HHS Public Access

Author manuscript

*Fertil Steril*. Author manuscript; available in PMC 2017 May 01.

Published in final edited form as:

*Fertil Steril*. 2016 May ; 105(5): 1322–1329.e1. doi:10.1016/j.fertnstert.2015.12.136.

# Sonographic markers of ovarian morphology, but not hirsutism indices, predict serum total testosterone in women with regular menstrual cycles

**Heidi Vanden Brink, MSc**[1], **Amy D Willis, MSc**[2], **Brittany Y Jarrett, BS**[1], **Annie W Lin, MSc**[1], **Steven Soler, BS**[1], **Siân Best, BA**[1], **Erica L Bender, MSN**[1], **Andrew K Peppin, MD**[1], **Kathleen M Hoeger, MD**[3], and **Marla E Lujan, PhD**[1]

[1]Human Metabolic Research Unit, Division of Nutritional Sciences, Cornell University, Ithaca, NY

[2]Department of Statistical Sciences, Cornell University, Ithaca, NY

[3]Division of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynecology, University of Rochester Medical Center, Rochester, NY

## Abstract

**Objective—**To determine whether sonographic markers of ovarian morphology or male pattern hair growth scores predict androgen levels in women with regular or irregular menstrual cycles.

**Design—**Cross-sectional observational study.

**Setting—**Clinical research unit.

**Patient(s)—**Seventy-six women of reproductive age (18–39 years) were evaluated for male-pattern hair growth (using modified Ferriman-Gallwey scoring system), ovarian morphology (transvaginal ultrasonography) and total serum testosterone (liquid chromatography tandem mass spectrometry).

**Interventions—**None

**Main Outcome Measures—**Regional and total modified Ferriman-Gallwey (mFG) scores, number of follicles per follicle size category, follicle number per ovary (FNPO), ovarian volume (OV), ovarian area (OA), stromal to ovarian area ratio (S/A), stromal echogenicity index (SI), total testosterone (TT) and menstrual cycle length.

**Results—**Neither regional nor total mFG scores correlated with TT concentrations in women with regular or irregular menstrual cycles as judged by the Least Absolute Shrinkage and Selection Operator (LASSO) technique. By contrast, a sonographic marker (FNPO 6–9mm) significantly

**CORRESPONDING AUTHOR:** Marla E Lujan, MSc PhD. Assistant Professor, Human Nutrition. 216 Savage Hall. Division of Nutritional Sciences, Cornell University, Ithaca, NY 14853; Tel.607.255.3153. Fax. 607.255.1033, MEL245@cornell.edu.



Portions of this work were presented at the annual meetings of the Canadian Fertility and Andrology Society (Quebec City, Quebec, Canada, September 11 – 14, 2014) and Androgen Excess and PCOS Society (Lihue, Hawaii, United States, October 22 – 23, 2014).

predicted TT concentrations in women with regular menstrual cycles but not in women with irregular menstrual cycles.

**Conclusion—**Sonographic markers of ovarian morphology, but not hirsutism scores, predicted TT levels. However, the predictive value of ovarian morphology for TT differed by menstrual cycle status. That sonographic markers did not predict androgen levels in a diverse cohort of women with cycle irregularity, suggests the potential for distinct variations in ovarian morphology for androgenic and non-androgenic types of cycle irregularity. Overall, our findings support that an assessment of ovarian morphology may be helpful in reflecting TT levels.

### Keywords

Ultrasonography; ovaries; hirsutism; testosterone; oligoamenorrhea

## INTRODUCTION

Biochemical assessments of androgens in women are controversial (1). Commercial assays for serum testosterone yield inconsistent results (2, 3) while direct measurements of free testosterone are technically challenging (3, 4) and influenced by metabolic status (5, 6). Overall, mass spectrometry assay performance is improved compared to commercially available products – albeit modest inter-lab differences in estimates have been reported (7, 8). To that end, a national effort to standardize androgen measurements across centers is underway and promises to have significant impact on future estimates of androgens in women (1, 3). Given these challenges in biochemical assessments of androgens in women, additional measures to evaluate androgen status are also needed.

Male-pattern hair growth is the most commonly accepted clinical indicator of androgen status (9). Atypical hair growth is commonly quantified using the Modified Ferriman-Gallwey (mFG) scoring system, which rates hair growth on nine androgen-sensitive regions of the body using a 0–4 scale (10). The utility of all nine regions in the prediction of androgen excess remains a topic of debate (11–15). This notion has merit since a more focused approach involving a subset of regions with the greatest sensitivity to androgen levels could help to obviate some of the subjectivity associated with hirsutism scoring (11, 13, 16). Uncertainty in the utility of hirsutism scores stems from findings of poor inter-rater agreement in hirsutism scores (6, 17, 18) as well as the known influence of age (15), race and ethnicity (19–22), and adiposity (6, 17, 18) on male-pattern hair growth. While hirsutism has shown better sensitivity for biochemical hyperandrogenism compared to acne or alopecia (23, 24), its specificity is low since idiopathic hirsutism occurs in 5–15% of the general population (reviewed in (22)) and in up to 50% of all mild hirsutism cases studied (25). The advent of more standardized approaches to measure serum androgens provides an opportunity to revisit the utility of hirsutism scores to reflect androgen levels.

