Validation of the American Society for Reproductive Medicine guidelines/recommendations in white European men presenting for couple's infertility

Eugenio Ventimiglia, M.D.,[a,b] Paolo Capogrosso, M.D.,[a,b] Luca Boeri, M.D.,[a] Silvia Ippolito, M.D.,[a] Roberta Scano, B.S.,[a] Marco Moschini, M.D.,[a] Giorgio Gandaglia, M.D.,[a,b] Enrico Papaleo, M.D.,[c] Francesco Montorsi, M.D.,[a,b] and Andrea Salonia, M.D., Ph.D.[a,b]

[a] Division of Experimental Oncology/Unit of Urology, Urological Research Institute, Istituto di Ricerca e Cura a Carattere Scientifico Ospedale San Raffaele; [b] Università Vita-Salute San Raffaele; [c] Infertility Unit, Unit of Obstetrics/Gynecology, Istituto di Ricerca e Cura a Carattere Scientifico Ospedale San Raffaele, Milan, Italy

Objective: To retrospectively validate the American Society for Reproductive Medicine (ASRM) guidelines/recommendations concerning endocrine evaluation in a cohort of white European men presenting for couple's infertility.
Design: Retrospective study.
Setting: Academic reproductive medicine outpatient clinic.
Patient(s): Cohort of 1,056 consecutive infertile men (noninterracial infertile couples).
Intervention(s): Testicular volume was assessed with a Prader orchidometer. Serum hormones were measured (8–10 A.M.) in all cases. Hypogonadism was defined as total T < 3 ng/mL, according to the Endocrine Society definition. Semen analysis values were assessed based on the 2010 World Health Organisation reference criteria.
Main Outcome Measure(s): ASRM indications for endocrine assessment in infertile men (sperm concentration <10 million/mL, impaired sexual function, and other clinical findings suggesting a specific endocrinopathy) were used to predict hypogonadism in our cohort. Moreover, a clinically user-friendly three-item nomogram was developed to predict hypogonadism and was compared to the ASRM guidelines assessment.
Result(s): Biochemical hypogonadism was diagnosed in 156 (14.8%) men. Overall, 669 (63.4%) patients would have necessitated total T assessment according to the ASRM criteria; of these, only 119 (17.8%) were actually hypogonadal according to the Endocrine Society classification criteria. Conversely, 37 (23.7%) out of 156 patients with biochemical hypogonadism would have been overlooked. The overall predictive accuracy, sensitivity, and specificity of the ASRM guidelines was 58%, 76%, and 39%, respectively. Our nomogram was not reliable enough to predict hypogonadism, despite demonstrating a significantly higher predictive accuracy (68%) than the ASRM guidelines.
Conclusion(s): The current findings show that the ASRM guidelines/recommendations for male infertility workup may not be suitable for application in white European infertile men. (Fertil Steril® 2016;106:1076–82. ©2016 by American Society for Reproductive Medicine.)
Key Words: Male infertility, guidelines, semen parameters, hormones, testosterone

Discuss: You can discuss this article with its authors and with other ASRM members at https://www.fertstertdialog.com/users/16110-fertility-and-sterility/posts/11005-validation-of-the-american-society-for-reproductive-medicine-guidelines-recommendations-in-white-european-men-presenting-for-couple-s-infertility

Received February 3, 2016; revised and accepted June 29, 2016; published online July 26, 2016.
E.V. has nothing to disclose. P.C. has nothing to disclose. L.B. has nothing to disclose. S.I. has nothing to disclose. R.S. has nothing to disclose. M.M. has nothing to disclose. G.G. has nothing to disclose. E.P. has nothing to disclose. F.M. has nothing to disclose. A.S. has nothing to disclose.
Reprint requests: Andrea Salonia, M.D., Ph.D., Università Vita-Salute San Raffaele, Division of Experimental Oncology/Unit of Urology, Urological Research Institute, IRCCS Ospedale San Raffaele, Via Olgettina 60, 20132 Milan, Italy (E-mail: salonia.andrea@hsr.it).

