A simplified questionnaire for self-assessment of hirsutism in population-based studies

Ligia Gabrielli[1,2] and Estela M L Aquino[1]

[1]Institute of Collective Health, Federal University of Bahia, Rua Basílio da Gama, s/n, Campus Universitário do Canela, 40110-040 Salvador, Bahia, Brazil and [2]Bahia State Center for Diabetes and Endocrinology, Avenida ACM, s/n, Iguatemi, 40275-350 Salvador, Bahia, Brazil

Correspondence
should be addressed
to L Gabrielli
Email
ligiagabrielli@uol.com.br

Abstract

Objective: The measurement of excess body hair is not straightforward. As the modified Ferriman–Gallwey (mFG) score is unsuitable for self-assessment and requires specialist training, a short, self-administered questionnaire to identify hirsutism was constructed and validated for large-scale application, particularly targeting population-based studies.

Design: A validation study was conducted to assess a new hirsutism questionnaire.

Methods: A total of 90 women aged 35–72 years who were enrolled in the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil) were evaluated. A self-administered instrument containing four questions was designed to evaluate five body areas: upper lip, chin, chest, lower abdomen, and thighs with respect to the current distribution of body hair and that before 35 years of age. A score of 0–4 was attributed to each region based on drawings provided in the instrument. Test–retest reliability was evaluated by reformulating the initial questions. An independent medical examination was conducted to apply the gold standard, the mFG score.

Results: The area under the receiver operating characteristic curve was 0.93 (95% CI: 0.87–0.99). A cut-off score of 5 showed the best balance between sensitivity (85%) and specificity (90%), with 88.9% accuracy. Spearman's correlation between current and past body hair score was calculated at 0.82 ($P = 0.000$), and showed a test–retest reliability of 0.49, with a trend toward similar answers regarding changes in the quantity of body hair over time, irrespective of how the questions were asked ($P = 0.000$).

Conclusion: The accuracy and internal consistency of this self-administered questionnaire for the identification of hirsutism were good. Therefore, this questionnaire represents a useful tool for self-assessment of hirsutism in population-based studies.

Introduction

Hirsutism is a common clinical condition defined by excess terminal body hair in androgen-dependent areas of the female body (1, 2, 3). It may be present in polycystic ovary syndrome (PCOS), which is responsible for 60–85% of cases, in non-classical congenital adrenal hyperplasia (3–5%) and in idiopathic hirsutism (5–10%) (2, 3, 4, 5, 6). Its occurrence is linked to an exaggerated androgenic effect that occurs through various mechanisms such as an increase in the production of ovarian or adrenal androgens (1), greater enzymatic activity in the pilosebaceous units (4, 7), an increase in the proportion of free androgens (7),

and the use of exogenous hormones. Hirsutism is the most common manifestation of hyperandrogenism and the clinical criterion most commonly used to identify hyper-androgenic women (8).

The measurement of excess body hair has been the subject of major controversies. For decades, investigators have tried to establish a method of quantifying body hair that would result in a more precise diagnosis of hirsutism and reduce the inter-observer variations found with the subjective methods that have been proposed. Described by David Ferriman and John Gallwey in 1961, the scoring

system that bears their names was originally designed to evaluate the amount of body hair in eleven regions of the body (9). Following revisions introduced 20 years later, the modified Ferriman–Gallwey (mFG) score is currently the method most commonly used to identify hirsutism (10). It consists of observing the quantity and distribution of terminal hair in nine body areas including the upper lip, chin, chest, upper and lower abdomen, back, sacroiliac region, thighs, and arms. These areas are given a score ranging from 0 to 4 according to quantity and density, with higher scores being indicative of a greater amount of body hair (9). The sum of these values is classified according to varying cut-off points, the most commonly used being between 6 and 8 points (11). Owing to its ease of application and low cost, the mFG is widely accepted. Nevertheless, the fact that this method is difficult to self-administer and requires specialist training (12) renders its use in studies involving a great number of participants largely unfeasible.

Objective methods of quantifying body hair by counting or weighing have been investigated; however, they appear to be more useful in the therapeutic control of hirsutism and are not recommended for use in diagnosis or for baseline quantification (13). The identification of ethnic differences constitutes another challenge. A study conducted in 915 Chinese women failed to identify hirsutism in any of the participants when the criteria used in Western countries were applied (14). On the other hand, a population-based study carried out in Iran estimated a prevalence of hirsute women of 25.6% (15). With such disparate populations insofar as hirsutism is concerned, it would appear relevant to develop methods of quantification that would take these differences into account and to evaluate these methods locally (8).

Some androgenically sensitive areas are known to be more closely associated with hyperandrogenemia, hirsutism, and PCOS than others. The chin, submandibular region, upper lip, upper thighs, the midline of the abdomen, and the chest were identified as being the most sensitive areas (16). The increase in body hair on the upper lip, the chin, and the lower abdomen proved to be better predictors of hirsutism (16, 17). Unpublished data from a study conducted to determine the prevalence of PCOS (18) estimated that, of the nine areas investigated using the mFG score, only five were associated with a diagnosis of hirsutism and were predictive of the syndrome: the upper lip, chin, chest, lower abdomen, and thighs. Another study reported that, in body areas that are more sensitive to androgens, a strong correlation was found between the clinical evaluation and self-assessment of hirsutism (19).