In view of the improved resolution afforded by the latest imaging systems (26–28), there is growing evidence supporting an expanded role for ovarian ultrasonography in the clinical evaluation of androgen excess. We (29), using mass spectrometry, and others (8, 30–32), using commercially available assays, have shown that ovarian markers, such as antral follicle count, ovarian size and stromal characteristics are significantly associated with total

testosterone concentrations. In the case of antral follicle count, this is consistent with the concept that small antral follicles are a significant source of androgen production by the ovaries (33). Whether the relationship between ovarian morphology and androgen production is conserved between women with and without regular menstrual cycles is uncertain. Studies to date have been limited primarily to women with hyperandrogenic causes of anovulation (31) and those undergoing assisted reproduction (34, 35). Given that androgen excess can manifest in women with regular menstrual cycles and is associated with increased risk for cardiometabolic disease (36–38), there is relevance in identifying clinical markers of androgen excess in women with both regular and irregular menstrual cyclicity.

The primary objective of this research was to assess the ability of mFG scores and sonographic markers of ovarian morphology to predict total testosterone levels in women. To this end, we enrolled women with regular and irregular cycles in order to assess any impact of menstrual cycle status on these relationships. We hypothesized that a sonographic marker from the ovary, the main site of androgen production, would significantly predict total testosterone concentrations whereas a marker reflecting a consequence of androgen action, such as a hirsutism index, would have limited ability to predict total testosterone. In this way, ultrasonography could represent an additional tool to predict androgen status.

## MATERIALS AND METHODS

### Study Subjects

Seventy-six women from the general population (Tompkins County, NY and surrounding area) were recruited to the study between 2009 and 2014. Participants were recruited using targeted ads seeking both healthy women of reproductive age with regular menstrual cycles (every 21–35 days) and women with a history of irregular or absent menstrual cycles (>35 days) with the goal of recruiting equal numbers of women in each group. Women who were 18–39 years of age with clear visualization of at least one ovary on ultrasonography were eligible to participate. Exclusion criteria included: evidence of reproductive aging as gauged by the Stages of Reproductive Aging (39) and/or premature ovarian insufficiency, use of hormonal therapy, insulin sensitizers, and/or statins in the previous 2 months, participation in a drug trial within the last 30 days, pregnancy, lactation, hyperprolactinemia, diabetes or uncontrolled thyroid disorders. Written, informed consent was obtained from all participants. This study was approved by the Institutional Review Board at Cornell University (Ithaca, NY).

### Study Procedures

Participants were evaluated at Cornell University's Human Metabolic Research Unit for the following: (1) an assessment of self-reported menstrual cycle history to determine the extent of any menstrual cycle disturbance, (2) a physical examination to assess height, weight, vitals and male-pattern terminal hair growth; (3) a transvaginal ultrasound scan to characterize ovarian morphology; and (4) fasting blood tests. Menstrual cycle history was taken at the time of enrollment as part of establishing eligibility to participate in the study. A baseline ultrasound scan was also conducted at this initial visit to corroborate visualization of ovaries and stage of cycle. A physical exam, repeat ultrasound scan and blood draw were

then conducted on the same day during a follow-up early morning study visit to the research unit. In the case of women with regular menstrual cycles, biochemical and sonographic evaluations occurred during a follow-up visit scheduled between Days 2 and 7 of their cycle. In women with irregular cycles, none demonstrated a dominant follicle or corpus luteum at the initial ultrasound scan or during the follow-up study visit (approximately 1 – 2 days later). In this way, all measures for this group were standardized to a time point when no dominant follicle or corpus luteum was observed.

## Hirsutism Scoring

Male-pattern hair growth was assessed on nine regions of the body using the mFG scoring system (10). Regions were ranked on a scale of 1–4 with 1 representing sparse terminal growth and 4 indicating frank male-pattern hair growth. When no terminal hair growth was present, a rank of 0 was assigned to that body region. Participants had not been asked to refrain from mechanical removal of hair prior to attending study visits. As such, each of the nine areas was assessed jointly by the investigator and participant to better gauge any impact of cosmetic measures which may have been taken to reduce the visibility of terminal hair growth and scores were based on combined visual inspection and on participant self-report and follow-up questions. In the event that grading of hair growth fell between rank categories, a value of 0.5 was assigned.

## Ultrasonography Measurements

Participants were evaluated by transvaginal ultrasonography by one of two experienced ultrasonographers. Participants with regular menstrual cycles were examined between days 2 and 7 of their cycle and women with cycle irregularity were examined at a time when there was no evidence of a morphologically dominant follicle (>10mm) or recent ovulation (i.e. active corpus luteum and/or endometrial thickening). Whole ovaries were imaged from their inner to outer margins in the longitudinal plane using a 6–12MHz transducer on a GE Voluson E8 Expert System (Milwaukee, WI, USA). Digital images of each ovary were archived for off-line analysis using Santesoft DICOM Editor (©Emmanouil Kannellopoulus, Athens, Greece). Images were de-identified such that offline evaluation of sonographic endpoints was conducted in a blinded manner.

The largest single cross-sectional view of each ovary was evaluated by a single investigator for ovarian volume (OV), ovarian area (OA), stroma-to-total area ratio (S/A), and stromal index (SI) as previously described (29). In short, OA was calculated using the equation $\pi/4$ (transverse diameter) × (longitudinal diameter) and OV calculated based on the equation $\pi/6$ (transverse diameter) × (anteroposterior diameter) × (longitudinal diameter). S/A ratio was calculated by dividing the traced stromal region of the ovary (providing the stromal area) by the trace of the periphery of the ovary (providing the ovarian area). Each of these tracings also produced a mean pixel echogenicity of each region. The SI was determined by dividing the mean stromal echogenicity by the mean echogenicity of the entire ovary. In this way, any adjustment in gain during the ultrasonographic examination was corrected. Ultrasonographic cineloops of each entire ovary were evaluated for the number and diameter of all antral follicles present using the grid-system approach (27). All follicle populations were reported as the mean of both ovaries (FNPO). Physiological cohorts of interest included: 1) number

of 2–5mm follicles per ovary, 2) number of 6–9mm follicles per ovary, and 3) number of 2–9mm follicles per ovary. Based on an intra-class correlation coefficient analysis, the level of agreement among 3 observers for FNPO was 0.89. Values reported for all sonographic endpoints represent the mean of both ovaries. In the event where a regressing corpus luteum was still detected in the early follicular phase (N=2) or the participant had 1 ovary (N=1), data for a single ovary was reported.

