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

Heidi Vanden Brink, M.Sc.,[a] Amy D. Willis, M.Sc.,[b] Brittany Y. Jarrett, B.S.,[a] Annie W. Lin, M.Sc.,[a] Steven Soler, B.S.,[a] Siân Best, B.A.,[a] Erica L. Bender, M.S.N.,[a] Andrew K. Peppin, M.D.,[a] Kathleen M. Hoeger, M.D.,[c] and Marla E. Lujan, M.Sc., Ph.D.[a]

[a] Human Metabolic Research Unit, Division of Nutritional Sciences, Cornell University, Ithaca; [b] Department of Statistical Sciences, Cornell University, Ithaca; and [c] Division of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynecology, University of Rochester Medical Center, Rochester, New York

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 a modified Ferriman-Gallwey scoring system), ovarian morphology (by transvaginal ultrasonography), and total serum testosterone (T) (by liquid chromatography tandem mass spectrometry).
Intervention(s): Not applicable.
Main Outcome Measure(s): Regional and total modified Ferriman-Gallwey scores, number of follicles per follicle size category, follicle number per ovary, ovarian volume, ovarian area, stromal to ovarian area ratio, stromal echogenicity index, total testosterone (total T), and menstrual cycle length.
Result(s): Neither regional nor total modified Ferriman-Gallwey scores correlated with total T concentrations in women with regular or irregular menstrual cycles, as judged by the Least Absolute Shrinkage and Selection Operator technique. By contrast, a sonographic marker (follicle number per ovary 6–9 mm) significantly predicted total T concentrations in women with regular menstrual cycles but not in women with irregular menstrual cycles.
Conclusion(s): Sonographic markers of ovarian morphology, but not hirsutism scores, predicted total T levels. However, the predictive value of ovarian morphology for total T 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 nonandrogenic types of cycle irregularity. Overall, our findings support that an assessment of ovarian morphology may be helpful in reflecting total T levels. (Fertil Steril® 2016;105: 1322–9. ©2016 by American Society for Reproductive Medicine.)
Key Words: Hirsutism, oligoamenorrhea, ovaries, testosterone, ultrasonography

Discuss: You can discuss this article with its authors and with other ASRM members at http://fertstertforum.com/brinkh-ovarian-morphology-androgen-levels/

Use your smartphone to scan this QR code and connect to the discussion forum for this article now.*

* Download a free QR code scanner by searching for "QR scanner" in your smartphone's app store or app marketplace.

Received June 5, 2015; revised December 24, 2015; accepted December 28, 2015; published online January 27, 2016.
H.V.B. has nothing to disclose. A.D.W. has nothing to disclose. B.Y.J. has nothing to disclose. A.W.L. has nothing to disclose. S.S. has nothing to disclose. S.B. has nothing to disclose. E.L.B. has nothing to disclose. A.K.P. has nothing to disclose. K.M.H. has nothing to disclose. M.E.L. has nothing to disclose.
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 or the National Institutes of Health.
Portions of this work were presented at the annual meetings of the Canadian Fertility and Andrology Society (September 11–14, 2014, Quebec City, Quebec, Canada) and the Androgen Excess and PCOS Society (October 22–23, 2014, Lihue, Hawaii).
Reprint requests: Marla E. Lujan, M.Sc., Ph.D., Division of Nutritional Sciences, Cornell University, 216 Savage Hall, Ithaca, New York 14853 (E-mail: MEL245@cornell.edu).

Fertility and Sterility® Vol. 105, No. 5, May 2016 0015-0282/$36.00

http://dx.doi.org/10.1016/j.fertnstert.2015.12.136

Biochemical assessments of androgens in women are controversial (1). Commercial assays for serum testosterone (T) yield inconsistent results (2, 3), whereas direct measurements of free T are technically challenging (3, 4) and influenced by metabolic status (5, 6). Overall, mass spectrometry assay performance is improved compared with commercially available products—albeit modest interlaboratory 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 because 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 interrater 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. Although hirsutism has shown better sensitivity for biochemical hyperandrogenism compared with acne or alopecia (23, 24), its specificity is low because idiopathic hirsutism occurs in 5%–15% of the general population (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 (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 (total T) 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 T levels in women. To this end, we enrolled women with regular and irregular cycles, 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 T concentrations, whereas a marker reflecting a consequence of androgen action, such as a hirsutism index, would have limited ability to predict total T. 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, New York and surrounding area) were recruited to the study between 2009 and 2014. Participants were recruited using targeted advertisements 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, vital signs, 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 examination, 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 to 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 before 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 that 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 (>10 mm) 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. Digital images of each ovary were archived for offline 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) $\times$ (longitudinal diameter) and OV calculated based on the equation $\pi/6$ (transverse diameter) $\times$ (anteroposterior diameter) $\times$ (longitudinal diameter). The 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 (follicle number per ovary, FNPO). Physiologic cohorts of interest included [1] number of 2–5 mm follicles per ovary, [2] number of 6–9 mm follicles per ovary, and [3] number of 2–9 mm follicles per ovary. On the basis of an intraclass correlation coefficient analysis, the level of agreement among three 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 one ovary (n = 1), data for a single ovary was reported.

