Diagnostic Criteria for Polycystic Ovary Syndrome: Pitfalls and Controversies

Marla E. Lujan, MSc, PhD, Donna R. Chizen, MD, FRCSC, Roger A. Pierson, MS, PhD, FEAS, FCAHS

Department of Obstetrics, Gynecology and Reproductive Sciences, College of Medicine, University of Saskatchewan, Saskatoon SK

Abstract

It is estimated that as many as 1.4 million Canadian women may be afflicted with polycystic ovary syndrome (PCOS). Although PCOS is heralded as one of the most common endocrine disorders occurring in women, its diagnosis, management, and associated long-term health risks remain controversial. Historically, the combination of androgen excess and anovulation has been considered the hallmark of PCOS. To date, while these symptoms remain the most prevalent among PCOS patients, neither is considered an absolute requisite for the syndrome. Inclusion of ultrasonographic evidence of polycystic ovaries as a diagnostic marker has substantially broadened the phenotypic spectrum of PCOS, yet much debate surrounds the validity of these newly identified milder variants of the syndrome. Difficulty in resolving the spectrum of PCOS stems from the continued use of inconsistent and inaccurate methods of evaluating androgen excess, anovulation, and polycystic ovaries on ultrasound. At present, there is no clear-cut definition of biochemical hyperandrogenemia, particularly since we depend on poor laboratory standards for measuring androgens in women. Clinical signs of hyperandrogenism are ill-defined in women with PCOS, and the diagnosis of both hirsutism and polycystic ovarian morphology remains alarmingly subjective. Lastly, there is an inappropriate tendency to assign ovulatory status solely on the basis of menstrual cycle history or poorly timed endocrine measurements. In this review, we elaborate on these limitations and propose possible resolutions for clinical and research settings. By stimulating awareness of these limitations, we hope to generate a dialogue aimed at solidifying the evaluation of PCOS in Canadian women.

Résumé

On estime que pas moins de 1,4 millions de Canadiennes pourraient être atteintes du syndrome des ovaires polykystiques (SOPK). Bien que le SOPK soit reconnu comme l'un des troubles endocriniens les plus courants chez les femmes, son diagnostic, sa prise en charge et les risques à long terme pour la santé qui lui sont associés demeurent controversés. Historiquement, la présence simultanée d'un excès d'androgènes et d'une anovulation a été considérée comme étant le signe distinctif du SOPK. À ce jour, bien que ces symptômes demeurent les plus prévalents chez les patientes atteintes du SOPK, aucun d'eux n'est considéré comme une nécessité absolue pour que l'on puisse envisager la présence du syndrome. Bien que l'inclusion de signes échographiques d'ovaires polykystiques à titre de marqueur diagnostique ait considérablement élargi le spectre phénotypique du SOPK, la validité de ces variantes moins graves du syndrome nouvellement identifiées fait l'objet de nombreux débats. L'utilisation continue de moyens hétérogènes et imprécis d'évaluer l'excès d'androgènes, l'anovulation et la présence de signes échographiques d'ovaires polykystiques est à l'origine des difficultés qui accablent la résolution du spectre du SOPK. À l'heure actuelle, il n'existe pas de définition tranchée de l'hyperandrogénémie biochimique, particulièrement en raison du fait que nous nous fions à des normes de laboratoire de faible qualité pour mesurer le taux d'androgènes chez la femme. Les signes cliniques de l'hyperandrogénie sont mal définis chez les femmes atteintes du SOPK et le diagnostic d'hirsutisme et de morphologie ovarienne polykystique demeure, de façon alarmante, subjectif. Enfin, il existe une tendance inappropriée selon laquelle l'on détermine l'état ovulatoire uniquement en fonction des antécédents quant au cycle menstruel ou en fonction de mesures endocriniennes mal planifiées. Dans le cadre de cette analyse, nous élaborons sur ces limites et proposons des solutions possibles pour les milieux cliniques et de recherche. En accentuant la sensibilisation à ces limites, nous espérons générer un dialogue qui permettra de solidifier l'évaluation du SOPK chez les Canadiennes.

J Obstet Gynaecol Can 2008;30(8):671–679

Key Words: Polycystic ovary syndrome, hyperandrogenism, hirsutism, menstruation disturbances, ultrasonography

Competing Interests: None declared.

Received on November 20, 2007

Accepted on February 27, 2008

INTRODUCTION

It is unlikely that either Stein or Leventhal could have anticipated the enormous amount of curiosity and controversy that would stem from their 1935 description of a unique gynaecological condition that would later be designated as PCOS.[1] PCOS was originally described in seven women in whom the syndrome could *at best* be described as the combination of hirsutism, obesity, amenorrhea, and enlarged bilateral polycystic ovaries.[1] Since then, our understanding of PCOS has evolved so far that none of the originally described features is considered to be a consistent finding in PCOS—not even the appearance of numerous tiny ovarian "cysts" for which the syndrome was named.[2] Women with PCOS present most frequently with complaints of infertility, menstrual irregularity, hirsutism, and/or other outward signs of androgen excess such as acne or alopecia.[3] Clues to the diagnosis also include commonly associated metabolic disturbances such as obesity,

insulin resistance, dyslipidemia, and hypertension.[3] Because of these diverse clinical and metabolic manifestations, considerable debate remains regarding what collection of symptoms constitutes a diagnosis of PCOS. At present, there is no agreement on definitive biochemical or imaging markers for the clinical diagnosis of PCOS. Rather, the diagnosis remains one of exclusion.