### Testosterone Assay

Fasting concentrations of total testosterone were measured by liquid chromatography tandem mass spectrometry (LC/MS/MS) with a sensitivity of 2ng/dL at a clinical chemistry lab participating in the Centers for Disease Control and Prevention (CDC) Hormone Standardization (HoSt) Program (Brigham Research Assay Core, Boston, MA, USA). As part of the HoSt Program, quality control samples provided by the CDC were run every 3 months to confirm that the bias in quality control samples was <6.4%.

### Statistical Analysis

Variables that best predicted total testosterone were determined using the Least Absolute Shrinkage and Selection Operator (LASSO) technique (40) using the lars (41) and covTest (42) packages in R (R, Version 3.2.0, Vienna, Austria). LASSO permits variable selection in the high-dimensional multiple linear regression context (i.e. a large number of predictors relative to the number of observations). LASSO identifies least-squares estimates of each parameter's regression coefficient subject to a constraint on the sum of the absolute value of the coefficient estimates. By tightly constraining the size of the coefficient estimates, and then relaxing constraint, the order in which the predictors enter the model may be used to indicate relative order of explanatory power. This approach allows identification of the most predictive covariates of the dependent variable based on a limited sample size, even when some collinearity is present. The LASSO procedure was performed with the following covariates: 1) nine regional hirsutism scores and the total mFG score for a total of 10 covariates in the model, and 2) ultrasonographic markers 2–5mm FNPO, 6–9mm FNPO, S/A, SI, and OV for a total of five covariates in the model. Due to high collinearity, 2–9mm FNPO and OA were excluded as they destabilized the model. Total testosterone values were log-transformed. P values reflect the hypothesis test of a significant improvement in predictive power when the first variable enters the model (43). Between-group comparisons were conducted using Mann Whitney U-tests (continuous variables) and Fisher's Exact (categorical variables) (SPSS Statistics V23, Armonk, NY, USA). Descriptive statistics ($5^{th}$, Median and $95^{th}$ percentile) of clinical and sonographic endpoints are provided for each cohort and for each cohort based on the increasing levels (quartiles) of testosterone.

### Ethical Considerations

This study was approved by the Institutional Review Board at Cornell University. All interactions with human participants occurred at the Human Metabolic Research Unit within the Division of Nutritional Sciences (Ithaca, NY, USA). Informed, written consent was obtained from all study participants.

# RESULTS

## Characteristics of the Study Population

Clinical and sonographic characteristics of the study participants are listed in Table 1. Women reporting regular menstrual cycles were similar in age, body mass index, age at menarche and total testosterone levels compared to women with irregular menstrual cycles. Total mFG scores were similar among groups – albeit hair growth scores on the upper lip and chin were higher in women with irregular cycles compared to those with regular cycles. Women with irregular menstrual cycles had more antral follicles (2–5mm and 2–9mm), as well as larger ovaries compared to those with regular cycles. The proportion of women identified as ever-users of hormonal contraception did not differ in women with regular or irregular menstrual cycles women (58% versus 74%, P=0.147). While none of the participants used hormonal agents in the 2 months preceding the study, 12 participants reported use of hormonal contraception in the year prior to enrollment. Of these 12, 6 had terminated use 6 months prior (3 were women with regular cycles and 3 had irregular cycles). The proportion of women across different races was also similar between women with regular (Caucasian 68%, Black 14%, Asian 19%, American Indian 0%) and irregular menstrual cycles (Caucasian 78%, Black 11%, Asian 11%, American Indian 0%; P=0.655). Similarly, ethnicity did not differ between groups (Regular Cycles: Hispanic 11% and non-Hispanic 70%, and Irregular Cycles: Hispanic 5% and non-Hispanic 81%; P=0.512).

## Clinical and Sonographic Predictors of Total Testosterone

Covariates most predictive of total testosterone as judged by LASSO analysis are listed in Tables 2 and 3. In a model involving mFG scores on nine body regions and total mFG score, none of the covariates significantly predicted total testosterone in women with regular or irregular cycles. By contrast, in a model including ovarian markers, the number of 6–9mm follicles was significantly correlated with total testosterone in the LASSO model in women with regular menstrual cycles (Table 3; P=0.001). In women with irregular cycles, sonographic markers did not significantly predict total testosterone (Table 3). This was despite larger ovarian size (P=0.038) and numerically higher follicle counts (2–9mm FNPO) in women with higher TT levels (Supplemental Table 1).

# DISCUSSION

This study was conducted to assess the predictive value of hirsutism scores and ovarian sonographic markers for total testosterone levels, using an accurate and reliable assay for androgen status. We noted that regional mFG scores were more predictive of total testosterone compared to the total mFG score. However, neither regional nor total mFG scores predicted total testosterone to any significant degree in either women with regular or irregular menstrual cycles. By contrast, we showed that sonographic markers of ovarian morphology were significant and strong predictors of total testosterone. This was the case only in women with regular menstrual cycles. Thus, our hypothesis that aspects of ovarian morphology – by virtue of being a site of androgen production – could reliably predict androgen levels in women of reproductive age was only partially supported.