Fertility and Sterility® Vol. 106, No. 5, October 2016 0015-0282/$36.00

http://dx.doi.org/10.1016/j.fertnstert.2016.06.044

Male factor infertility still remains a poorly understood disease; although extensively studied, up to 30%–45% of cases remain unexplained (1). This number is significant considering the fact that outcomes in infertility management are often rather frustrating and that nonidiopathic causes might be treatable (1). Therefore, a proper diagnostic workup would be advantageous in terms of both etiologic assessment and, above all, subsequent effective treatment (2). Several attempts have been made to achieve standardized and evidence-based criteria for the evaluation of the infertile male. Nevertheless, available recommendations and guidelines are of poor support to date. If, on one hand, international guidelines substantially agree concerning semen evaluation in the infertile man, things somehow become less clear when dealing with hormone determination. Considering the European Association of Urology (EAU) guidelines (1), despite suggesting a hormone assessment for infertile men with primary testicular failure, no clear indications are provided regarding when/how hormone measures should be used for patient assessment. The American Society of Reproductive Medicine (ASRM) also lacks guidelines related to this issue but published a 2015 updated Committee Opinion about diagnostic evaluation of infertile men (3). This document clearly states that the identification and treatment of correctable conditions will improve the male partner's fertility and allow conception to be achieved naturally. The ASRM recommendations for endocrine evaluation of infertile men are straightforward and include an initial evaluation with total T (tT) and FSH for men with [1] abnormal semen parameters, particularly when the sperm concentration is <10 million/mL; [2] impaired sexual function; or [3] other clinical findings that suggest a specific endocrinopathy. Further workup (second early morning measurement of tT, serum free T, LH, and PRL) remains indicated whenever tT is found to be <3 ng/mL. To the best of our current knowledge, neither these recommendations nor the EAU guidelines have been externally validated.

Considering these aspects, a strong need emerges for shared and reliable evidence-based support for the basic management of the infertile male in everyday real-life clinical practice. Thus, the aim of our study was to retrospectively validate the ASRM recommendations for endocrine evaluation in the infertile man in a cohort of white European men presenting at the same academic outpatient clinic for couple's infertility and to develop a novel nomogram capable of predicting tT levels <3 ng/mL.

MATERIALS AND METHODS

Patients

The analyses of this cross-sectional study were based on a cohort of 1,056 consecutive white European men assessed at a single academic center for primary couple's infertility (noninterracial infertile couples only) between September 2005 and January 2015. Patients were enrolled if they were ≥ 18 year old and had either male factor infertility or mixed factor infertility. Male factor infertility was defined after a comprehensive diagnostic evaluation of the female partners. According to the World Health Organisation (WHO) criteria, infertility is defined as not conceiving a pregnancy after at least 12 months of unprotected intercourse regardless of whether or not a pregnancy ultimately occurred (4). Secondary infertility was defined as the inability to conceive after a previous pregnancy (4).

During the office visit, patients were assessed by means of a physical examination and a thorough self-reported medical history including age, comorbidities, and previous drug prescriptions. Comorbidities were then scored with the Charlson Comorbidity Index (CCI) (5). We used the International Classification of Diseases, 9th revision, because its coding algorithms were used to define the 17 comorbidities that constitute the most widely used CCI score. For the specific purpose of the analysis, CCI was categorized as 0, 1, ≥ 2. Body mass index (BMI), defined as weight in kilograms by height in square meters, was measured for each patient. Testes volume was assessed using a Prader orchidometer. Patients underwent at least two consecutive semen analyses, both showing at least one parameter below standard values for normal semen parameters according to the WHO criteria (6).

A complete hormone profile was requested and obtained for every patient as per our academic department internal guidelines; according to these, venous blood samples were drawn from each patient between 7 A.M. and 11 A.M. after an overnight fast. FSH and LH were measured using a heterogeneous competitive magnetic separation assay (Bayer Immuno 1 System, Bayer). Inhibin B (InhB) was measured with an enzyme-linked immunosorbent assay (Beckman Coulter AMH Gen II ELISA). Total T levels were measured via a direct chemiluminescence immunoassay (ADVIA Centaur; Siemens Medical Solutions Diagnostics), and sex hormone–binding globulin levels were measured via a solid-phase chemiluminescent immunometric assay on Immulite 2000 (Medical Systems SpA). Calculated free testosterone was derived from the Vermeulen formula (7). Hypogonadism was defined as tT $<$ 3 ng/mL (8). The same laboratory was used for all patients.