PCOS is one of the most common endocrine disorders occurring in women. Epidemiological studies have resulted in estimates of prevalence, in women of reproductive age, that range from 6.5% to 8% using biochemical and/or clinical evidence,[4–7] and ultrasound-based studies have reported a prevalence of 20% or more.[8–11] Therefore, in a population of seven million Canadian women aged between 15 and 44 years, as many as 1.4 million women may be afflicted with this disorder.[12] It has been our experience that many, if not most, women are first given a diagnosis of PCOS when they present to a reproductive endocrinologist with infertility, and that in the years following last delivery and continuing to reproductive senescence there is a clear tendency to forgo long-term management of symptoms. Our difficulty in diagnosing PCOS and maintaining long-term follow-up emphasizes the current state of controversy and confusion surrounding diagnostic criteria, patient management, and long-term health risks for PCOS. This condition should invite early diagnosis and intervention because there is considerable evidence that women with PCOS are at increased risk of infertility, dysfunctional uterine bleeding, metabolic syndrome, type II diabetes, and cardiovascular disease.[3] There is also growing evidence that women with PCOS are at increased risk of obstructive sleep apnea, depression, nonalcoholic fatty liver disease, and certain cancers.[13–16]

The purpose of this review is to highlight the current state of controversy surrounding the clinical definition and diagnosis of PCOS. The limitations in methods used to assess clinical, biochemical and ultrasound features of PCOS are discussed with the goal of producing unified methods of evaluation in Canadian women. Suggestions for overcoming these limitations are provided. Our hope is to stimulate collaborative efforts to improve the timely diagnosis of PCOS and facilitate appropriate clinical intervention.

ABBREVIATIONS

AES	Androgen Excess Society
ASRM	American Society for Reproductive Medicine
ESHRE	European Society for Human Reproduction and Embryology
NICHD	National Institute of Child Health and Human Development
PCOS	polycystic ovary syndrome
SHBG	sex hormone binding globulin

LITERATURE SEARCH METHOD AND QUALITY OF EVIDENCE

A search of Medline was conducted using subject headings "polycystic ovary syndrome," "hyperandrogenism," and "menstruation disturbances" on the Ovid database server. Each subject heading and all of its more specific terms were used to retrieve results. Searches were limited to studies conducted in humans and articles published in English from 2002 to 2007. Combining search terms "polycystic ovary syndrome" and "hyperandrogenism" yielded 177 articles, and the combination "polycystic ovary syndrome" and "menstruation disturbances" yielded 100 articles. All articles were examined, and in instances where referenced articles were considered to have essential information, earlier publications were also reviewed. Evidence presented corroborating the phenotypic spectrum of PCOS is derived primarily from descriptive studies, reports of expert committees, and opinions of respected authorities based on clinical experience that are designated as Level III evidence.[17]

DEFINING CONSENSUS CRITERIA OF POLYCYSTIC OVARY SYNDROME

In 1990, the first formal attempt to consolidate a clinical definition of PCOS by the National Institute of Child Health and Human Development resulted in PCOS being defined as the *combined* presence of androgen excess and oligo-anovulation in the absence of all other reasons for anovulatory infertility (Table 1).[18] The NICHD criteria were deliberately listed in order of perceived importance.[19] The use of these criteria defined PCOS as a syndrome whose primary determinant was a derangement in androgen homeostasis with consequent effects on menstrual cyclicity. Ultrasonographic evidence of polycystic ovaries was concluded to be "suggestive" of PCOS but not necessarily diagnostic.[3] This prevailing opinion reflected the paucity of British and European attendees at the meeting to define the NICHD criteria, because ultrasonographic evidence of polycystic ovaries had long been considered definitive evidence of PCOS in the UK and most of Europe.[20]

The NICHD criteria represented a very important first step towards establishing a universally accepted clinical definition for PCOS. However, it is important to recognize that the criteria were based on majority opinion and not clinical trial evidence.[3] In the years that followed, it became apparent that the clinical presentation of PCOS was much more

Table 1. A comparison of diagnostic criteria for polycystic ovary syndrome

1990 National Institute of Child Health and Human Development (NICHD) Guidelines
Patient demonstrates both:
1. Clinical and/or biochemical signs of hyperandrogenism
2. Oligo- or chronic anovulation
Exclusion of other etiologies of androgen excess and anovulatory infertility is necessary.
2003 European Society for Human Reproduction and Embryology and American Society for Reproductive Medicine (ESHRE/ASRM or Rotterdam) Guidelines
Patient demonstrates two of three criteria:
1. Oligo- or chronic anovulation
2. Clinical and/or biochemical signs of hyperandrogenism
3. Polycystic ovaries
Exclusion of other etiologies of androgen excess and anovulatory infertility is necessary.
2006 Androgen Excess Society (AES) Guidelines
Patient demonstrates both:
1. Hirsutism and/or hyperandrogenemia
2. Oligo-anovulation and/or polycystic ovaries
Exclusion of other etiologies of androgen excess and anovulatory infertility is necessary.