Our finding that chin and lower abdominal hair growth scores were among the top predictors of total testosterone complements previous reports that facial and lower abdominal hair are the most common places for male-pattern hair growth in women of reproductive age (9, 10, 16, 44). Others have proposed that lower abdominal hair growth, in addition to facial hair, serve as an adequate proxy for overall hirsutism scoring in both women from the general population (10, 14, 15) and women presenting with clinical or biochemical evidence of androgen excess (12, 13, 45). These studies were initiated in part, to examine the need for scoring all nine regions comprising the mFG scale, which can increase the likelihood for error in scores and can be deemed invasive by certain patients. Indeed, total mFG score emerged second-last (ninth) in a series of covariates predicting total testosterone in women with regular menstrual cycles and never entered the model predicting testosterone levels in women with irregular cycles. Together, these data provide increasing evidence of the poor predictive value and sensitivity of total mFG scores to reflect current androgen levels in women of reproductive age. This is consistent with hirsutism best reflecting the action of bioavailable testosterone on susceptible areas of the skin over time rather than some aspect of current androgen production.

Our finding that total mFG scores were neither predictive of, or related to, total testosterone levels is supported by some (24, 29, 44, 46) but not all studies (9, 47, 48). Differences in findings might relate to the inclusion of more severe manifestations of hirsutism in women with overt androgen excess disorders by some studies (12, 13, 45) as well as differences in the racial and ethnic groups investigated (12–15). In our study, the 95$^{th}$ percentiles for total mFG scores were 13 and 15 for women with regular and irregular menstrual cycles, meaning that severe cases of androgen excess were not represented. Hence, our study was not able to evaluate the sensitivity of more severe cases of hirsutism to predict total testosterone. Also, our current study was comprised of mainly non-Hispanic Caucasian women, which is consistent with our local demographic (Tompkins County, NY). Differences in hair growth scores across diverse races and ethnicities are generally accepted (reviewed in (22)). However, differences in hair growth within Caucasian populations have been noted, with individuals from Northern Europe having lower hirsutism scores compared to their counterparts from North America (32). Because we did not collect more comprehensive information on race and ethnicity, we are unable to fully appreciate the racial and ethnic origins of our study population, which likely span numerous global regions. Last, the use of different testosterone assays may have also contributed to differences among studies. Currently, there is support that the use of LC/MS/MS may be expected to yield more accurate results compared to other techniques (2, 3). However, it is worth noting that there are several studies employing commercial testosterone assays that have also reported lack of associations among androgen levels and hirsutism scores (24, 44). Hence, the contribution of technical differences among studies is not fully known.

Unlike assessments of regional and total hirsutism scores, sonographic markers predicted total testosterone levels. In women with regular menstrual cycles, the number of larger follicles (6–9mm) – which physiologically corresponds to follicles recruited to a wave-like cohort – emerged as a significant predictor of total testosterone levels. These findings complement recent reports by Jeppesen et al. who showed that 5–8mm follicles represent a physiologically informative follicular pool compared to smaller follicle populations by

contributing the majority of circulating anti-Müllerian hormone levels (49). Healthy 6–9mm follicles would be expected to be a significant source of testosterone given their potential for preferential growth and development which is a steroidogenic-dependent process (33). In the case of women with cycle irregularity, there is increased likelihood for this follicular pool to represent follicle arrest, disordered steroidogenesis and/or atresia(50, 51). This may have served to obviate direct associations between total testosterone and aspects of ovarian morphology in women with irregular cycles. That said, our findings contrast with previous work, which demonstrated that smaller follicles measuring 2–5mm were positively correlated with total testosterone while follicles 6–9mm were negatively correlated with total testosterone, in an unadjusted analysis (31). Our data differ in that assessments by Dewailly and colleagues were limited to women with hyperandrogenic anovulation. In addition, our study was conducted more than 10 years later and used newer imaging technology. Given the higher resolution afforded by newer technology, it may not be wholly appropriate to make direct comparisons of follicle size populations among studies (52). Nevertheless, it was surprising that follicle populations were predictive of total testosterone in women with regular menstrual cycles but not in those with irregular cycles. When we explored bivariate associations among follicle populations and testosterone levels, we confirmed that total testosterone correlated with the number of 6–9mm ($\rho$=0.452; p=0.004), but not with the 2–5mm follicles ($\rho$=−0.025, p=0.881) in women with regular menstrual cycles (correlation analyses not shown). Likewise, we noted that OV ($\rho$=0.422, p=0.01), but neither the number of 2–5mm ($\rho$=0.274, p=.100) nor 6–9mm ($\rho$=0.187, p=0.267) follicles, correlated with total testosterone in women with irregular menstrual cycles. When data for women with irregular menstrual cycles were stratified by increasing testosterone levels, we saw the expected increase in ovarian size, and numerically higher follicle populations, with higher androgen levels. Together, these data point to the etiology of menstrual cycle dysfunction as being a significant effect modifier in the association between ovarian morphology and androgen status. Indeed, the cause of cycle irregularity in our cohort was not uniform. Ovarian insufficiency, hypothyroidism and hyperprolactinemia were excluded. However, the cohort included women with and without clinical evidence of androgen excess. It is plausible that the variation in pathophysiology within this group, in addition to a small sample size, explains, in part, why we did not detect significance in the LASSO models. Our study supports the need for further research to fully clarify how relationships between ovarian morphology and total testosterone levels vary across the spectrum of reproductive dysfunction.

