclusion of "probable cases" may introduce case misclassification as some "probable bases" may not be "true PCOS cases". Based on the data presented here and the criteria used to define PCOS (according to the Consensus papers), it could be that the opposite may be true too.

REVIEWER	Dr. Channa Jayasena Department of Medicine, Imperial College London, UK
REVIEW RETURNED	24-May-2016

GENERAL COMMENTS	This is an interesting study reporting results of the interrogation of a large database of primary care health data. The authors endeavour to determine how common PCOS (diagnosed or probable) is within primary care, and investigate associated factors such as social deprivation, age, year. The authors also report the spectrum of medications given to women with PCOS (diagnosed or probable). I make some suggestions principally aimed to improve its appeal to the non-specialist reader. Strengths: 1. This is a well-written paper which reads clearly. 2. The authors clearly acknowledge that database studies such as this will have 'data noise' e.g. we presume that some of the women given metformin had T2DM as the indication rather than PCOS - nevertheless, I agree with the authors that the data can be interpreted with the caveat of these limitations. Weaknesses / major comments: 1. I am currently unclear about the specific message that the reader (such as a GP) should take from this paper. PCOS commonly misclassified? No consensus on medications given? Please be more explicit. Furthermore, what should be changed in clinical practice / policy / attitudes? If nothing, then it is difficult to justify what this adds. 2. Some of the more specialist medications (e.g. spironolactone) warrant brief explanation for their indication. In the case of spironolactone, its anti-androgen effects are off-licence so important to advice / caution the non-specialist reader that these are likely to be initiated in secondary care. 3. related to point 2, I find the 'velocity' of initiating medications following diagnosis of PCOS intriguing (Fig S1). Presumably, COC, acne related drugs and metformin have such a rapid uptake because they are generally initiated in primary care, which contrasts with clomid, spironolactone. Worth discussing. 4. Since the authors have observed wide variation in prescribing habits for clinicians re PCOS, it would be worthwhile discussing why these differences exist - please consider clinical guidelines / consensus statements, the varied nature of clinical presentation etc 5. I suggest the authors speculate why the incidence may change

	with year and be more common in those most deprived. Minor comment: Please make unit of time in legend of Fig S1 (months) and x axis (days) match Background lines 21. Presumably 'real'... rather than 'read'

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Renato Pasquali
Institution and Country: University of Bologna, Italy
Competing Interests: No conflict of interest

The authors planned this study to estimate the incidence and prevalence of PCOS in UK primary care and investigate prescribing patterns before and after a PCOS diagnosis. They conclude that, compared to rates estimated in community samples, the incidence of women presenting in primary care with PCOS diagnoses and features is low. Among the women that do present, only 50% were observed to have a recorded PCOS diagnosis. The data may be of importance, but there are a few methodological aspects that should be considered.

Specific comments:
• Background: please note that hyperandrogenemia is a fundamental criterion (and biomarker) to define hyperandrogenism in PCOS and that an androgen profile rather than testosterone by itself should be considered, according to the recent literature.

This is an interesting point and one that we considered in developing our case definition. Our decision to include raised testosterone alone as a marker of hyperandrogenism was based on a 2015 publication in Endocrine Practice. In this publication the American Association of Clinical Endocrinologists (AACE), the American College of Endocrinology and the Androgen Excess and PCOS Society (AEPCOS) reported that “the value of measuring the levels of androgens other than testosterone in patients with PCOS is relatively low” as only “5% of patients with PCOS have an exclusive increase in DHEAS” and “measurements of either 11β-hydroxyandrostenedione or androstenedione reportedly add only a few patients and are thus generally not needed in clinical use”. Despite this, we understand reviewer 1’s concerns that testosterone tests may be unreliable when used as markers of hyperandrogenism and have therefore added a line to the discussion acknowledging this.

• Please add full name for AHD

AHD stands for additional health data, which is a file containing variables such as the body mass index and test results of the patients. We have now specified this abbreviation in the manuscript.