variable than that described by the NICHD criteria, and that polycystic morphology of the ovaries was a consistent finding in women demonstrating biochemical and clinical evidence of the syndrome.[11,21–23] In 2003, the European Society for Human Reproduction and Embryology and the American Society for Reproductive Medicine amended the consensus criteria to include polycystic ovaries as a third diagnostic marker and to allow for a diagnosis of PCOS if *two of three* criteria were met (Table 1). These "Rotterdam criteria" were intended to broaden the phenotypic expression of the syndrome and to redefine PCOS as primarily a syndrome of ovarian dysfunction (i.e., one that occurs in the presence of anovulation and/or ovarian dysmorphology).[19]

The Rotterdam criteria are controversial.[24] Fulfilling two of three diagnostic criteria implies that PCOS can be diagnosed in the absence of androgen excess or menstrual irregularity—the very factors that were once considered absolute requisites for the syndrome. While most agree that PCOS exists as a spectrum, it has been difficult to reconcile the absence of androgen excess in the diagnosis. In 2006, the Androgen Excess Society formed a task force to review existing data on the phenotypic expression of PCOS.[24] The AES concluded that although there was good evidence for features of PCOS (e.g., mild insulin resistance and mild ovarian dysfunction) in women with polycystic ovaries, androgen excess, and regular menstrual cycles,[21,25] there was conflicting evidence supporting the presence of such features of PCOS in women with polycystic ovaries and ovulatory dysfunction but without clinical or biochemical signs of hyperandrogenism.[11,26] The AES has proposed a new set of diagnostic criteria that acknowledge the wide prevalence of morphologic polycystic ovaries and the wide heterogeneity of PCOS. They do not, however, recognize a mild variant of the syndrome in which little is known about metabolic status or long-term health risks (Table 1).[24]

ELUCIDATING THE PHENOTYPIC SPECTRUM OF POLYCYSTIC OVARY SYNDROME

Assuming a broad spectrum, PCOS can be categorized into four main phenotypes (Table 2).[27] These categories are useful in clinical practice because health risks have been defined for at least two subtypes, and this dictates careful evaluation of metabolic disturbances for women with frank or classic PCOS.[27] However, in research settings, subdividing PCOS into more discrete categories is imperative if we are to clearly define incidence, degree of symptomology and health risks among all variants of PCOS (Table 3).[28–32] In the most comprehensive study aimed at evaluating the phenotypic spectrum of PCOS, Diamanti-Kandarakis et al. showed hyperandrogenic and anovulatory phenotypes of PCOS to be the most insulin resistant—irrespective of BMI or central adiposity.[33] Their approach to elucidating differences among PCOS phenotypes involved a number of physicians and ultrasonographers working cooperatively to diagnose and evaluate a large study population. Inclusion of each study participant rested strictly on agreement between at least two physicians that symptoms and signs of PCOS were apparent. Cut-off levels for biochemical

Table 2. Four major clinical phenotypes of polycystic ovary syndrome

	Frank PCOS	Classic PCOS	Ovulatory PCOS	Mild PCOS
Biochemical/clinical hyperandrogenism	✓	✓	✓	
Chronic anovulation	✓	✓		✓
Polycystic ovaries	✓		✓	✓
Prevalence[33, 64, 65]	46–71%	7–40%	7–18%	7–16%
Long-term health risks	known	known	unknown	unknown

hyperandrogenism were carefully established from a large population of non-hirsute, regularly menstruating women with proven ovulatory cycles. Lastly, an independent ultrasonographer interpreted all transvaginal ultrasound recordings. While the efforts of these investigators were exemplary, it is likely that future attempts to substantiate these findings for different ethnic populations will be hampered by the lack of accuracy and reliability that is apparent in the evaluation of PCOS features.[34,35]

EVALUATING ANDROGEN EXCESS

Hyperandrogenemia

It is estimated that 60% to 80% of women with PCOS demonstrate elevated circulating androgen levels.[24] However, the actual prevalence of hyperandrogenemia among women with PCOS is debatable since there is no definitive agreement on (1) which androgen(s) should be measured, (2) when and how often they should be measured, (3) normal androgen levels in women, and (4) which analytical techniques should be used.[35,36] The appropriateness and reliability of analytical methods has been a major issue in this debate. The Endocrine Society recently appealed to laboratories worldwide to refine their methods of assessing androgens in women.[37] Cost-effective, direct commercial assays for measuring androgens in serum perform well in the male range but poorly in the lower female range, yet can be substantially improved following serum extraction and purification.[37] Nevertheless, clinical laboratories continue to avoid these costly extraction processes and report normal ranges that are so broad that women with hyperandrogenemia and/or severe clinical hirsutism are included.[27]