There were several strengths to this study. First, reliable and standardized methods for assaying total testosterone were used. Second, we used high resolution ultrasonographic technology and validated methods of follicle counting (27). Third, we recruited women from the general population and evaluated clinical markers of androgen action in a diverse cohort of women that spanned the androgenic spectrum. However, this study also had several limitations. First, the degree of hirsutism represented in our cohort was narrow as mentioned earlier. Because our study did not capture severe cases of androgen excess, the generalizability of our findings for women with higher hirsutism scores is limited. Second, despite our attempts to standardize all measurements, assessments of hirsutism are subjective and relied in part, on participant disclosure of current and previous cosmetic

practices. Given that women were not asked to refrain from cosmetic practices prior to attending study visits, there was risk of underestimating actual hair growth scores. Third, 12 participants had used hormonal contraception within one year of study participation – with half terminating use in the six months prior to enrollment. We acknowledge that recent contraceptive use had the potential to influence hair growth, ovarian morphology, and total testosterone levels (9, 53, 54). However, since there are limited data on the time course of clinical and biochemical manifestations following cessation of treatment, we felt it reasonable to exclude only those more recent users of hormonal contraception (two months or less). Fourth, we recognize that there is currently no statistical procedure for estimating power of the LASSO calculations. As such, we can only modestly estimate that our study had at least 75% power to detect a significant correlations (rho= 0.300) among clinical and sonographic markers with total testosterone levels (G*Power, Version 3.1.9.2, Universität Kiel, Germany). Last, our study did not include sufficient clinical evaluations to confirm the nature of the cycle irregularity in the cohorts studied. It is likely that the women studied include those with defined anovulatory disorders such as hypothalamic amenorrhea and polycystic ovary syndrome. As such, future sufficiently powered studies will aim to more comprehensively characterize how folliculogenesis is differentially impacted in these conditions and how unique disturbances in folliculogenesis might be reflected in cross-sectional sonographic evaluations of ovarian morphology.

In summary, sonographic markers of ovarian morphology, but neither regional nor total hirsutism scores, predicted testosterone levels. As such, sonographic markers may serve as a clinical biomarker for androgen status in instances where access to high-performance assays is more limited. The ability of ovarian morphology to predict total testosterone levels was modified by menstrual cycle status. Our findings support the use of ultrasonography as a potentially informative tool in the detection of hyperandrogenemia in women with regular menstrual cycles, where other clinical indicators of androgen excess may not necessarily be present. Alternate screening mechanisms could help to prevent or minimize the cardiovascular and metabolic sequelae associated with androgen excess in women. Future research is needed to fully elaborate how aspects of ovarian morphology reflect androgen levels in the context of variable etiologies for ovulatory dysfunction.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We are grateful to our participants, without whom this study would not be possible. We are also thankful to Lynn Johnson for her additional technical assistance with the statistical analysis.

**Funding:** This research was supported by Cornell University's Division of Nutritional Sciences, the President's Council of Cornell Women, the United States Department of Agriculture (#34324-20769), and the National Institutes of Health (T32-DK007158). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) or the National Institutes of Health.

## References


1. Rosner W, Vesper H, Endocrine Society, American Association for Clinical Chemisty, American Association of Clinical Endocrinologists, Androgen Excess. et al. Toward excellence in testosterone testing: a consensus statement. J Clin Endocrinol Metab. 2010; 95:4542–8. [PubMed: 20926540]
2. Moal V, Mathieu E, Reynier P, Malthiery Y, Gallois Y. Low serum testosterone assayed by liquid chromatography-tandem mass spectrometry. Comparison with five immunoassay techniques. Clin Chim Acta. 2007; 386:12–9. [PubMed: 17706625]
3. 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; 92:405–13. [PubMed: 17090633]
4. Miller KK, Rosner W, Lee H, Hier J, Sesmilo G, Schoenfeld D, et al. Measurement of free testosterone in normal women and women with androgen deficiency: comparison of methods. J Clin Endocrinol Metab. 2004; 89:525–33. [PubMed: 14764757]
5. Pugeat M, Crave JC, Elmidani M, Nicolas MH, Garoscio-Cholet M, Lejeune H, et al. Pathophysiology of sex hormone binding globulin (SHBG): relation to insulin. J Steroid Biochem Mol Bio. 1991; 40:841–9. [PubMed: 1958579]
6. Cupisti S, Dittrich R, Binder H, Kajaia N, Hoffmann I, Maltaris T, et al. Influence of body mass index on measured and calculated androgen parameters in adult women with hirsutism and PCOS. Exp Clin Endocrinol Diabetes. 2007; 115:380–6. [PubMed: 17701884]
7. Vesper HW, Bhasin S, Wang C, Tai SS, Dodge LA, Singh RJ, et al. Interlaboratory comparison study of serum total testosterone [corrected] measurements performed by mass spectrometry methods. Steroids. 2009; 74:498–503. [PubMed: 19428438]
8. Legro RS, Schlaff WD, Diamond MP, Coutifaris C, Casson PR, Brzyski RG, et al. Total testosterone assays in women with polycystic ovary syndrome: precision and correlation with hirsutism. J Clin Endocrinol Metab. 2010; 95:5305–13. [PubMed: 20826578]
9. Azziz R, Sanchez LA, Knochenhauer ES, Moran C, Lazenby J, Stephens KC, et al. Androgen excess in women: experience with over 1000 consecutive patients. J Clin Endocrinol Metab. 2004; 89:453–62. [PubMed: 14764747]
10. Ferriman D, Gallwey JD. Clinical assessment of body hair growth in women. J Clin Endocrinol Metab. 1961; 21:1440–7. [PubMed: 13892577]
11. Api M, Badoglu B, Akca A, Api O, Gorgen H, Cetin A. Interobserver variability of modified Ferriman-Gallwey hirsutism score in a Turkish population. Arch Gynecol Obstet. 2009; 279:473–9. [PubMed: 18677501]
12. Cook H, Brennan K, Azziz R. Reanalyzing the modified Ferriman-Gallwey score: is there a simpler method for assessing the extent of hirsutism? Fertil Steril. 2011; 96:1266–70. [PubMed: 21924716]
13. Derksen J, Moolenaar AJ, Van Seters AP, Kock DF. Semiquantitative assessment of hirsutism in Dutch women. Br J Dermatol. 1993; 128:259–63. [PubMed: 8471509]
14. Ramezani Tehrani F, Minooee S, Azizi F. Validation of a simplified method to assess hirsutism in the Iranian population. Eur J Obstet Gynecol Reprod Biol. 2014; 174:91–5. [PubMed: 24393448]
15. Rashidi H, Parizi ZT, Mohammadi M. Evaluation of only the chin or lower abdomen for predicting hirsutism. Indian J Endocrinol Metab. 2013; 17:896–8. [PubMed: 24083173]
16. Hines G, Moran C, Huerta R, Folgman K, Azziz R. Facial and abdominal hair growth in hirsutism: a computerized evaluation. J Am Acad Dermatol. 2001; 45:846–50. [PubMed: 11712028]
17. Clark NM, Podolski AJ, Brooks ED, Chizen DR, Pierson RA, Lehotay DC, et al. Prevalence of Polycystic Ovary Syndrome Phenotypes Using Updated Criteria for Polycystic Ovarian Morphology: An Assessment of Over 100 Consecutive Women Self-reporting Features of Polycystic Ovary Syndrome. Reprod Sci. 2014; 21:1034–43. [PubMed: 24520081]
18. Kiddy DS, Sharp PS, White DM, Scanlon MF, Mason HD, Bray CS, et al. Differences in clinical and endocrine features between obese and non-obese subjects with polycystic ovary syndrome: an analysis of 263 consecutive cases. Clin Endocrinol. 1990; 32:213–20.
19. Escobar-Morreale HF, Carmina E, Dewailly D, Gambineri A, Kelestimur F, Moghetti P, et al. Epidemiology, diagnosis and management of hirsutism: a consensus statement by the Androgen