All the above listed data were then collected in our prospectively maintained institutional database. Data collection followed the principles outlined in the Declaration of Helsinki; all patients had signed an informed consent agreeing to deliver their own anonymous information for future studies. The study was approved by our local ethics committee.

Statistical Analyses

Statistical analysis consisted of different steps. First, descriptive statistics focused on frequencies and proportions for categorical variables. Medians and interquartile (IQ) ranges were reported for continuously coded variables. The chi-square test and Mann-Whitney *U*-test were used to compare proportions and medians, respectively. Second, the ASRM recommendations of [1] abnormal semen parameters, particularly when the sperm concentration is <10 million/mL; [2] impaired sexual function (assessed through the International Index of Erectile Function questionnaire (9), the International Society for Sexual Medicine 2014 definition of premature ejaculation (10), and the previous study by Salonia et al. (11)); or [3] other

clinical findings that suggest a specific endocrinopathy (this last point was implemented mainly focusing on the Endocrine Society clinical practice guidelines (8); the presence of those clinical signs was evaluated in each patient by a single expert with competence in both urology and endocrinology [A.S.]) were retrospectively validated in our population, evaluating their accuracy as for the area under the curve (AUC) estimates. Third, univariable (UVA) and multivariable (MVA) logistic regression models predicting the presence of biochemical hypogonadism were fitted, including age, BMI, mean testicular size, and azoospermia as covariates. Mean testicular size was chosen neither to underestimate nor to overstate the impact of varicocele that could occur when choosing a single testis volume; preliminary analyses showed significant correlation among left, right, and mean testicular volume (data not shown). A logistic regression–based nomogram predicting hypogonadism in our population was then developed relying on the obtained regression coefficients. Our multivariate model was designed in a way to be reader friendly and useful at the first clinical evaluation; as a matter of fact, it includes variables easily accessible with a simple physical evaluation and a semen analysis. The AUC was used to quantify the tool's predictive accuracy (PA). The nomogram was subjected to 200 bootstrap resamples for reduction of overfit bias and for internal validation. Bootstrap-corrected AUC of the model was then compared to the ASRM recommendations using the DeLong method (12). The extent of overestimation or underestimation of the biochemically confirmed versus the nomogram-predicted hypogonadism rates was graphically explored using a calibration plot. Fourth, we used decision curve analyses (13) to determine the clinical net benefit of the two prediction models.

RESULTS

Table 1 lists the characteristics and descriptive statistics of the entire cohort of patients as segregated according to gonadal status. Overall, biochemical hypogonadism was diagnosed in 156 (14.8%) men. Hypogonadal patients had significantly higher BMI and FSH values and lower testicular volume (all $P<.03$). Likewise, hypogonadal individuals showed lower values for sperm concentration and progressive motility (all $P<.03$) but higher rates of nonobstructive azoospermia patients ($P<.001$) as compared with eugonadal men.

Considering the ASRM recommendations, a sperm concentration of <10 million/mL, impaired sexual function, and other clinical findings suggesting a specific endocrinopathy were found in 490 (46.4%), 221 (20.9%), and 591 (56%) men, respectively. Overall, 669 (63.4%) patients would have deserved tT assessment according to the ASRM guidelines. Of those, 119 (17.8%; 11.3% of the total cohort of patients) were actually hypogonadal according to the Endocrine Society classification criteria. Conversely, hypothetically relying on the ASRM guidelines, 37 (23.7%) out of 156 patients with biochemical hypogonadism would not have been candidates for further endocrine assessment. Moreover, 550 (52.1% of the total) patients would have been candidates for hormone workup without having low tT values. As a whole, the ASRM recommendations' sensitivity, specificity, and PA were 76.3%, 39%, and 58%, respectively.

Table 2 details logistic regression models predicting hypogonadism in our cohort. At UVA, patient BMI ($P<.001$), mean testicular volume ($P=.004$), and azoospermia ($P<.001$) were independent predictors of tT < 3 ng/mL. Similar findings were observed at MVA (all $P\leq.05$). After 200 bootstrap resamples, the PA value for the MVA model was 68% ($P<.001$ as compared with ASRM PA).