• Severe insulin resistant states should also be considered as an exclusion criteria for PCOS, according to the consensus (i.e. the AEPCOS consensus)

We found that the specific Readcodes for severe insulin resistant states have been used very infrequently (i.e. the code for type A insulin resistance has only been used 70 times from the establishment of the THIN database in 2003), suggesting that severe insulin resistance is a very rare condition and therefore, is unlikely to generate strong impacts on the current incidence/prevalence. Furthermore, excluding patients with severe insulin resistance from PCOS patients is challenging due to the overlap in a range of symptoms (i.e. acanthosis nigricans, hirsutism, oligomenorrhea).

Considering the infrequent use of the relevant codes and the possibility of excluding a great number of overlapping 'probable' cases, we do not think it should be applied as an exclusion criterion.

• The definition of "diagnosed cases" vs that of "probable cases" clearly depends on the Code list. If I understand correctly, "probable cases" are much likely to be PCOS. This is probably due to the fact that ≥ 2 criteria are considered in defining PCOS. Table 1 clearly shows that this was not considered (only 8.3% cases) in defining PCOS in the group of sc "Diagnosed cases". I wonder whether this represents a bias in the definition of PCOS. In addition, this approach could clearly imply a diagnostic bias. The authors should explain why they decided to use this approach. In fact, if the data related to "diagnosed PCOS" are not considered, all findings including prevalence of PCOS and treatment proportion do not change the final message.

We wondered whether there may be a misunderstanding regarding probable cases and diagnosed cases. To clarify this issue it should be noted that upon making a diagnosis of PCOS, or receiving information that a PCOS diagnosis has been made in secondary/tertiary care, general practitioners are likely to record the main diagnosis (i.e. PCOS) in the coded patient records but are unlikely to code all of the features related to the diagnosis (e.g. hyperandrogenemia, menstrual dysfunction). They may record information on features/symptoms in an unstructured text field however we do not have access to this information. We therefore consider women with a specific diagnostic code for PCOS to be the most certain cases. In the case where women have two PCOS features recorded but no specific PCOS diagnostic code (i.e. our probable cases) we are less certain that they have PCOS as we do not know whether PCOS has been considered and ruled out by the GP and/or specialist, or whether the GP has not linked the two features and has therefore not referred the woman to a specialist for diagnosis. We have added text to the manuscript to make this clear to the reader.

• Definitely, the number of 597 in the "diagnosed cases" compared to the 7057 in the "probable cases" needs a specific comment. In particular, it should be clarified why, by using the diagnostic criteria (≥ 2), such a large difference may exist. Does it depend on the code used (see supplement table SI? (for example acne is not so frequent in adult PCOS women, whereas hirsutism is much more common).

As explained in our response to the previous comment the small number of diagnosed cases with >= 2 PCOS features recorded is related to the nature of data recording in UK primary care. We have now clarified this in the manuscript.

• It should be noted that the prevalence rates are rather low compared to those expected on the basis of epidemiological surveys performed in Europe (>6%)

We agree with the reviewer and have added text to the manuscript noting the difference between our rates and those in epidemiological surveys performed in Europe.

• In the Discussion section you mention that "…the inclusion of "probable cases" may introduce case misclassification as some "probable bases" may not be "true PCOS cases". Based on the data presented here and the criteria used to define PCOS (according to the Consensus papers), it could be that the opposite may be true too.

We agree with the reviewer and have added some text discussing the potential for misclassification of diagnosed cases as 'probable' cases.

Reviewer: 2
Reviewer Name: Dr. Channa Jayasena

Institution and Country: Department of Medicine, Imperial College London, UK
Competing Interests: None declared.

This is an interesting study reporting results of the interrogation of a large database of primary care health data. The authors endeavour to determine how common PCOS (diagnosed or probable) is within primary care, and investigate associated factors such as social deprivation, age, year. The authors also report the spectrum of medications given to women with PCOS (diagnosed or probable). I make some suggestions principally aimed to improve its appeal to the non-specialist reader.