There are no validated methods for quantifying body hair in menopausal women. Nevertheless, the identification of hirsutism in these women might be important, first because not all women with PCOS are diagnosed before menopause. Furthermore, after a woman's reproductive years, various situations related to excess body hair may develop including hirsutism induced by obesity, ovarian hyperthecosis, and androgen-producing tumors (18, 20). In addition, the clinical effects of hyperandrogenism alone or within the context of PCOS at this time of life may be related to atherogenesis, insulin resistance, breast cancer, or cardiovascular disease (21, 22, 23).

Hirsutism is an important element in the identification of PCOS, which, in turn, has been strongly associated with the occurrence of obesity, diabetes mellitus, and cardiovascular disease. The pandemic of this triad has increased the interest of the scientific community in the subject of PCOS; however, adequate conditions must be created in order to conduct population-based studies.

The objective of this study was to construct and validate a simplified questionnaire to identify hirsutism that could be used in research studies, could be applied to large sample populations, and could be administered by the various means available, including the internet.

Materials and methods

Study design and population

This study was conducted in 90 consecutive women voluntarily participating at the baseline time point of a longitudinal study on adult health (The Brazilian Longitudinal Study of Adult Health (ELSA-Brasil)) (24) at the Federal University of Bahia's Research Center in Salvador, Bahia, Brazil.

A short instrument containing four questions was developed with the aim of enabling it to be administered by the participant herself, and was named *Hirsuta* (HST). It was based on the mFG score and was designed to evaluate the five body areas considered most sensitive to the circulating androgens associated with hirsutism and with a diagnosis of PCOS: i) upper lip; ii) chin and neck; iii) breasts and chest; iv) lower abdomen; and v) thighs. Drawings were elaborated to represent the different scores, ranging from 0 to 4 for each of the aforementioned regions in accordance with the description of the score and the drawings provided previously in various papers published on this subject (9, 10, 25). The participants were asked to answer questions on the current status of their body hair

as it would be if they had not used any hair removal methods. Although written explanations were added to each one of the drawings in case the participants required further help, it was believed that the drawings alone would provide sufficient information to allow them to answer the questions.

Two moments in time were taken into consideration when formulating the questions: the first referring to the current quantity and distribution of body hair and the second to body hair status before the participant reached 35 years of age. Information on hirsutism in the past was collected because it was assumed that modifications might occur in the quantitative pattern of body hair over time. As the participants in the ELSA-Brasil study were 35–74 years old at baseline, with a median of 51 years, it was decided to compare and analyze responses referring to two moments in the women's lives, based on the hypothesis that there could be a dependency relationship between these two variables. The lower age limit of 35 years was chosen because this was also one of the inclusion criteria for the ELSA-Brasil study, thus representing a moment in the past that was common to all the participants. This limit was also considered adequate because at this age the full effect of androgens on the hair follicles has been already present (9). A third question was included with the intention of reformulating the previous ones, thus enabling test–retest reliability to be evaluated in the same instrument, avoiding the need to reapply the questionnaire. This approach not only avoided changes in answers due to a putative time interval between two applications of the same questionnaire, but also minimized time and resources. The final question referred to the use of permanent hair removal methods that could account for non-hormonal changes in body hair patterns over time.

Preliminary tests and a pilot study were conducted with participants of the ELSA-Brasil study under similar circumstances to those planned for the validation study.

The participants completed the questionnaire privately and were then examined clinically and independently by one single investigator with a vast experience in applying the mFG score, which was the gold standard for validating the answers to this short-form questionnaire.

Statistical analysis

Internal consistency reliability was analyzed using the test–retest method, with different formulations of the same question being presented within the instrument. Differences between current and past body hair scores that were obtained following objective questions, and their responses based on drawings, were compared with the more subjective third question, which asked the participant to compare their current and past body hair statuses. For this final question, the following possible answers were provided: 'much less hair', 'less hair', 'the same amount of hair', 'more hair', or 'much more hair'; these alternative answers being scored from 1 to 5. Spearman's rank correlation test and Cuzick's test for trend across ordered groups were applied to these values. To evaluate the reliability of the answers on body hair status in the past, Spearman's correlation was also used to compare the scores obtained for the current and past quantifications and Student's t-test for paired samples was used to compare means.

Linear regression estimated the proportion contributed by each area of the body to the variability in the overall score, when applying both the mFG and the HST. This was achieved by calculating R^2, i.e., the square of the correlation between the score found in the test and that found with the gold standard. For the contribution of each area in the final diagnosis of hirsutism, R^2 was calculated within a logistic regression model.

For the validation study, sensitivity, specificity, and the percentage of correctly classified tests (accuracy) were calculated using the various cut-off points suggested by the receiver operating characteristic (ROC) curve. In addition, the area under the ROC curve was calculated, showing the overall performance of the test. To define the cut-off value to be selected for the new instrument, limits were estimated with a view to maximizing sensitivity and specificity. These analyses were performed using the cut-off points most commonly used in the mFG score (11).