Excess and Polycystic Ovary Syndrome Society. Hum Reprod Update. 2012; 18:146–70. [PubMed: 22064667]

20. Javorsky E, Perkins AC, Hillebrand G, Miyamoto K, Boer Kimball A. Race, rather than skin pigmentation, predicts facial hair growth in women. J Clin Aesthet Dermatol. 2014; 7:24–6. [PubMed: 24847405]
21. Legro RS, Myers ER, Barnhart HX, Carson SA, Diamond MP, Carr BR, et al. The pregnancy in polycystic ovary syndrome study: baseline characteristics of the randomized cohort including racial effects. Fertil Steril. 2006; 86:914–33. [PubMed: 16963034]
22. Yildiz BO, Bolour S, Woods K, Moore A, Azziz R. Visually scoring hirsutism. Hum Reprod Update. 2010; 16:51–64. [PubMed: 19567450]
23. Azziz R. The evaluation and management of hirsutism. Obstet Gynecol. 2003; 101:995–1007. [PubMed: 12738163]
24. Karrer-Voegeli S, Rey F, Reymond MJ, Meuwly JY, Gaillard RC, Gomez F. Androgen dependence of hirsutism, acne, and alopecia in women: retrospective analysis of 228 patients investigated for hyperandrogenism. Medicine. 2009; 88:32–45. [PubMed: 19352298]
25. Reingold SB, Rosenfield RL. The relationship of mild hirsutism or acne in women to androgens. Arch Dermatol. 1987; 123:209–12. [PubMed: 2949707]
26. Jayaprakasan K, Campbell BK, Clewes JS, Johnson IR, Raine-Fenning NJ. Three-dimensional ultrasound improves the interobserver reliability of antral follicle counts and facilitates increased clinical work flow. Ultrasound Obstet Gynecol. 2008; 31:439–44. [PubMed: 18330873]
27. Lujan ME, Brooks ED, Kepley AL, Chizen DR, Pierson RA, Peppin AK. Grid analysis improves reliability in follicle counts made by ultrasonography in women with polycystic ovary syndrome. Ultrasound Med Biol. 2010; 36:712–8. [PubMed: 20381953]
28. Lujan ME, Chizen DR, Peppin AK, Kriegler S, Leswick DA, Bloski TG, et al. Improving inter-observer variability in the evaluation of ultrasonographic features of polycystic ovaries. Reprod Biol Endocrinol. 2008; 6:30. [PubMed: 18638401]
29. Christ JP, Willis AD, Brooks ED, Vanden Brink H, Jarrett BY, Pierson RA, et al. Follicle number, not assessments of the ovarian stroma, represents the best ultrasonographic marker of polycystic ovary syndrome. Fertil Steril. 2014; 101:280–7. [PubMed: 24188871]
30. Fulghesu AM, Angioni S, Frau E, Belosi C, Apa R, Mioni R, et al. Ultrasound in polycystic ovary syndrome–the measuring of ovarian stroma and relationship with circulating androgens: results of a multicentric study. Hum Reprod. 2007; 22:2501–8. [PubMed: 17635847]
31. 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; 18:598–603. [PubMed: 12615832]
32. 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; 91:4842–8. [PubMed: 17003085]
33. Gougeon, A. Dynamics of human follicular growth: morphologic, dynamic, and functional aspects. In: Leung, PCK.; Adashi, EY., editors. The Ovary. San Diego, CA: Elsevier; 2004. p. 34-5.
34. Ben-Haroush A, Farhi J, Zahalka Y, Sapir O, Meizner I, Fisch B. Correlations between antral follicle count and ultrasonographic ovarian parameters and clinical variables and outcomes in IVF cycles. Gynecol Endocrinol. 2012; 28:432–5. [PubMed: 22122694]
35. Nardo LG, Christodoulou D, Gould D, Roberts SA, Fitzgerald CT, Laing I. Anti-Mullerian hormone levels and antral follicle count in women enrolled in in vitro fertilization cycles: relationship to lifestyle factors, chronological age and reproductive history. Gynecol Endocrinol. 2007; 23:486–93. [PubMed: 17852428]
36. Adams JM, Taylor AE, Crowley WF, Hall JE. Polycystic ovarian morphology with regular ovulatory cycles: insights into the pathophysiology of polycystic ovarian syndrome. J Clin Endocrinol Metab. 2004; 89:4343–50. [PubMed: 15356031]
37. Carmina E, Chu MC, Longo RA, Rini GB, Lobo RA. Phenotypic variation in hyperandrogenic women influences the findings of abnormal metabolic and cardiovascular risk parameters. J Clin Endocrinol Metab. 2005; 90:2545–9. [PubMed: 15728203]