Figure 1A displays our logistic regression–based nomogram predicting tT < 3 g/mL. The calibration plot (Fig. 1B) shows predicted probabilities against the observed hypogonadism rate. The decision curve analysis (Fig. 1C) shows how our predictive models clearly display a higher clinical net benefit compared with the ASRM recommendations.

DISCUSSION

We retrospectively validated the ASRM recommendations for endocrine assessment in infertile males, comparing them with a novel clinic-friendly logistic regression–based nomogram. The ASRM recommendations were tested in a large cohort of white European men consecutively presenting at the same academic outpatient clinic for primary couple's infertility. Total T values were assessed in every patient, diagnosing biochemical hypogonadism whenever tT < 3 ng/mL was found. The ASRM recommendations were later challenged in this population to establish their performances. What emerged was a not completely satisfactory performance of those recommendations in our setting of white European men. Moreover, our newly developed nomogram, although showing some improvements, did not achieve clinically significant results (Supplemental Table 1).

Endocrine milieu assessment in the infertile male is of major importance since hypogonadism might coexist with infertility in up to 10%–16% of cases (14). However, the questions are whether every patient seeking medical help for infertility would actually benefit from an endocrine workup and what the proper timing is for such an assessment. Available guidelines do not seem to agree in this regard. In this context, the EAU guidelines stress the importance of hormone values in selected clusters of patients (e.g., Klinefelter, primary testicular failure), without any clear recommendation, especially for all other groups of infertile men (1). Unfortunately, little evidence exists to support this statement. A complete hormonal workup, besides being of undoubtful diagnostic utility, is of fundamental importance in terms of deciding on further therapeutic options: glaring is the case of secondary hypogonadism (15, 16), where substantial therapeutic efficacy can be achieved. Things get less clear and more frustrating for both patients and clinicians when it comes to primary hypogonadism (1). To date, no study has addressed the clinical predictors/criteria for selecting patients harboring clinically significant endocrine alterations among infertile men. Therefore, the current indications/recommendations are poorly supported by clinical evidence or lack appropriate clinical data (17–19). Moreover, none of the currently available guidelines have yet been externally validated. For this reason, we considered it important to

TABLE 1

Characteristics and descriptive statistics of patients according to patients' gonadal status.

Characteristic	Eugonadal (n = 900)	Hypogonadal (n = 156)	*P* value[a]	Overall
Patients, n	900	156		
Age, y				
Median	36	36.5	.25	36
IQ range	33–40	32–40		33–40
BMI, kg/m^2				
Median	25.11	27.18	<.001	25.31
IQ range	23.36–27.41	24.55–30.17		23.46–27.78
CCI (continuously coded), n (%)				
Median	0	0	.50	0
IQ range	0–0	0–0		0–0
CCI, n (%)				
CCI 0	834 (92.7)	139 (89.1)	.28; $\chi^2 = 2.57$	973 (92.1)
CCI 1	35 (3.9)	10 (6.4)		45 (4.3)
CCI ≥2	31 (209)	7 (4.5)		38 (3.6)
Testis volume (mL), Prader estimation				
Right testis				
Median	20	20	.016	20
IQ range	15–25	12–20		12–25
Left testis	20	15	.014	15
Median	12–20	10–20		12–20
IQ range				
FSH, mIU/mL				
Median	5.0	5.7	.023	5.1
IQ range	3.0–9.1	2.9–12.4		3.0–9.4
LH, mIU/mL				
Median	4.0	3.5	.61	3.9
IQ range	2.8–5.7	2.0–5.7		2.7–5.7
InhB, pg/mL				
Median	112.0	80.0	.21	106.3
IQ range	58.0–169.3	35.1–132.3		53.0–165.9
tT, ng/mL				
Median	5.02	2.50	–	4.67
IQ range	4.06–6.29	2.12–2.80		3.50–5.93
Semen volume, mL				
Median	3.0	3.0	.94	3.0
IQ range	2.0–4.0	1.8–4.0		2.0–4.0
Semen volume <1.5 mL, n (%)	127 (14.1)	28 (18.0)	.21; $\chi^2 = 1.56$	16 (15.6)
Sperm concentration, 10^6/mL[b]				
Median	13.0	12.9	.005	13.0
IQ range	3.1–38.5	2.4–29.0		3.0–37.5
Sperm concentration <15 × 10^6/mL, n (%)[b]	411 (53)	67 (52.8)	.97; $\chi^2 = 0.01$	478 (52.9)
Progressive motility[b]				
Median	25	15	.023	25
IQ range	10–38	4–36		9–38
Progressive motility <32%, n (%)[b]	521 (67.1)	93 (73.2)	.17; $\chi^2 = 1.86$	614 (68.0)
Normal morphology[b]				
Median	4	3.5	.10	4
IQ range	0–15	0–13.5		0–15
Normal morphology <4%, n (%)[b]	434 (55.9)	72 (56.7)	.87; $\chi^2 = 0.03$	506 (56.0)
Nonobstructive azoospermia, n (%)	92 (10.1)	32 (20.5)	<.001; $\chi^2 = 13.59$	124 (11.7)
Obstructive azoospermia, n (%)	24 (2.7)	5 (3.2)	.70; $\chi^2 = 0.14$	29 (2.7)
ASMR+, n (%)	550 (61.1)	119 (76.3)	<.001; $\chi^2 = 13.2$	669 (63.4)
Endocrinopaties, n (%)	476 (52.9)	115 (73.7)	<.001; $\chi^2 = 23.4$	591 (56)
Sexual dysfunctions, n (%)	187 (20.8)	34 (21.8)	.77; $\chi^2 = 0.1$	221 (20.9)
Sperm concentration <10 × 10^6/mL, n (%)	411 (45.7)	79 (50.6)	.25; $\chi^2 = 1.32$	490 (46.4)