Bland–Altman plots were constructed to evaluate agreement between the short-form instrument proposed in this study and the gold standard, and between the HST scores for current and past body hair statuses. This method allows a graphic analysis to be performed, which evaluates differences and means in the measurements, establishing 95% limits of agreement (26, 27).

The Stata software program, version 10.0 (StataCorp., College Station, TX, USA), was used throughout the statistical analysis.

Ethical aspects

This study was evaluated and approved by the internal review board of the Collective Health Institute, Federal University of Bahia. All subjects of this investigation signed an informed consent.

Results

The questionnaire was answered by 90 women aged 35–72 years (median: 50.5 years) and the participation refusal rate was 5.3%. Overall, 78.7% of the women described themselves as black- or brown-skinned. One-third of the participants had high school education or less, while 38.1% reported having undergone postgraduate training (Table 1).

Mean scores, obtained by applying either the mFG or the HST questionnaire, did not vary significantly with age, race/skin color, schooling, number of pregnancies, menstrual pattern, history of infertility, prior diagnosis of PCOS, or menopause. Mean scores differed as a function of

BMI, although the differences were only of borderline significance ($P=0.06$ for the mFG and $P=0.05$ for the HST; Table 1).

To analyze the performance of the test, the three most commonly used cut-off points for studies on hirsutism were applied (≥ 6, ≥ 7 and ≥ 8). Analyses of the ROC curves showed that for both current and past body hair statuses evaluated according to the HST, the cut-off point of ≥ 5 points resulted in good accuracy, with very satisfactory sensitivity and specificity levels. The best performance of the HST questionnaire was found with an mFG cut-off point of ≥ 6, as shown by the area under the ROC curve for both moments in time (Fig. 1).

Table 1 Sociodemographic, reproductive, and anthropometric characteristics distributed according to the modified Ferriman–Gallwey (mFG) score and the proposed *Hirsuta* questionnaire (HST).

Variables	n (%)	mFG mean (s.d.)	HST mean (s.d.)
Age (years)			
35–44	19 (21.1)	5.47 (5.8)	4.26 (4.3)
45–54	49 (43.3)	3.54 (2.4)	2.64 (2.5)
55+	32 (35.6)	3.53 (2.9)	2.41 (2.5)
Student's t-test[a]		0.42	0.25
Race/skin color			
Black	48 (78.7)	3.63 (3.0)	2.50 (7.7)
Non-black	13 (21.3)	3.23 (2.7)	2.69 (2.8)
Student's t-test		0.67	0.82
Schooling (years)			
≤11	28 (33.3)	3.79 (2.9)	2.71 (2.7)
>11	56 (66.7)	3.89 (3.9)	2.91 (3.1)
Student's t-test		0.90	0.78
Number of pregnancies			
0–2	43 (51.8)	3.88 (2.9)	3.14 (3.0)
2+	40 (48.2)	3.18 (2.5)	2.13 (1.9)
Student's t-test		0.24	0.07
Menstrual dysfunction[b]			
No	70 (84.3)	3.83 (3.6)	2.86 (3.1)
Yes	13 (15.7)	3.62 (3.2)	2.62 (2.5)
Student's t-test		0.84	0.78
History of infertility[c]			
No	67 (79.8)	3.78 (3.7)	2.70 (2.9)
Yes	17 (20.2)	3.82 (2.9)	3.24 (3.1)
Student's t-test		0.96	0.51
Diagnosis of PCOS[d]			
No	77 (91.7)	3.68 (3.6)	2.69 (2.9)
Yes	7 (8.3)	5.00 (3.2)	4.14 (3.1)
Student's t-test		0.34	0.22
Menopause			
No	55 (65.5)	3.68 (3.5)	2.62 (2.5)
Yes	29 (34.5)	4.21 (4.7)	3.17 (3.7)
Student's t-test		0.43	0.42
BMI (weight (kg)/height (m^2))			
<25	40 (48.2)	3.05 (2.14)	2.08 (1.8)
≥25	43 (51.8)	4.49 (4.4)	3.42 (3.7)
Student's t-test		0.06	0.05

[a]Student's t-test comparing <55 and ≥55 years.
[b]Intervals between menstrual periods were considered regular when ≥25 and <35 days.
[c]One year or more trying to achieve pregnancy without success.
[d]Previous medical diagnosis of polycystic ovary syndrome.

In the simple linear regression analysis, four of the five body areas used in the test were found to be better predictors of the total mFG score, as the predictive capacity of the chest was below the expected value. The thighs, upper lip, and chin were found to be the areas that most contributed toward the overall HST score. In the simultaneous multiple linear regression analysis, all the areas in the HST taken together were able to predict the total mFG score in 97% of cases (Table 2). When the chest was excluded from the analysis, calculation of the area under the ROC curve showed an overall performance of the test of 0.9364 and accuracy of 90%, higher than that of the proposed instrument with five body areas, despite the slight decrease in its ability to predict the total mFG score (adjusted R^2=93%).