Serum levels of free testosterone, and not total testosterone, are more frequently elevated in women with PCOS. Serum free testosterone is therefore considered to be the most sensitive biochemical marker supporting a diagnosis of PCOS.[38] Measurements of total testosterone in serum include a portion bound to SHBG. Because PCOS is often associated with decreased SHBG levels (because of obesity and insulin resistance), increased testosterone clearance does not allow for an accurate reflection of increased androgen production.[39] The most accurate method of measuring free testosterone in serum is equilibrium dialysis, yet very few laboratories have adopted this standard because the process is complicated, expensive and labour-intensive.[40] Assays that directly measure free testosterone have for the most part been abandoned since they are notoriously inaccurate.[41] Instead, surrogates such as the free androgen index (FAI: the ratio of total testosterone to SHBG multiplied by 100) or bioavailable testosterone (BioT: law of mass action involving total testosterone, SHBG and albumin) have become widely accepted, but only if the necessary assays have been validated. Nevertheless, it should be recognized that the unavoidable drawback to FAI and BioT relates to their use of SHBG measurements. Unlike direct measurements of free testosterone, FAI and BioT are not markers of hyperandrogenemia independent of obesity.[39]

Hirsutism

Hirsutism is the most common clinical manifestation of hyperandrogenism in women.[42] Approximately 60% to 70% of women with PCOS have hirsutism.[24] Hirsutism is defined as excessive terminal hair growth that takes on a male pattern distribution.[42] The clinical assessment of hirsutism is overtly subjective, and it is therefore prudent in clinical assessment to consider the patient's perception of unwanted hair growth in addition to the perceived rate and timing of hair growth onset. Rapid and sudden appearance of thick pigmented hair suggests the presence of an androgen-secreting neoplasm, whereas hair growth in PCOS tends to be more gradual and commonly occurs following cessation of long-term hormonal contraceptive use.[42] Age and ethnicity significantly also influence hair growth due to genetic variances in 5 α-reductase activity.[43] Asian women and adolescents can therefore be expected to demonstrate less terminal hair growth than older women of other ethnic groups (e.g., Mediterranean or East Indian origin).[43]

In an effort to reduce some of the subjectivity associated with the clinical evaluation of hirsutism, excessive hair growth in women is generally quantified by the Ferriman-Gallwey scoring system. This system grades

Table 3. Discrete clinical phenotypes represented by consensus guidelines for polycystic ovary syndrome

1990 NICHD Guidelines
1. Hirsutism, hyperandrogenemia, and oligo-anovulation
2. Hirsutism and oligo-anovulation
3. Hyperandrogenemia and oligo-anovulation
2003 ESHRE/ASRM (Rotterdam) Guidelines
1. Oligo-anovulation, hirsutism, hyperandrogenemia, and polycystic ovaries
2. Oligo-anovulation, hirsutism, and hyperandrogenemia
3. Oligo-anovulation, hirsutism, and polycystic ovaries
4. Oligo-anovulation, hyperandrogenemia, and polycystic ovaries
5. Oligo-anovulation and hirsutism
6. Oligo-anovulation and hyperandrogenemia
7. Oligo-anovulation and polycystic ovaries
8. Polycystic ovaries, hirsutism, hyperandrogenemia, and regular cycles
9. Polycystic ovaries, hirsutism, and regular cycles
10. Polycystic ovaries, hyperandrogenemia, and regular cycles
2006 AES Guidelines
1. Hirsutism, hyperandrogenemia, oligo-anovulation, and polycystic ovaries
2. Hirsutism, hyperandrogenemia, and oligo-anovulation
3. Hirsutism, oligo-anovulation, and polycystic ovaries
4. Hyperandrogenemia, oligo-anovulation, and polycystic ovaries
5. Hirsutism and oligo-anovulation
6. Hyperandrogenemia and oligo-anovulation
7. Hirsutism, hyperandrogenemia, polycystic ovaries, and regular cycles
8. Hirsutism, polycystic ovaries, and regular cycles
9. Hyperandrogenemia, polycystic ovaries, and regular cycles

terminal hair growth on a scale from 0 to 4 (i.e., no terminal hairs to extensive terminal hair growth) on 11 anatomical sites and uses the sum of nine areas to generate an overall hirsutism score. Scores of ≥ 8 or ≥ 5 have been commonly accepted as abnormal,[44,45] although recently a score as low as 3 was proposed as the upper limit of normal.[46] The Ferriman-Gallwey scoring system has been criticized not only for cut-off scores that are debatable, but also for being too general.[35] A total score assumes hair growth on the trunk or thighs is equivalent to hair growth on the face or chest, yet no evidence exists to support this assumption. Moreover, an overall score does not allow for a description of hair growth patterns, which can be variable among women.[34,35,43] Lastly, and probably most importantly, Ferriman-Gallwey scoring does not adequately overcome the subjectivity associated with assessing hirsutism. In an evaluation of the level of interobserver variability associated with ascribing hirsutism scores, Wild et al. showed that agreement in scores was alarmingly poor when three clinicians evaluated the same 21 women, noting a discrepancy of up to 10 points between scores.[47] Since unreliable methods for measuring and interpreting serum androgens make clinical signs of hyperandrogenism sufficient evidence for androgen excess,[24] the use of an unreliable scoring system cannot be viewed as inconsequential.