38. Sung YA, Oh JY, Chung H, Lee H. Hyperandrogenemia is implicated in both the metabolic and reproductive morbidities of polycystic ovary syndrome. Fertil Steril. 2014; 101:840–5. [PubMed: 24424368]
39. Harlow SD, Gass M, Hall JE, Lobo R, Maki P, Rebar RW, et al. Executive summary of the Stages of Reproductive Aging Workshop + 10: addressing the unfinished agenda of staging reproductive aging. Fertil Steril. 2012; 97:843–51. [PubMed: 22341880]
40. Tibshirani R. Regression shrinkage and selection via the Lasso. J Roy Stat Soc B Met. 1996; 58:267–88.
41. Hastie T. lars: Least Angle Regression, Lasso and Forward Stagewise. R package version 12 [Internet]. 2013
42. Lockhart R, Taylor J, Tibshirani R, Tibshirani R. covTest: Computes covariance test for adaptive linear modelling. R package version 102 [Internet]. 2013
43. Lockhart R, Taylor J, Tibshirani RJ, Tibshirani R. A Significance Test for the Lasso. Ann Stat. 2014 Apr; 42(2):413–68. [PubMed: 25574062]
44. Puri N. A study on the clinical and hormonal profile of patients with hirsutism. Our Dermatol Online. 2012; 3:88–91.
45. Knochenhauer ES, Hines G, Conway-Myers BA, Azziz R. Examination of the chin or lower abdomen only for the prediction of hirsutism. Fertil Steril. 2000; 74:980–3. [PubMed: 11056244]
46. Lobo RA, Goebelsmann U, Horton R. Evidence for the importance of peripheral tissue events in the development of hirsutism in polycystic ovary syndrome. J Clin Endocrinol Metab. 1983; 57:393–7. [PubMed: 6223045]
47. Arduc A, Saricam O, Dogan BA, Tuna MM, Tutuncu YA, Isik S, et al. Should insulin resistance be screened in lean hirsute women? Gynecol Endocrinol. 2015; 31:291–5. [PubMed: 25561024]
48. Rosenfield RL. Clinical practice. Hirsutism. N Engl J Med. 2005; 353:2578–88. [PubMed: 16354894]
49. Jeppesen JV, Anderson RA, Kelsey TW, Christiansen SL, Kristensen SG, Jayaprakasan K, et al. Which follicles make the most anti-Mullerian hormone in humans? Evidence for an abrupt decline in AMH production at the time of follicle selection. Mol Hum Reprod. 2013; 19:519–27. [PubMed: 23562944]
50. Franks S, Mason H, White D, Willis D. Etiology of anovulation in polycystic ovary syndrome. Steroids. 1998; 63:306–7. [PubMed: 9618791]
51. Gilling-Smith C, Willis DS, Beard RW, Franks S. Hypersecretion of androstenedione by isolated thecal cells from polycystic ovaries. J Clin Endocrinol Metab. 1994; 79:1158–65. [PubMed: 7962289]
52. Dewailly D, Lujan ME, Carmina E, Cedars MI, Laven J, Norman RJ, et al. Definition and significance of polycystic ovarian morphology: a task force report from the Androgen Excess and Polycystic Ovary Syndrome Society. Hum Reprod Update. 2014; 20:334–52. [PubMed: 24345633]
53. Deb S, Campbell BK, Pincott-Allen C, Clewes JS, Cumberpatch G, Raine-Fenning NJ. Quantifying effect of combined oral contraceptive pill on functional ovarian reserve as measured by serum anti-Mullerian hormone and small antral follicle count using three-dimensional ultrasound. Ultrasound Obstet Gynecol. 2012; 39:574–80. [PubMed: 21997961]
54. Mes-Krowinkel MG, Louwers YV, Mulders AG, de Jong FH, Fauser BC, Laven JS. Influence of oral contraceptives on anthropomorphometric, endocrine, and metabolic profiles of anovulatory polycystic ovary syndrome patients. Fertil Steril. 2014; 101:1757–65. [PubMed: 24680368]

**Table 1**

Clinical and sonographic characteristics of study participants with regular and irregular menstrual cycles.