Testosterone Assay

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

Statistical Analysis

Variables that best predicted total T 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). The LASSO technique permits variable selection in the high-dimensional multiple linear regression context (i.e., a large number of predictors relative to the number of observations). The LASSO technique 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 on the basis of 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 2–5 mm FNPO, 6–9 mm FNPO, S/A, SI, and OV, for a total of five covariates in the model. Because of high collinearity, 2–9 mm FNPO and OA were excluded because they destabilized the model. Total T 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 test (categorical variables) (SPSS Statistics V23). Descriptive statistics (5th percentile, median, and 95th percentile) of clinical and sonographic endpoints are provided for each cohort and for each cohort according to the increasing levels (quartiles) of T.

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). 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 T levels compared with 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 with those with regular cycles. Women with irregular menstrual cycles had more antral follicles (2–5 mm and 2–9 mm), as well as larger ovaries compared with 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% vs. 74%, $P=.147$). Although none of the participants used hormonal agents in the 2 months preceding the study, 12 participants reported use of hormonal contraception in the year before 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%; $P=.655$). Similarly, ethnicity did not differ between groups (regular cycles: Hispanic 11% and non-Hispanic 70%; and irregular cycles: Hispanic 5% and non-Hispanic 81%; $P=.512$).

Clinical and Sonographic Predictors of Total T

Covariates most predictive of total T 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 T in women with regular or irregular cycles. By contrast, in a model including ovarian markers, the number of 6–9 mm follicles significantly predicted total T in the LASSO model in women with regular menstrual cycles (Table 3; $P=.001$). In women with irregular cycles, sonographic markers did not significantly predict total T (Table 3). This was despite larger ovarian size ($P=.038$) and numerically higher follicle counts (2–9 mm FNPO) in women with higher total T levels (Supplemental Table 1, available online).

DISCUSSION

This study was conducted to assess the predictive value of hirsutism scores and ovarian sonographic markers for total T

TABLE 1

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

	Regular cycles (n = 39)			Irregular cycles (n = 37)		
Characteristic	**P^{05}**	**P^{50}**	**P^{95}**	**P^{05}**	**P^{50}**	**P^{95}**
Demographics						
Age (y)	19	28	38	19	24	35.0
Body mass index (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 T (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 score	0.0	0.0	2.0	0.0	0.0	3.0
Lower abdomen score	0.0	1.0	3	0.0	1.0	3.0
Sonographic markers						
2–5 mm FNPO	5.5	14.5**	40	7.5	23**	74.0
6–9 mm FNPO	0.5	3.0	12.5	0.0	3.0	14
2–9 mm 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
SI	1.21	1.33	1.61	1.19	1.36	1.55
OA (cm^2)	2.74	4.67**	7.05	3.78	5.66**	9.32
OV (mL)	3.86	7.62**	13.75	5.36	9.55**	19.64

*$P<.05$.
**$P<.01$.

Vanden Brink. Ovarian morphology reflects androgen levels. Fertil Steril 2016.

TABLE 2

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

	Regular cycles		Irregular cycles	
Rank order	Clinical marker	*P* value	Clinical marker	*P* value
1	Lower back	.944	Lower abdomen	.415
2	Chin	.904	Upper abdomen	.956
3	Upper back	.910	Chin	.957
4	Lower abdomen	.916	Lower back	.943
5	Upper abdomen	.792	Thigh	.972
6	Chest	.998	Upper lip	.736
7	Thigh	.961	Upper back	.640
8	Upper arm	.953	Upper arm	.865
9	Total mFG	.953	Lower back[a]	N/A
10	Upper lip	.943	Chest	.910

[a] Covariate exited model to improve model fit.

Vanden Brink. Ovarian morphology reflects androgen levels. Fertil Steril 2016.

levels, using an accurate and reliable assay for androgen status. We noted that regional mFG scores were more predictive of total T compared with the total mFG score. However, neither regional nor total mFG scores predicted total T to any significant degree in either women with regular or irregular menstrual cycles. By contrast, we showed that a sonographic marker of ovarian morphology was a significant and strong predictor of total T. 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 T 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, serves 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 T in women with regular menstrual cycles and never entered the model predicting T 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 T on susceptible areas of the skin over time rather than some aspect of current androgen production.