Note: ASRM+: patients deserving hormonal workup according to ASRM recommendations. BMI = body mass index; CCI = comorbidity index; InhB = inhibin B; IQ = interquartile; tT = total testosterone.

[a] *P* value according to Mann-Whitney *U*-test or χ^2-test, as indicated.

[b] Values obtained considering 903 nonazoospermic men.

Ventimiglia. Validation of the ASRM recommendations. Fertil Steril 2016.

validate the aforementioned recommendations in a large contemporary homogenous cohort of consecutive primary infertile patients. If, on one hand, the EAU guidelines are not straightforward to validate due to their intrinsic vagueness, the ASRM recommendations, on the other hand, were found to have poor accuracy (AUC, 58%). Considering

TABLE 2

Logistic regression models predicting hypogonadism (tT < 3 ng/mL) in the whole cohort of patients (N = 1,056).

	UVA		MVA	
Variable	**OR (CI)**	***P* value**	**OR (CI)**	***P* value**
BMI, kg/m^2	1.15 (1.10–1.20)	<.001	1.15 (1.09–1.21)	<.001
Mean testis size, mL (Prader)	0.95 (0.92–0.99)	.004	0.97 (0.93–1.00)	.05
Azoospermia	2.10 (1.38–3.19)	<.001	1.77 (1.05–2.99)	.0 3

Note: BMI = body mass index; CI = confidence interval; MVA = multivariable; OR = odds ratio; tT = total testosterone; UVA = univariable.

Ventimiglia. Validation of the ASRM recommendations. Fertil Steril 2016.

our results, roughly one out of four patients with biochemical hypogonadism would have missed a proper diagnosis if the ASRM recommendations were applied in our population, with a simultaneous consistent rate of unnecessarily screened patients. There might be a number of explanations for this lack of accuracy. The ASRM guidelines recommend clinical intervention based on several aspecific signs/symptoms of hypogonadism/endocrinopathy. For instance, low T signs/symptoms were mainly identified in the context of late-onset hypogonadism (17, 19), which should be regarded as a different clinical entity from infertile male hypogonadism. Therefore, the use of criteria mainly developed for late-onset hypogonadism (19) in infertile men is questionable, at the very least. Infertile patients are usually considered a relatively young population, although a consistent increase in paternal age has been recently described (20). According to our findings, age does not seem to differ in the presence of endocrine abnormalities, calling into question one of the foremost indicators of hypogonadism in the general population (21).