Strengths:
1. This is a well-written paper which reads clearly.
2. The authors clearly acknowledge that database studies such as this will have 'data noise' e.g. we presume that some of the women given metformin had T2DM as the indication rather than PCOS - nevertheless, I agree with the authors that the data can be interpreted with the caveat of these limitations.

Weaknesses / major comments:
1. I am currently unclear about the specific message that the reader (such as a GP) should take from this paper. PCOS commonly misclassified? No consensus on medications given? Please be more explicit. Furthermore, what should be changed in clinical practice / policy / attitudes? If nothing, then it is difficult to justify what this adds.

We thank the reviewer for the helpful comments. We have strengthened the messages in our discussion section by including the following:
-PCOS cases are often undiagnosed in primary care since many potential cases with features indicating evidence of PCOS have never been clinically diagnosed. We suggest that when general practitioners identify patients with features indicative of PCOS, a detailed screening (i.e. ultrasound scan) may need to be recommended so patients would not miss the chance for early monitoring (i.e. weight control, life style modification) to avoid any worsening of the condition or the rapid conversion into metabolic disorders.
-Our exploration of prescribing reveals the fact that although there is wide variation in the drugs prescribed for PCOS, metformin and oral contraceptives are the two drugs that have been prescribed most frequently following a confirmed diagnosis. This is in line with the long-term metabolic concerns of this syndrome stated by PCOS consensuses.

2. Some of the more specialist medications (e.g. spironolactone) warrant brief explanation for their indication. In the case of spironolactone, its anti-androgen effects are off-licence so important to advice / caution the non-specialist reader that these are likely to be initiated in secondary care.

Again, this is a very good point and we have added text to the discussion describing the potential for drugs to be prescribed in secondary care and the impact this may have on our results.

3. Related to point 2, I find the 'velocity' of initiating medications following diagnosis of PCOS intriguing (Fig S1). Presumably, COC, acne related drugs and metformin have such a rapid uptake because they are generally initiated in primary care, which contrasts with clomid, spironolactone. Worth discussing.

We have extended our discussion to highlight the differing velocity of treatment initiation across drug types.

4. Since the authors have observed wide variation in prescribing habits for clinicians re PCOS, it would be worthwhile discussing why these differences exist - please consider clinical guidelines /

consensus statements, the varied nature of clinical presentation etc.

In the previous version of the manuscript we highlighted that the variation in prescribing is likely to be related to the presenting symptoms of PCOS and to the lack of clarity on which is the most effective treatment for the condition. The clinical presentation and treatment of PCOS may also change gradually with age; we have therefore added comments to the discussion to highlight this.

5. I suggest the authors speculate why the incidence may change with year and be more common in those most deprived.

The incidence of PCOS increased slightly over the study period however no significant changes in yearly rates were observed. This might reflect the increasing awareness of the syndrome after the establishment of the Rotterdam and AES criteria during the study period.
Women who lived in more deprived areas had a higher incidence of PCOS than those living in the less deprived areas. A possible explanation is that obesity (a factor strongly associated with PCOS) is more prevalent among women living in more deprived areas. Alternatively, these women may consult their GP more frequently than those in less deprived areas, for other morbidities (i.e. type 2 diabetes), and therefore have more opportunity for PCOS to be diagnosed and recorded.
We have added text to the manuscript reflecting these responses.

Minor comment:
Please make unit of time in legend of Fig S1 (months) and x axis (days) match.
Background lines 21. Presumably 'real'... rather than 'read'.
We have made the unit of time in the legend of Fig S1 matched to the x-axis and corrected the wording in lines 21.

VERSION 2 – REVIEW

REVIEWER	Renato Pasquali University of Bologna. S Orsola Hospital, Italy
REVIEW RETURNED	13-Jun-2016

GENERAL COMMENTS	The authors adequately replied to all my comments

REVIEWER	Dr. Channa Jayasena Imperial College London, UK
REVIEW RETURNED	15-Jun-2016

GENERAL COMMENTS	The reviewer comments have been addressed.