TABLE 3

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

	Regular cycles		Irregular cycles	
Rank order	Sonographic marker	*P* value	Sonographic marker	*P* value
1	6–9-mm FNPO	.001	OV	.623
2	S/A Ratio	.699	6–9-mm FNPO	.652
3	SI	.396	S/A Ratio	.129
4	OV	.561	2–5-mm FNPO	.999
5	2–5-mm FNPO	.965	SI	.685

Vanden Brink. Ovarian morphology reflects androgen levels. Fertil Steril 2016.

Our finding that total mFG scores were neither predictive of, or related to, total T 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 95th percentiles for total mFG scores were 13 and 15 for women with regular and irregular menstrual cycles, respectively, 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 T. Additionally, the present study was composed of mainly non-Hispanic Caucasian women, which is consistent with our local demographic (Tompkins County, New York). Differences in hair growth scores across diverse races and ethnicities are generally accepted (22). However, differences in hair growth within Caucasian populations have been noted, with individuals from Northern Europe having lower hirsutism scores compared with 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 T 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 with other techniques (2, 3). However, it is worth noting that there are several studies using commercial T 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 T levels. In women with regular menstrual cycles, the number of larger follicles (6–9 mm)—which physiologically corresponds to follicles recruited to a wave-like cohort—emerged as a significant predictor of total T levels. These findings complement recent reports by Jeppesen et al. (49), who showed that 5–8 mm follicles represent a physiologically informative follicular pool compared with smaller follicle populations by contributing the majority of circulating antimüllerian hormone levels. Healthy 6–9 mm follicles would be expected to be a significant source of T 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 T and aspects of ovarian morphology in women with irregular cycles. That said, our findings contrast with previous work, which demonstrated that smaller follicles measuring 2–5 mm were positively correlated with total T, whereas follicles of 6–9 mm were negatively correlated with total T, in an unadjusted analysis (31). Our data differ, in that assessments by Dewailly et al. 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 T in women with regular menstrual cycles but not in those with irregular cycles. When we explored bivariate associations among follicle populations and T levels, we confirmed that total T correlated with the number of 6–9-mm ($\rho = 0.452$; $P = .004$) but not with the 2–5 mm follicles ($\rho = -0.025$, $P = .881$) in women with regular menstrual cycles (correlation analyses not shown). Likewise, we noted that OV ($\rho = 0.422$, $P = .01$) but neither the number of 2–5-mm ($\rho = 0.274$, $P = .100$) nor 6–9 mm ($\rho = 0.187$, $P = .267$) follicles, correlated with total T in women with irregular menstrual cycles. When data for women with irregular menstrual cycles were stratified by increasing T 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 T levels vary across the spectrum of reproductive dysfunction.

There were several strengths to this study. First, reliable and standardized methods for assaying total T 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 before attending study visits, there was risk of underestimating actual hair growth scores. Third, 12 participants had used hormonal contraception within 1 year of study participation—with half terminating use in the 6 months before enrollment. We acknowledge that recent contraceptive use had the potential to influence hair growth, ovarian morphology, and total T levels (9, 53, 54). However, because there are limited data on the time course of clinical and biochemical manifestations after cessation of treatment, we thought it reasonable to exclude only those more recent users of hormonal contraception (2 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 T 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 T 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 T 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.

Acknowledgments: The authors thank the participants, without whom this study would not be possible; and Lynn Johnson for her additional technical assistance with the statistical analysis.

REFERENCES

1. Rosner W, Vesper H, et al, Endocrine Society, American Association for Clinical Chemistry, American Association of Clinical Endocrinologists, Androgen Excess/PCOS Society. Toward excellence in testosterone testing: a consensus statement. J Clin Endocrinol Metab 2010;95:4542–8.
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.
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.