Acne

One third of women with PCOS, particularly younger women, demonstrate acne.[24] Androgens participate in the development of acne by stimulating sebum production, thereby providing optimal conditions for bacterial colonization with organisms such as *Propionibacterium acnes*.[48] Scoring systems that classify and/or grade acne severity (i.e., numbers and types of acne lesions) are reliable and widely used in dermatology to facilitate therapeutic decisions and assess response to treatment.[49] We know very little about the severity of acne in PCOS because acne scores are seldom used or reported. Moreover, it is unclear whether the actual prevalence of acne is increased in women with PCOS compared with the population at large.[24] Some form of acne occurs in virtually all teenage girls and in more than one half of women over the age of 25.[50] It is difficult to accept acne

Table 4. Recommendations for reducing variability in the evaluation of polycystic ovary syndrome

PCOS feature	Limitations	Recommendations for clinical practice & research (R)
Androgen excess		
1.Hyperandrogenemia	• Inconsistent evaluation of androgens • Inaccurate normative ranges • Imprecise analytical techniques	Measure testosterone (FAI or BioT) in a single sample, on a morning of day 1–5 of the menstrual cycle, following at least 3 months furlough from hormonal therapies. R: Develop age related reference ranges for androgens from a group of women with regular cycles, no hirsutism or polycystic ovaries; work with local laboratories to optimize normative ranges for the health region; develop mass spectrometry techniques to measure androgens.
2. Hirsutism	• Subjective scoring system	Use a validated scoring system to grade hirsutism. Determine and optimize intra- and inter-observer variability among group members. R: Agreement between 2 observers is necessary to confirm hirsutism for each study subject, determine differences in hair distribution patterns among PCOS subtypes and age/ ethnic-matched controls.
3. Acne and alopecia	• Unconfirmed increase in prevalence or severity	Use validated scoring systems to grade acne and alopecia. R: Report intra- and inter-investigator variability; determine differences in acne severity and incidence among PCOS subtypes and age-matched controls.
Anovulation	• Untimely progesterone measurements • Inappropriate designation of ovulatory status based on normal menses	Measure mid-luteal progesterone levels 7 days before anticipated menses in two consecutive cycles. R: Confirm ovulatory cycles by luteal phase progesterone and ultrasound (whenever possible) in two consecutive cycles in both control and PCOS subjects.
Polycystic Ovaries	• Use of multiple ultrasound (US) criteria • Subjective evaluation of US imaging	Use ESHRE/ASRM US consensus criteria only. Communicate new US criteria to local medical imaging groups. R: Report intra- and inter-observer variability in making the US diagnosis using consensus US guidelines; validate US criteria in accurate study populations; generate US criteria that maximize diagnostic accuracy and minimize observer variability; collaborate with medical imaging specialists.

as a clinical feature of PCOS when it has not been conclusively identified as an abnormal finding.[35]

Alopecia

Women may experience a diffuse pattern of thinning hair over the vertex of the scalp with the frontal hairline commonly preserved.[51] Well-established scoring systems for hair loss in women exist but are seldom used in the evaluation of PCOS. Historically, alopecia was recognized as a symptom of PCOS because it is an androgen-mediated process[51]; however, it is a poor predictor of biochemical hyperandrogenemia, and low serum iron levels and aging are more common causes of hair loss in women.[24,35] While the actual prevalence of alopecia in women with PCOS is relatively low compared with other androgenic symptoms (approximately 5%[24]), an association with polycystic ovaries has been reported,[22] and this observation merits an investigation of whether alopecia is actually increased in prevalence or severity in women with PCOS.

EVALUATING ANOVULATION

Menstrual disturbances in PCOS generally present in the form of oligo-amenorrhea (fewer than eight episodes of menstrual bleeding per year or menses that occur at intervals greater than 35 days).[27] Menstrual irregularity in women with PCOS is primarily the consequence of anovulation. Ovulatory dysfunction may be present in women with PCOS who report regular menstrual cycles.[27] For these reasons, menstrual history alone is insufficient for defining PCOS phenotypes in women who describe having regular cycles. Ovulation must be confirmed, generally by serum progesterone measurements >10 nmol/L taken at a random time during the luteal phase.[52] A single mid-luteal phase measurement >30 nmol/L on cycle day 21 is another cost-effective standard used to confirm the quality of ovulation.[52] These commonly adopted criteria are not without limitations, and the validity of these cut-off values have been argued.[53] Inappropriate timing of progesterone measurements is the most obvious reason for an inaccurate

diagnosis of anovulation.[35] Lack of routine ultrasonography to visualize corpus luteum development is also a factor.[54]

EVALUATING POLYCYSTIC OVARIAN MORPHOLOGY

The current ultrasonography guidelines, supported by the ESHRE/ASRM consensus group,[3] define the polycystic ovary as containing 12 or more follicles measuring 2–9 mm and/or an increased ovarian volume of > 10 cm^3. Unlike previous definitions, this requires no subjective assessment of stromal echogenicity and/or follicle distribution pattern. The cut-off value for increased ovarian volume was based on cumulative evidence reporting a larger mean volume of > 10 cm^3 for polycystic ovaries.[55] The cut-off of ≥ 12 follicles *throughout the entire ovary* was based on a single report demonstrating this value to have 99% specificity and 75% sensitivity in distinguishing between polycystic and normal ovaries.[56] The reproducibility of these values has not been confirmed in repeated studies, and there is evidence to support the postulate that these cut-off values cannot be used to distinguish between women with and without PCOS[57,58] (e.g., 58% of the control subjects in the study of Diamanti-Kandarakis et al.[33] had > 12 follicles per ovary). As a result, many reproductive endocrinologists use their own criteria in clinical practice (e.g., > 20 follicles per ovary[58] or increased stroma/total area ratio[59]), while medical imaging specialists tend to use older criteria because of the slow dissemination of findings among disciplines.[60]