Cut-off points	Sensitivity (%)	Specificity (%)	Accuracy (%)	Area under the ROC curve (95%CI)
mFG cut-off point ≥6				0.929 (0.871 - 0.987)
≥4	85.0	82.9	83.3	
≥5	85.0	90.0	88.9	
≥6	60.0	95.7	87.8	
mFG cut-off point ≥7				0.928 (0.867 - 0.990)
≥4	87.5	79.7	81.1	
≥5	87.5	86.5	86.7	
≥6	62.5	93.2	87.8	
mFG cut-off point ≥8				0.922 (0.851 - 0.993)
≥3	92.7	69.7	73.3	
≥4	85.7	77.6	78.9	
≥5	85.7	84.2	84.4	
≥6	64.3	92.1	87.8	
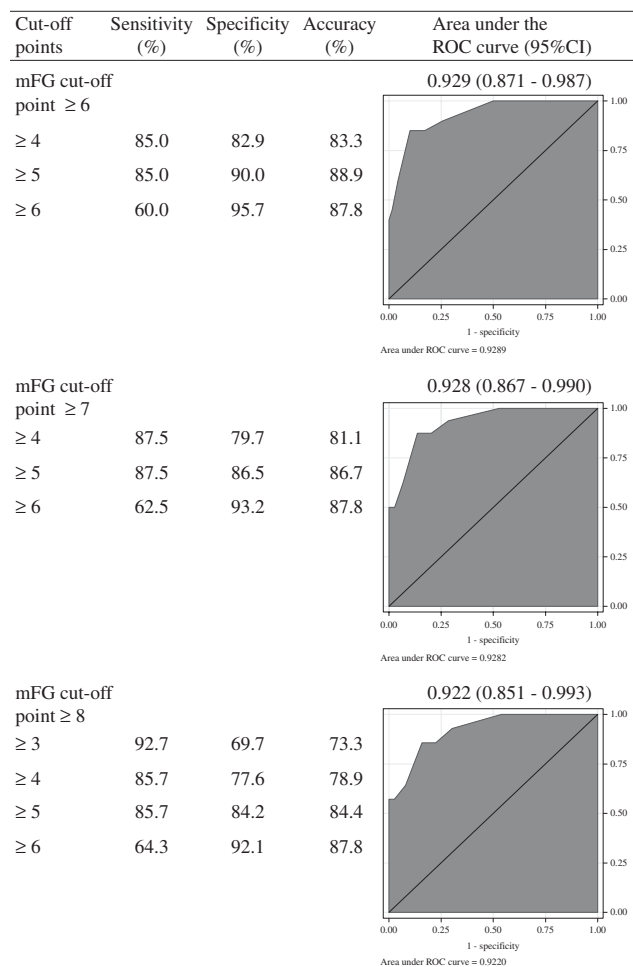
Figure 1
General performance of the *Hirsuta* questionnaire (proposed test with five body areas) at the various cut-off points of the mFG score (gold standard) and analysis of the receiver operating characteristic (ROC) curves.

Considering hirsutism identified by the mFG as the dependent variable of the regression model, for all the cut-off points studied, the five areas that most contributed toward predicting this diagnosis were, in decreasing order, the upper lip, thighs, chin, lower abdomen, and sacroiliac region. Yet again, it was found that four of the areas evaluated in the proposed test (HST) were among them. The three regions of the body that most contributed to hirsutism identified according to the HST score at cut-off points of ≥4 and ≥5 were, also in decreasing order, the upper lip, thighs, and lower abdomen, differing from the cut-off point of ≥6 in which this latter region was substituted by the chin (Table 2).

When the HST scores for the two moments in time, the current score and that for the period before 35 years of age, were compared, a strong, positive Spearman's rank correlation coefficient was found (r=0.82), which was statistically significant (P=0.000), showing that as the current score changes, the score referring to the past also changes. This change was linear and in the same direction. The boxplot shows that both the current and the past scores are different when hirsute and non-hirsute women are compared, hirsutism being defined using the mFG at all the cut-off points evaluated, with paired t-tests indicating statistically significant differences in mean scores (Fig. 2).

When the HST scores were considered in relation to the five corresponding areas scored according to the gold standard (mFG), they were found to remain strongly correlated (Spearman's r=0.84, P=0.000).

For the test–retest analysis, a new variable (M) was created to describe the changes perceived in the quantity of body hair throughout time, which represents the differences between the HST scores for current body hair status and that before 35 years of age. By correlating this difference (M) with the value attributed to the responses given for current status compared with that before 35 years of age, represented in the third question of the questionnaire, a positive, statistically significant Spearman's rank correlation was found (r=0.49; P=0.000). There was also a clearly positive trend toward greater perception of body hair as the difference between the scores at the two moments in time increased (P=0.000). For the new variable, 'M', sensitivity was 95.2% (95% CI: 90.8–99.6%) and specificity 72.5% (95% CI: 63.2–81.7%) for the perception in change of the quantity of body hair over time, as elicited by the third question of the instrument.

Use of the Bland–Altman plot to evaluate agreement between the HST and the mFG revealed a mean difference between the two methods of 0.966 (95% CI: 0.612–1.321).