The Endocrine Society includes small testis size (<5 mL) among the signs of low T level (8); this criterion was one of those used in our analysis to implement the ASRM statement "other clinical signs suggestive for a specific endocrinopathy." According to our data, only 38.2% of biochemically hypogonadal patients have at least one testis smaller than 5 mL (data not shown). Of clinical importance, such a cutoff is not suitable in this setting, thus supporting a more complex predictive model.

Of importance, the recommendations made by the ASRM are addressed mainly to U.S. physicians. This is not of scant relevance since American and European men might differ significantly for several factors and these differences may be responsible for the suboptimal performance of American recommendations in our cohort. This is the case for obesity and metabolic syndrome: their respective prevalence among men in the United States is consistently higher than what is observed in European men (22, 23), although direct comparative data are lacking. This is of paramount importance since obesity and its pathological correlates are capable of modulating hormone homeostasis and its deriving alterations (24).

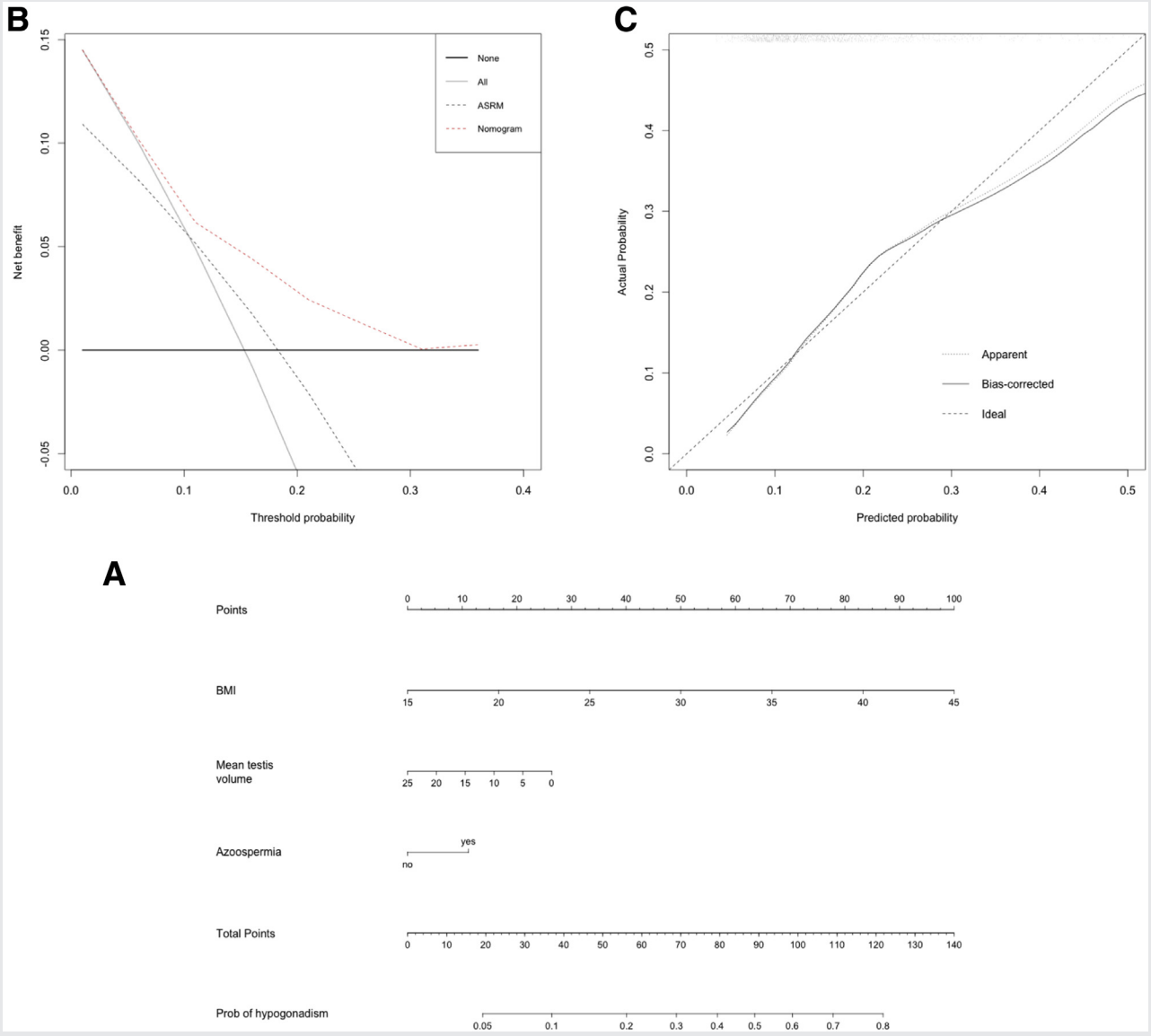
For these reasons, we secondly aimed to develop a novel risk prediction tool capable of detecting hypogonadal patients by relying on easily obtainable clinical variables. Such variables, which have been included in our study, were derived from a review of the published data and therefore, at first sight, resemble those included in the guidelines. Another condition, deemed fundamental for our variables, was that they be easily and promptly obtainable during the male infertility workup. As a matter of fact, a physical examination and a semen analysis provide clinical predictors included in our model (i.e., BMI, testis size, azoospermia). We did not include clinical symptoms suggestive of low T for two main reasons: [1] available questionnaires were not validated in the infertile (and therefore younger) population and [2] we wanted to make our model as objective as possible. Moreover, azoospermia was chosen rather than a particular categorized sperm parameter to minimize both intra- and interpersonal differences and laboratory biases. Remarkably, when trying to add other seminal parameters to the predictive model, no gain in PA was observed ($P>.05$ at DeLong test; data not shown). In spite of the aforementioned caveats, when trying to predict tT < 3 ng/mL in our population, the newly developed nomogram was not astonishingly better than the ASRM guidelines (overall PA 68%).