Despite the need for further validation of these criteria, it is important to recognize that the interpretation of ultrasonographic images of polycystic ovaries is alarmingly subjective. In an analysis of 54 scans in which images of polycystic and normal ovaries were duplicated and randomized for evaluation by four observers, Amer et al. showed that observers agreed on a diagnosis of PCOS only 51% of the time, and they agreed with themselves only 69% of the time.[61] In that study, the polycystic ovary was defined as having ≥ 10 follicles (2–8 mm), an ovarian volume ≥ 12 cm^3, and a bright echogenic stroma.[61] The variability demonstrated by the observers indicated that the criteria employed were either too subjective or that the measurements were too insensitive to allow for good agreement.[56] Unfortunately, the extent to which any feature contributed to the subjectivity of the diagnosis was not evaluated, nor has a similar evaluation using the ESHRE/ASRM ultrasound criteria been performed since they were proposed in 2003. If ultrasonographic evidence of polycystic ovaries is to remain a diagnostic criterion for PCOS, then reduction in observer variability is needed.[62,63]

Recommendations

Suggestions for overcoming limitations in the evaluation of androgen excess, anovulation, and polycystic ovaries on ultrasonography are presented in Table 4.[34,35] The recommendations are intended to motivate clinicians to evaluate their methods for diagnosing PCOS carefully in both clinical and research settings. Addressing these issues will ultimately require large, well-controlled research studies. However, thoughtful steps can be taken in the interim to improve the diagnosis of PCOS by first-line practitioners.

CONCLUSION

Polycystic ovary syndrome remains a highly controversial topic because of its undetermined and potentially variable etiology and an undetermined phenotypic spectrum. In clinical and research practice, a conservative and broadly based definition of PCOS is warranted. We believe a conservative diagnosis is more likely to motivate appropriate education, judicious treatment, and long-term follow-up of patients as the actual health risks of these distinct phenotypes are slowly revealed. In a research setting, it is hard to argue that anything other than the Rotterdam criteria be used for diagnosis, because these criteria allow for investigations of all PCOS-related features as well as all possible phenotypic combinations. Having an accepted definition is just the first step to unravelling markers of syndrome severity and predictors of long-term health. The remaining and arguably most important steps, in both clinical and research settings, relate to improving the accuracy and reliability of our methods for evaluating features of PCOS. Our attempts at early diagnosis and intervention will continue to be hindered if these inconsistencies remain ignored or regarded as inherent and insurmountable issues in the evaluation of PCOS.

ACKNOWLEDGEMENTS

This work was supported by a scholarship from the Canadian Institutes of Health Research (CIHR) funded Strategic Training Initiative in Research in Reproductive Health Sciences (STIRRHS) and a Saskatchewan Health Research Foundation (SHRF) Fellowship Award to MEL.

REFERENCES

1. Stein IF, Leventhal ML. Amenorrhea associated with bilateral polycystic ovaries. Am J Obstet Gynecol 1935;29:181–91.
2. Dunaif A, Thomas A. Current concepts in the polycystic ovary syndrome. Annu Rev Med 2001;52:401–19.
3. 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(1):19–25.
4. 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(9):3078–82.