Table 2 Proportion of prediction (R^2) of each area for the overall score and for the diagnosis of hirsutism at the various mFG and HST cut-off points.

| Body areas | Total score | | Hirsutism (mFG) | | | Hirsutism (HST) | | |
| | mFG | HST | mFG ≥ 6 | mFG ≥ 7 | mFG ≥ 8 | HST ≥ 4 | HST ≥ 5 | HST ≥ 6 |
	R^2	R^2	R^2	R^2	R^2	R^2	R^2	R^2
Thighs	0.55	0.56	0.38	0.43	0.33	0.39	0.36	0.36
Upper lip	0.53	0.62	0.48	0.56	0.50	0.45	0.44	0.44
Chin	0.51	0.38	0.26	0.19	0.19	0.19	0.20	0.28
Lower abdomen	0.44	0.34	0.23	0.17	0.22	0.23	0.27	0.10
Upper abdomen	0.41	–	0.11	0.14	0.17	–	–	–
Upper back	0.36	–	0.04	0.05	0.06	–	–	–
Lower back	0.31	–	0.17	0.24	0.21	–	–	–
Chest	0.25	0.29	0.08	0.04	0.07	0.22	0.21	0.21
Arms	0.15	–	0.05	0.07	0.08	–	–	–
HST	0.97	–	0.65[a]	0.68[a]	0.63[a]	–	–	–

[a]Adjusted R^2.

Graphic analysis indicated that these values were strongly correlated ($r=0.84$; $P=0.00$) and concordant, as the difference was close to zero (Pitman's $r=0.12$), hence not statistically significant ($P=0.28$). Similar results were found in the agreement test between the HST scores for current and past body hair statuses. These measurements were also strongly correlated ($r=0.82$; $P=0.00$) and concordant, as Pitman's $r=0.06$ is also close to zero and not statistically significant ($P=0.59$). In this case, there was a mean difference of 0.61 (95% CI: 0.27–0.95), slightly over half a point (Fig. 3).

The areas in which the amount of body hair most increased over time were the upper lip and the chin (27.8%), whereas a reduction occurred in the lower abdomen area (10%) and on the thighs (7.8%).

Discussion

The purpose of this paper is to present a short instrument for identifying hirsutism, which would be useful in scientific research, particularly in population-based studies. Compared with the gold standard, the sensitivity and specificity of this instrument were high at all the cut-off points evaluated. In addition, test–retest reliability was good, and the instrument was easily self-administered, showing that it is adaptable for use as an online research tool.

Based on the most commonly used scoring scale, the simplification of the test consisted of reducing the number of areas studied and its self-administration, which was made possible through the use of colored, self-explanatory drawings designed to facilitate understanding and make it easier for the respondent to select a score for each body region.

The performance of the test

The best evaluation of the validity of this instrument consisted of the combination of the cut-off points from the gold standard and the test: ≥ 6 and ≥ 5, respectively,
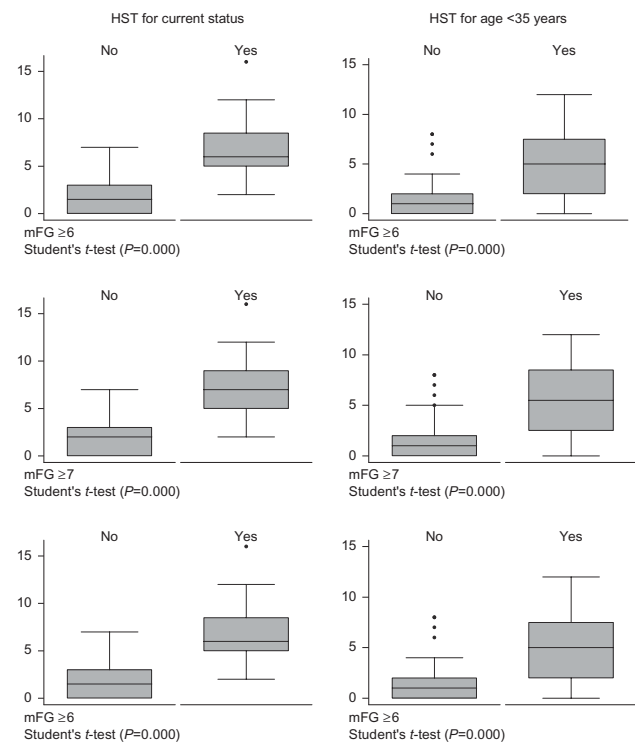
Figure 2

Hirsuta questionnaire (HST) scores for current body hair status and that before age 35 between hirsute and non-hirsute women as defined by several cut-off points of the modified FG (mFG) score.