| | REGULAR CYCLES (N=39) | | | IRREGULAR CYCLES (N=37) | | |
|---|---|---|---|---|---|---|
| | **$P^{05}$** | **$P^{50}$** | **$P^{95}$** | **$P^{05}$** | **$P^{50}$** | **$P^{95}$** |
| **Demographics** | | | | | | |
| Age (y) | 19 | 28 | 38 | 19 | 24 | 35.0 |
| BMI (kg/m$^2$) | 19.5 | 23.9 | 42.1 | 19.1 | 26.8 | 52.8 |
| Menarche (y) | 10 | 12 | 14.5 | 10 | 13 | 15.5 |
| Total Testosterone (ng/dL) | 15.0 | 33.4 | 71.1 | 17.3 | 35.6 | 98.9 |
| Mean Menstrual Cycle Length (d) | 26.5 | **30.0** ** | 34.0 | 37.0 | **61.0** ** | 365.0 |
| **Clinical Markers** | | | | | | |
| Total mFG Score | 0.0 | 4.0 | 13.0 | 1.0 | 6.0 | 15.0 |
| Upper Lip Score | 0.0 | **0.0** * | 2.0 | 0.0 | **1.0** * | 3.0 |
| Chin Score | 0.0 | **0.0** * | 3.0 | 0.0 | **1.0** * | 3.0 |
| Upper Back Score | 0.0 | 0.0 | 2.0 | 0.0 | 0.0 | 1.0 |
| Lower Back Score | 0.0 | 0.0 | 2.5 | 0.0 | 0.0 | 2.0 |
| Upper Arm Score | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 1.0 |
| Thigh Score | 0.0 | 0.0 | 3.0 | 0.0 | 1.0 | 2.0 |
| Chest Score | 0.0 | 0.0 | 2.0 | 0.0 | 1.0 | 3.0 |
| Upper Abdomen | 0.0 | 0.0 | 2.0 | 0.0 | 0.0 | 3.0 |
| Lower Abdomen | 0.0 | 1.0 | 3 | 0.0 | 1.0 | 3.0 |
| **Sonographic Markers** | | | | | | |
| 2–5mm FNPO | 5.5 | **14.5** ** | 40 | 7.5 | **23** ** | 74.0 |
| 6–9mm FNPO | 0.5 | 3.0 | 12.5 | 0.0 | 3.0 | 14 |
| 2–9mm FNPO | 7.5 | **19** ** | 41.5 | 9.5 | **27** ** | 77 |
| S/A Ratio | 0.25 | 0.44 | 0.63 | 0.29 | 0.40 | 0.57 |
| Stromal Index | 1.21 | 1.33 | 1.61 | 1.19 | 1.36 | 1.55 |
| Ovarian Area (cm$^2$) | 2.74 | **4.67** ** | 7.05 | 3.78 | **5.66** ** | 9.32 |

| | REGULAR CYCLES (N=39) | | | IRREGULAR CYCLES (N=37) | | |
|---|---|---|---|---|---|---|
| | **$P^{05}$** | **$P^{50}$** | **$P^{95}$** | **$P^{05}$** | **$P^{50}$** | **$P^{95}$** |
| Ovarian Volume (mL) | 3.86 | **7.62** ** | 13.75 | 5.36 | **9.55** ** | 19.64 |

Abbreviations include: BMI, Body Mass Index; mFG Score, modified Ferriman-Gallwey Score; FNPO, Follicle Number Per Ovary; S/A Ratio, Stromal-to-total Area Ratio. Significant differences between groups denoted by

*P<0.05;

**P<0.01.

## Table 2

Order of entry of clinical covariates (regional and total hirsutism scores) into models predicting total testosterone.

| | REGULAR CYCLES | | IRREGULAR CYCLES | |
|---|---|---|---|---|
| **Rank Order** | **Clinical Marker** | **P-value** | **Clinical Marker** | **P-value** |
| 1 | Lower Back | 0.944 | Lower Abdomen | 0.415 |
| 2 | Chin | 0.904 | Upper Abdomen | 0.956 |
| 3 | Upper Back | 0.910 | Chin | 0.957 |
| 4 | Lower Abdomen | 0.916 | Lower Back | 0.943 |
| 5 | Upper Abdomen | 0.792 | Thigh | 0.972 |
| 6 | Chest | 0.998 | Upper Lip | 0.736 |
| 7 | Thigh | 0.961 | Upper Back | 0.640 |
| 8 | Upper Arm | 0.953 | Upper Arm | 0.865 |
| 9 | Total mFG Score | 0.953 | Lower Back * | N/A |
| 10 | Upper Lip | 0.943 | Chest | 0.910 |

* Indicates covariate exited model to improve model fit. Abbreviations include: mFG Score, modified Ferriman-Gallwey Score.

## Table 3

Order of entry of sonographic covariates into models predicting total testosterone.

| | REGULAR CYCLES | | IRREGULAR CYCLES | |
|---|---|---|---|---|
| **Rank Order** | **Sonographic Marker** | **P** | **Sonographic Marker** | **P** |
| 1 | 6–9mm FNPO | 0.001 | OV | 0.623 |
| 2 | S/A Ratio | 0.699 | 6–9mm FNPO | 0.652 |
| 3 | SI | 0.396 | S/A Ratio | 0.129 |
| 4 | OV | 0.561 | 2–5mm FNPO | 0.999 |
| 5 | 2–5mm FNPO | 0.965 | SI | 0.685 |

Abbreviations include: FNPO, Follicle Number Per Ovary; S/A Ratio, Stromal-to-total Area Ratio; OV, Ovarian Volume; SI, Stromal Index.