5. Diamanti-Kandarakis E, Kouli CR, Bergiele AT, Filandra FA, Tsianateli TC, Spina GG, et al. A survey of the polycystic ovary syndrome in the Greek island of Lesbos: hormonal and metabolic profile. J Clin Endocrinol Metab 1999;84(11):4006–11.
6. 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(7):2434–8.
7. 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(6):2745–9.
8. Farquhar CM, Birdsall M, Manning P, Mitchell JM, France JT. The prevalence of polycystic ovaries on ultrasound scanning in a population of randomly selected women. Aust N Z J Obstet Gynaecol 1994;34(1):67–72.
9. Botsis D, Kassanos D, Pyrgiotis E, Zourlas PA. Sonographic incidence of polycystic ovaries in a gynecological population. Ultrasound Obstet Gynecol 1995;6(3):182–5.
10. 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(9085):1131–5.
11. Michelmore K, Ong K, Mason S, Bennett S, Perry L, Vessey M, et al. Clinical features in women with polycystic ovaries: relationships to insulin sensitivity, insulin gene VNTR and birth weight. Clin Endocrinol (Oxf) 2001;55(4):439–46.
12. Canadian Socio-economic Information Management System (CANSIM). Population by sex and age group: Statistics Canada, 2007.
13. Tasali E, Van Cauter E, Ehrmann DA. Relationships between sleep disordered breathing and glucose metabolism in polycystic ovary syndrome. J Clin Endocrinol Metab 2006;91(1):36–42.
14. Giudice LC. Endometrium in PCOS: Implantation and predisposition to endocrine CA. Best Pract Res Clin Endocrinol Metab 2006;20(2):235–44.
15. Cerda C, Perez-Ayuso RM, Riquelme A, Soza A, Villaseca P, Sir-Petermann T, et al. Nonalcoholic fatty liver disease in women with polycystic ovary syndrome. J Hepatol 2007;47(3):412–7.
16. Hollinrake E, Abreu A, Maifeld M, Van Voorhis BJ, Dokras A. Increased risk of depressive disorders in women with polycystic ovary syndrome. Fertil Steril 2007;87(6):1369–76.
17. Canadian Task Force on Preventive Health Care. New grades for recommendations from the Canadian Task Force on Preventive Health Care. CMAJ 2003;169(3):207–8.
18. Zawadzki JK, Dunaif, A. Diagnostic criteria for polycystic ovary syndrome: towards a rational approach. Boston: Blackwell Scientific Publications, 1992.
19. Azziz R. Diagnostic criteria for polycystic ovary syndrome: a reappraisal. Fertil Steril 2005;83(5):1343–6.
20. Balen A, Michelmore K. What is polycystic ovary syndrome? Are national views important? Hum Reprod 2002;17(9):2219–27.
21. Carmina E, Lobo RA. Polycystic ovaries in hirsute women with normal menses. Am J Med 2001;111(8):602–6.
22. Cela E, Robertson C, Rush K, Kousta E, White DM, Wilson H, et al. Prevalence of polycystic ovaries in women with androgenic alopecia. Eur J Endocrinol 2003;149(5):439–42.
23. Cresswell J, Fraser R, Bruce C, Egger P, Phillips D, Barker DJ. Relationship between polycystic ovaries, body mass index and insulin resistance. Acta Obstet Gynecol Scand 2003;82(1):61–4.
24. Azziz R, Carmina E, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Futterweit W, et al. Position statement: criteria for defining polycystic ovary syndrome as a predominantly hyperandrogenic syndrome: an Androgen Excess Society Guideline. J Clin Endocrinol Metab 2006; Nov;91(11):4237–45. Epub 2006 Aug 29.
25. Carmina E, Lobo RA. Do hyperandrogenic women with normal menses have polycystic ovary syndrome? Fertil Steril 1999;71(2):319–22.
26. Norman RJ, Hague WM, Masters SC, Wang XJ. Subjects with polycystic ovaries without hyperandrogenaemia exhibit similar disturbances in insulin and lipid profiles as those with polycystic ovary syndrome. Hum Reprod 1995;10(9):2258–61.
27. Norman RJ, Dewailly D, Legro RS, Hickey TE. Polycystic ovary syndrome. Lancet 2007;370(9588):685–97.
28. Chang WY, Knochenhauer ES, Bartolucci AA, Azziz R. Phenotypic spectrum of polycystic ovary syndrome: clinical and biochemical characterization of the three major clinical subgroups. Fertil Steril 2005;83(6):1717–23.
29. Carmina E, Orio F, Palomba S, Longo RA, Lombardi G, Lobo RA. Ovarian size and blood flow in women with polycystic ovary syndrome and their correlations with endocrine parameters. Fertil Steril 2005;84(2):413–9.
30. Belosi C, Selvaggi L, Apa R, Guido M, Romualdi D, Fulghesu AM, et al. Is the PCOS diagnosis solved by ESHRE/ASRM 2003 consensus or could it include ultrasound examination of the ovarian stroma? Hum Reprod 2006;21(12):3108–15.
31. Broekmans FJ, Knauff EA, Valkenburg O, Laven JS, Eijkemans MJ, Fauser BC. PCOS according to the Rotterdam consensus criteria: change in prevalence among WHO-II anovulation and association with metabolic factors. BJOG 2006;113(10):1210–7.
32. Barber TM, Wass JA, McCarthy MI, Franks S. Metabolic characteristics of women with polycystic ovaries and oligo-amenorrhoea but normal androgen levels: implications for the management of polycystic ovary syndrome. Clin Endocrinol (Oxf) 2007;66(4):513–7.
33. Diamanti-Kandarakis E, Panidis D. Unravelling the phenotypic map of polycystic ovary syndrome (PCOS): a prospective study of 634 women with PCOS. Clin Endocrinol (Oxf) 2007;67(5):735–42.
34. Franks S. How good are we at diagnosing polycystic ovary syndrome? Clin Endocrinol (Oxf) 2007;67(6):809–10.
35. Barth JH, Yasmin E, Balen AH. The diagnosis of polycystic ovary syndrome: the criteria are insufficiently robust for clinical research. Clin Endocrinol (Oxf) 2007;67(6):811–5.
36. Fraser IS, Kovacs G. Current recommendations for the diagnostic evaluation and follow-up of patients presenting with symptomatic polycystic ovary syndrome. Best Pract Res Clin Obstet Gynaecol 2004;18(5):813–23.
37. Rosner W, Auchus RJ, Azziz R, Sluss PM, Raff H. Position statement: utility, limitations, and pitfalls in measuring testosterone: an Endocrine Society position statement. J Clin Endocrinol Metab 2007;92(2):405–13.
38. Escobar-Morreale HF, Asuncion M, Calvo RM, Sancho J, San Millan JL. Receiver operating characteristic analysis of the performance of basal serum hormone profiles for the diagnosis of polycystic ovary syndrome in epidemiological studies. Eur J Endocrinol 2001;145(5):619–24.
39. Steck T, Wernze H. Is determination of the "free androgen index" for hormone screening in polycystic ovaries of value? [article in German]. Gynakol Geburtshilfliche Rundsch 1993;33(3):173–9.
40. Rosner W. An extraordinarily inaccurate assay for free testosterone is still with us. J Clin Endocrinol Metab 2001;86(6):2903.
41. 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(10):3666–72.
42. Rosenfield RL. Clinical practice. Hirsutism. N Engl J Med 2005;353(24):2578–88.
43. Williamson K, Gunn AJ, Johnson N, Milsom SR. The impact of ethnicity on the presentation of polycystic ovarian syndrome. Aust N Z J Obstet Gynaecol 2001;41(2):202–6.
44. Ferriman D, Gallwey JD. Clinical assessment of body hair growth in women. J Clin Endocrinol Metab 1961;21:1440–7.