European Journal of Endocrinology
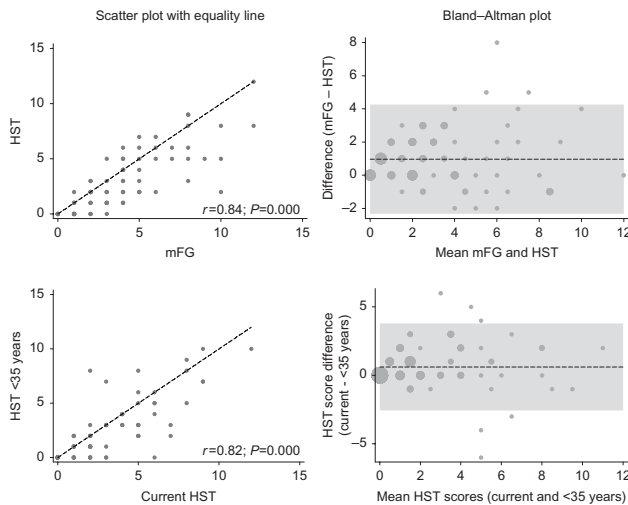
Figure 3

Scatter plots and Bland–Altman plots between the *Hirsuta* questionnaire (HST) and the modified FG (mFG) scores, and between current HST score and that before age 35.

resulting in sensitivity of 85% and specificity of 90%, with a positive predictive value of 70.8% and a negative predictive value of 95.5%. This cut-off point also resulted in the highest proportion of correctly classified tests and the best overall performance of the test. These indicators of validity were higher than those reported in a US study proposing a simplified test in which only one of the two body areas was rated: the chin or the lower abdomen (13, 28). Additionally, the results obtained using the instrument presented in this study proved better than those reported in other recent studies in which simplified tests have also been proposed. One of those studies tested only three areas (29), while another, which also evaluated three regions, used a cut-off point of ≥ 2 (8). As this latter study was conducted exclusively with Chinese patients, in whom the prevalence of hirsutism is extremely low compared with women in other countries, application of this instrument will probably be restricted to that population (8, 15, 18, 30).

The challenge of identifying parts that represent the whole

The heterogeneity of the distribution patterns of body hair involves a multiplicity of combinations of sensitive regions. Women may have excess body hair in one, two, or several of these areas, thus determining various phenotypes. This makes it necessary to include the examination of various body regions in scores designed to quantify hirsutism. In this study, five areas were used that have also been thoroughly evaluated in scores proposed by another two investigators, with the present findings being consistent with their reports (3, 31). In this study, one important aspect is particularly noteworthy: of the five areas evaluated, four were found to be highly predictive of the total score (upper lip, chin, lower abdomen, and thighs), both with respect to the gold standard and to the instrument proposed in this study, whereas one (the chest) contributed only modestly to the overall score. These data are consistent with findings reported by other authors (32). When only these four areas were considered, using the same cut-off point, the number of correctly classified tests was greater.

Marks from the past in the present

This questionnaire was tested in women aged 35–72 years, irrespective of hormone status. As a result, it was found necessary to include a question on the quantification of body hair before that age interval, a time at which body hair is at peak abundance (3, 17). The answer to this question resulted in high correlation rates and predictive values with respect to current body hair, as verified using the gold standard. With this finding, it is reasonable to consider that body hair status in the present is not an independent variable and that, by quantifying current body hair, information may be obtained regarding body hair in the past. This information may be relevant for identifying women who were exposed to a hyperandrogenic state over a considerable period of time, with all the cardiometabolic consequences that may result from this hormonal environment (33, 34). Data from this study confirm that, over time, facial hair increases in women, while on the thighs body hair decreases. This finding is consistent with the reports of other investigators who have stated that body hair growth, particularly facial hair, is highly dependent on age (9, 17). It may be that, as with men, women may develop physiological or non-physiological androgen sensitivity over time, in a caudocephalic direction, with hair growth initially on the legs, thighs, and abdomen, continuing toward the face and terminating on the scalp, where androgenic action results in hair loss (4).

In relation to the findings reported in this study, some observations have to be made. By including a question on body hair status in the past, which, depending on the participant's age, may have been many years previously, the possibility of answers in the present instrument that cannot be verified using the gold standard is introduced.

An attempt was made to resolve this problem by evaluating the correlation between the gold standard and the current answers and these with responses regarding past status, with this latter correlation showing positive results. One study demonstrated that hirsute women tend to overestimate the amount of their body hair; however, this was not confirmed in this study (19). It is important to emphasize that these findings refer to a population composed of university staff in Salvador, Brazil. The women evaluated had a particular ethnic profile compared with the other centers participating in the ELSA-Brasil study, with a greater proportion of black- and brown-skinned women in the sample (24). Additionally, the women had a higher education level compared with the general population. Taken together, these are aspects that may have an impact on the prevalence of hyperandrogenic signs (18, 35).

Self-administration: a desirable characteristic for large-scale use

The feasibility of self-administration, one of the principal objectives behind the development of this questionnaire, was confirmed by the strong correlation between the test and the gold standard, in addition to the agreement level between the two instruments, which was three times that described in the literature (12). In relation to the score attributed by the women themselves and that attributed by clinicians, another study reported a correlation that was much lower than that found in this study, even when the areas most sensitive to androgens were taken into consideration (19).

In the evaluation of agreement between the test and the gold standard, one aspect merits particular consideration. The difference between the two measurements may result from the fact that the mFG establishes an overall score by grading nine body areas, whereas the HST measures only five of these areas. It is reasonable to believe that two methods with different spectrums of possible measurements (mFG: 0–36; HST: 0–20) will result in different scores, and it is possible that the mFG score is higher than that obtained with the HST. The difference of 0.966 shown in this study represents approximately one point and appears to be acceptable for the purpose of differentiating between hirsute and non-hirsute women.