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.
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.
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.
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.
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.
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.
10. Ferriman D, Gallwey JD. Clinical assessment of body hair growth in women. J Clin Endocrinol Metab 1961;21:1440–7.
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.
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.
13. Derksen J, Moolenaar AJ, Van Seters AP, Kock DF. Semiquantitative assessment of hirsutism in Dutch women. Br J Dermatol 1993;128:259–63.
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.
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.
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.
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.
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.
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.
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.
22. Yildiz BO, Bolour S, Woods K, Moore A, Azziz R. Visually scoring hirsutism. Hum Reprod Update 2010;16:51–64.
23. Azziz R. The evaluation and management of hirsutism. Obstet Gynecol 2003;101:995–1007.
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.
25. Reingold SB, Rosenfield RL. The relationship of mild hirsutism or acne in women to androgens. Arch Dermatol 1987;123:209–12.
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.
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.
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.
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.
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.
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.
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.
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; 2004:34–5.
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.
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.
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.
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.
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.
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.
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 2013.
42. Lockhart R, Taylor J, Tibshirani R, Tibshirani R. covTest: Computes covariance test for adaptive linear modelling. R package version 102 2013.
43. Lockhart R, Taylor J, Tibshirani RJ, Tibshirani R. A significance test for the Lasso. Ann Stat 2014;42:413–68.
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.
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.
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.
48. Rosenfield RL. Clinical practice. Hirsutism. N Engl J Med 2005;353:2578–88.
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.
50. Franks S, Mason H, White D, Willis D. Etiology of anovulation in polycystic ovary syndrome. Steroids 1998;63:306–7.
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.
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.
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.
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.

SUPPLEMENTAL TABLE 1

Clinical and sonographic features of women with regular and irregular menstrual cycles stratified by increasing total T (TT) levels.

	Irregular cycles												Regular cycles											
	TT quartile 1 (n = 10)			TT quartile 2 (n = 10)			TT quartile 3 (n = 10)			TT quartile 4 (n = 9)			TT quartile 1 (n = 10)			TT quartile 2 (n = 10)			TT quartile 3 (n = 10)			TT quartile 4 (n = 9)		
Feature	**P^{05}**	**P^{50}**	**P^{95}**	**P^{05}**	**P^{50}**	**P^{95}**	**P^{05}**	**P^{50}**	**P^{95}**	**P^{05}**	**P^{50}**	**P^{95}**	**P^{05}**	**P^{50}**	**P^{95}**	**P^{05}**	**P^{50}**	**P^{95}**	**P^{05}**	**P^{50}**	**P^{95}**	**P^{05}**	**P^{50}**	**P^{95}**
TT (ng/dL)[a,b]	15.8	22.2	25.3	26.7	32.8	35.6	36.7	39.4	47.5	52.5	57.6	100.8	8.6	20.3	22.9	23	28.7	33.4	34.5	36.6	38.4	38.6	43.8	82.4
Total mFG score	0.0	8.0	12.0	1.0	3.5	10.0	2.0	6.0	12.0	1.0	7.0	15.0	0.0	4.5	11.0	0.0	5.5	13.0	0.0	4.0	7.5	0.0	4.0	13.0
2–5 mm FNPO	4.5	16.0	77.5	7.5	23.5	49.5	8.0	23.0	46.5	20.5	24.5	74.0	4.0	16.3	40	6.0	18.0	27.0	5.5	12.5	28.5	8.5	21.0	43.0
6–9 mm FNPO[b]	0.5	3.5	7.0	0.5	2.0	12.0	0.0	3.5	12.5	0.0	3.0	20.0	0.0	1.8	6.5	1.0	3.0	9.0	1.5	3.5	7.5	0.5	5.0	18.0
2–9 mm FNPO	9.5	19.0	82.5	9.5	24.8	60.5	10.0	29.5	54.0	22.0	37.0	77.0	7.0	17.8	41.5	7.5	20.5	33.0	8.5	17.3	30.0	13.5	22.5	61.0
S/A Ratio	0.32	0.43	0.67	0.27	0.41	0.57	0.29	0.36	0.54	0.29	0.42	0.46	0.29	0.39	0.6	0.25	0.46	0.63	0.34	0.41	0.51	0.23	0.47	0.66
SI	1.19	1.33	1.64	1.23	1.38	1.52	1.25	1.31	1.55	1.17	1.36	1.52	1.23	1.37	1.52	1.21	1.34	1.52	1.24	1.33	1.61	1.20	1.31	1.66
OA (cm^2)	3.2	5.2	8.9	3.9	5.7	7.1	3.8	4.7	11.0	4.2	5.8	9.3	3.0	4.7	6.1	3.2	4.6	6.7	3.0	4.7	5.8	2.4	4.5	7.8
OV (mL)[a]	4.3	7.43	14.6	5.42	7.8	15.6	6.9	9.3	27.7	6.6	12.3	19.6	3.9	7.8	10.7	3.2	7.6	10.9	4.6	6.9	9.8	4.3	8.3	14.7

[a] Significant differences across quartiles for the irregular cycles group ($P < .05$).
[b] Significant differences across quartiles for the regular cycles group ($P < .05$).

Vanden Brink. Ovarian morphology reflects androgen levels. Fertil Steril 2016.