45. Ferriman D, Purdie AW. The aetiology of oligomenorrhoea and/or hirsuties: a study of 467 patients. Postgrad Med J 1983;59(687):17–20.

46. 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(4):1345–50.

47. 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(7):4112–4.

48. Burkhart CN, Gottwald L. Assessment of etiologic agents in acne pathogenesis. Skinmed 2003;2(4):222–8.

49. Strauss JS, Krowchuk DP, Leyden JJ, Lucky AW, Shalita AR, Siegfried EC, et al. Guidelines of care for acne vulgaris management. J Am Acad Dermatol 2007;56(4):651–63.

50. Goulden V, Stables GI, Cunliffe WJ. Prevalence of facial acne in adults. J Am Acad Dermatol 1999;41(4):577–80.

51. Ludwig E. Classification of the types of androgenetic alopecia (common baldness) occurring in the female sex. Br J Dermatol 1977;97(3):247–54.

52. Practice Committee of the American Society for Reproductive Medicine. Optimal evaluation of the infertile female. Fertil Steril 2006;86(5 Suppl):S264–7.

53. Jeffcoate SL. Interpretation of single progesterone measurement in diagnosis of anovulation and defective luteal phase. Br Med J (Clin Res Ed) 1984;288(6420):864.

54. Baerwald AR, Adams GP, Pierson RA. Form and function of the corpus luteum during the human menstrual cycle. Ultrasound Obstet Gynecol 2005;25(5):498–507.

55. Balen AH, Laven JS, Tan SL, Dewailly D. Ultrasound assessment of the polycystic ovary: international consensus definitions. Hum Reprod Update 2003;9(6):505–14.

56. Jonard S, Robert Y, Cortet-Rudelli C, Pigny P, Decanter C, Dewailly D. Ultrasound examination of polycystic ovaries: is it worth counting the follicles? Hum Reprod 2003;18(3):598–603.

57. Baerwald AR, Adams GP, Pierson RA. Characterization of ovarian follicular wave dynamics in women. Biol Reprod 2003;69(3):1023–31.

58. Allemand MC, Tummon IS, Phy JL, Foong SC, Dumesic DA, Session DR. Diagnosis of polycystic ovaries by three-dimensional transvaginal ultrasound. Fertil Steril 2006;85(1):214–9.

59. Fulghesu AM, Ciampelli M, Belosi C, Apa R, Pavone V, Lanzone A. A new ultrasound criterion for the diagnosis of polycystic ovary syndrome: the ovarian stroma/total area ratio. Fertil Steril 2001;76(2):326–31.

60. Weissleder R, Wittenberg J, Harisinghani MG. Primer of Diagnostic Imaging, 2nd ed. St. Louis: Mosby Incorporated; 1997.

61. Amer SA, Li TC, Bygrave C, Sprigg A, Saravelos H, Cooke ID. An evaluation of the inter-observer and intra-observer variability of the ultrasound diagnosis of polycystic ovaries. Hum Reprod 2002;17(6):1616–22.

62. Lujan ME, Peppin AK, Bloski TG, Leswick D, Kriegler S, Pierson RA, et al. Improving inter-observer variability in the evaluation of ultrasonographic features of polycystic ovaries. Proceedings of the 53rd Annual Meeting of the Canadian Fertility and Andrology Society, Halifax, Canada; 2007.

63. Lujan ME, Peppin AK, Dhir A, Pierson RA, Chizen RA. Inter-observer variability in the identification and quantification of ultrasonographic features of polycystic ovaries. Proceedings of The Endocrine Society's 89th Annual Meeting, Toronto, Canada; 2007.

64. Dewailly D, Catteau-Jonard S, Reyss AC, Leroy M, Pigny P. Oligoanovulation with polycystic ovaries but not overt hyperandrogenism. J Clin Endocrinol Metab 2006;91(10):3922–7.

65. Welt CK, Gudmundsson JA, Arason G, Adams J, Palsdottir H, Gudlaugsdottir G, et al. Characterizing discrete subsets of polycystic ovary syndrome as defined by the Rotterdam criteria: the impact of weight on phenotype and metabolic features. J Clin Endocrinol Metab 2006; 91(12):4842–8.