Conclusion

The HST represents an important contribution to scientific research, not only with respect to hirsutism, but also principally to PCOS, offering a mechanism of large-scale quantification of a variable that is very important to identify and much more challenging to obtain. There are advantages such as the fact that it can be self-administered, that it is appropriate for women both before and after menopause, and that it presents good validity. As hirsutism is one of the diagnostic criteria for PCOS, identifying it at more advanced ages may allow studies of its association with cardiovascular disease to be conducted in the age group in which this outcome is most common. Therefore, this instrument may contribute to an advance in the study of PCOS and its consequences, benefiting the significant percentage of women affected by this condition.

Author contribution statement

L Gabrielli worked on the conception of the study and on the protocol, coordinated and supervised data collection, carried out the statistical analysis and interpretation of data, and drafted the manuscript. EMLA contributed to the analysis, worked on the discussion of the results, and reviewed the manuscript. Both authors read and approved the final manuscript.

References

1 Cebeci F, Onsun N & Mert M. Insulin resistance in women with hirsutism. *Archives of Medical Science* 2012 **8** 342–346. (doi:10.5114/aoms.2012.28563)

2 Escobar-Morreale HF. Diagnosis and management of hirsutism. *Annals of the New York Academy of Sciences* 2010 **1205** 166–174. (doi:10.1111/j.1749-6632.2010.05652.x)

3 Hassa H, Tanir HM, Yildirim A, Senses T, Eskalen M & Mutlu FS. The hirsutism scoring system should be population specific. *Fertility and Sterility* 2005 **84** 778–780. (doi:10.1016/j.fertnstert.2005.03.047)

4 Rosenfield RL. Pilosebaceous physiology in relation to hirsutism and acne. *Clinics in Endocrinology and Metabolism* 1986 **15** 341–362. (doi:10.1016/S0300-595X(86)80029-9)

5 Randall VA. Androgens and hair growth. *Dermatologic Therapy* 2008 **21** 314–328. (doi:10.1111/j.1529-8019.2008.00214.x)

6 Azziz R, Sanchez LA, Knochenhauer ES, Moran C, Lazenby J, Stephens KC, Taylor K & Boots LR. Androgen excess in women: experience with over 1000 consecutive patients. *Journal of Clinical Endocrinology and Metabolism* 2004 **89** 453–462. (doi:10.1210/jc.2003-031122)

7 Paparodis R & Dunaif A. The hirsute woman: challenges in evaluation and management. *Endocrine Practice* 2011 **17** 807–818. (doi:10.4158/EP11117.RA)

8 Li R, Qiao J, Yang D, Li S, Lu S, Wu X & Wei Z. Epidemiology of hirsutism among women of reproductive age in the community: a simplified scoring system. *European Journal of Obstetrics, Gynecology, and Reproductive Biology* 2012 **163** 165–169. (doi:10.1016/j.ejogrb.2012.03.023)

9 Ferriman D & Gallwey JD. Clinical assessment of body hair growth in women. *Journal of Clinical Endocrinology and Metabolism* 1961 **21** 1440–1447. (doi:10.1210/jcem-21-11-1440)

10 Hatch R, Rosenfield RL, Kim MH & Tredway D. Hirsutism: implications, etiology, and management. *American Journal of Obstetrics and Gynecology* 1981 **140** 815–830.

11 Yildiz BO. Diagnosis of hyperandrogenism: clinical criteria. *Best Practice & Research. Clinical Endocrinology & Metabolism* 2006 **20** 167–176. (doi:10.1016/j.beem.2006.02.004)

12 Wild RA, Vesely S, Beebe L, Whitsett T & Owen W. Ferriman Gallwey self-scoring I: performance assessment in women with polycystic ovary syndrome. *Journal of Clinical Endocrinology and Metabolism* 2005 **90** 4112–4114. (doi:10.1210/jc.2004-2243)

13 Yildiz BO, Bolour S, Woods K, Moore A & Azziz R. Visually scoring hirsutism. *Human Reproduction Update* 2010 **16** 51–64. (doi:10.1093/humupd/dmp024)

14 Chen X, Yang D, Mo Y, Li L, Chen Y & Huang Y. Prevalence of polycystic ovary syndrome in unselected women from southern China. *European Journal of Obstetrics, Gynecology, and Reproductive Biology* 2008 **139** 59–64. (doi:10.1016/j.ejogrb.2007.12.018)

15 Tehrani FR, Rashidi H & Azizi F. The prevalence of idiopathic hirsutism and polycystic ovary syndrome in the Tehran Lipid and Glucose Study. *Reproductive Biology and Endocrinology* 2011 **9** 144. (doi:10.1186/1477-7827-9-144)

16 Redmond GP. Androgenic disorders of women: diagnostic and therapeutic decision making. *American Journal of Medicine* 1995 **98** 120S–129S. (doi:10.1016/S0002-9343(99)80070-9)

17 Lunde O & Grøttum P. Body hair growth in women: normal or hirsute. *American Journal of Physical Anthropology* 1984 **64** 307–313. (doi:10.1002/ajpa.1330640313)

18 Gabrielli L & Aquino EM. Polycystic ovary syndrome in Salvador, Brazil: a prevalence study in primary healthcare. *Reproductive Biology and Endocrinology* 2012 **10** 96. (doi:10.1186/1477-7827-10-96)

19 Espinós JJ, Calaf J, Estadella J & Checa MA. Hirsutism scoring in polycystic ovary syndrome: concordance between clinicians' and patients' self-scoring. *Fertility and Sterility* 2010 **94** 2815–2816. (doi:10.1016/j.fertnstert.2010.05.022)

20 Rothman MS & Wierman ME. How should postmenopausal androgen excess be evaluated? *Clinical Endocrinology* 2011 **75** 160–164. (doi:10.1111/j.1365-2265.2011.04040.x)

21 Shaw LJ, Bairey Merz CN, Azziz R, Stanczyk FZ, Sopko G, Braunstein GD, Kelsey SF, Kip KE, Cooper-Dehoff RM, Johnson BD *et al*. Postmenopausal women with a history of irregular menses and elevated androgen measurements at high risk for worsening cardiovascular event-free survival: results from the National Institutes of Health-National Heart, Lung, and Blood Institute sponsored Women's Ischemia Syndrome Evaluation. *Journal of Clinical Endocrinology and Metabolism* 2008 **93** 1276–1284. (doi:10.1210/jc.2007-0425)

22 Endogenous Hormones and Breast Cancer Collaborative Group. Circulating sex hormones and breast cancer risk factors in postmenopausal women: reanalysis of 13 studies. *British Journal of Cancer* 2011 **105** 709–722. (doi:10.1038/bjc.2011.254)

23 Diamanti-Kandarakis E, Lambrinoudaki I, Economou F, Christou M, Piperi C, Papavassiliou AG & Creatsas G. Androgens associated with advanced glycation end-products in postmenopausal women. *Menopause* 2010 **17** 1182–1187. (doi:10.1097/gme.0b013e3181e170af)

24 Aquino EM, Barreto SM, Bensenor IM, Carvalho MS, Chor D, Duncan BB, Lotufo PA, Mill JG, Molina MC, Mota EL *et al*. Brazilian Longitudinal Study of Adult Health (ELSA-Brasil): objectives and design. *American Journal of Epidemiology* 2012 **175** 315–324. (doi:10.1093/aje/kwr294)

25 Rosenfield RL. Clinical practice. Hirsutism. *New England Journal of Medicine* 2005 **353** 2578–2588. (doi:10.1056/NEJMcp033496)

26 Bland JM & Altman DG. Measuring agreement in method comparison studies. *Statistical Methods in Medical Research* 1999 **8** 135–160. (doi:10.1191/096228099673819272)

27 Hirakata VC & Camey SA. Análise de concordância entre métodos Bland–Altman. *Revista do Hospital de Clínicas de Porto Alegre* 2009 **29** 8.

28 Knochenhauer ES, Hines G, Conway-Myers BA & Azziz R. Examination of the chin or lower abdomen only for the prediction of hirsutism. *Fertility and Sterility* 2000 **74** 980–983. (doi:10.1016/S0015-0282(00)01602-2)

29 Cook H, Brennan K & Azziz R. Reanalyzing the modified Ferriman–Gallwey score: is there a simpler method for assessing the extent of hirsutism? *Fertility and Sterility* 2011 **96** 1266–1270. (doi:10.1016/j.fertnstert.2011.08.022)

30 Zhao J-L, Chen Z-J, Shi Y-H, Geng L, Ma Z-X, Li Y & Tang R. Investigation of body hair assessment of Chinese women in Shandong region and its preliminary application in polycystic ovary syndrome patients. *Chinese Journal of Obstetrics and Gynecology* 2007 **42** 5.

31 Moncada-Lorenzo E. Familial study of hirsutism. *Journal of Clinical Endocrinology and Metabolism* 1970 **31** 556–564. (doi:10.1210/jcem-31-5-556)

32 Api M, Badoglu B, Akca A, Api O, Gorgen H & Cetin A. Interobserver variability of modified Ferriman–Gallwey hirsutism score in a Turkish population. *Archives of Gynecology and Obstetrics* 2009 **279** 473–479. (doi:10.1007/s00404-008-0747-8)

33 Practice Committee of the American Society for Reproductive Medicine. The evaluation and treatment of androgen excess. *Fertility and Sterility* 2004 **82** (Suppl 1) S173–S180.

34 Cakir E, Doğan M, Topaloglu O, Ozbek M, Cakal E, Vural MG, Yeter E & Delibasi T. Subclinical atherosclerosis and hyperandrogenemia are independent risk factors for increased epicardial fat thickness in patients with PCOS and idiopathic hirsutism. *Atherosclerosis* 2013 **226** 291–295. (doi:10.1016/j.atherosclerosis.2012.11.004)

35 Merkin SS, Azziz R, Seeman T, Calderon-Margalit R, Daviglus M, Kiefe C, Matthews K, Sternfeld B & Siscovick D. Socioeconomic status and polycystic ovary syndrome. *Journal of Women's Health* 2011 **20** 413–419. (doi:10.1089/jwh.2010.2303)

European Journal of Endocrinology