The decision curve analyses allowed us to make some interesting considerations as the inferior performance of the ASRM recommendations clearly emerged; their clinical net benefit was not only partly inferior to the "treat-all option," but also consistently worse compared with our nomogram for all reasonable threshold probabilities. This clearly demonstrates that using the ASRM indications for endocrine assessment in the infertile man would be, if not properly harmful, at least entirely nonbeneficial compared with the other aforementioned strategies.

Several factors might account for the poor performance displayed by our nomogram. First, the most accurate predictor of tT < 3 ng/mL in our cohort, BMI, accounted only for a 66% of PA (data not shown). We were not able to find any other more accurate predictor among clinical and seminal variables. Variables known to be associated with hypogonadism in the general population lost clinical usefulness in our sample. This might be related to patient age. As previously mentioned, no relationship emerged between low tT levels and age in our cohort. This may be due to both the relative homogeneity observed in our population in terms of age and to the fact that age is not the main pathophysiological drive of hypogonadism in infertile men. Moreover, younger age allows for a shorter period of time for low tT effects to manifest. Second, Scovell et al. (25) suggested that a serum tT threshold of 4 ng/mL might be associated with hypogonadal symptoms in young men; according to the authors, the Endocrine Society cutoff of 3 ng/mL would exclude a discreet share of patients from the hypogonadal population. This would add another clue to explaining the low diagnostic power of the several tools analyzed in this study. Third, impaired semen parameters, often used as a proxy of low tT values (3, 8), have a very limited clinical utility in our setting; moreover, owing to their value fluctuations, an isolated sperm analysis may be unreliable or at least misleading.

Thus, according to what emerged from our data, trying to [1] predict hypogonadism in infertile white European men to

FIGURE 1



(**A**) Logistic regression–based nomogram predicting tT < 3 ng/mL. (**B**) Calibration plot showing predicted probabilities against the observed hypogonadism rate. (**C**) Decision curve analyses assessing the effect of ASRM guidelines and of our nomogram in detecting hypogonadism in our population. The *red dashed line* represents the endocrine workup based on our nomogram, the *black dashed line* represents the endocrine workup based on ASRM guidelines, and the *continuous line* represents the endocrine workup of all patients.

Ventimiglia. Validation of the ASRM recommendations. Fertil Steril 2016.

therefore [2] tailor hormonal screening among infertile men is definitely a tough and intricate task.

Our study is not devoid of limitations. This was a hospital-based study, raising the possibility of selection biases. Patients were recruited from a single academic outpatient clinic, and despite the fact that this particular study likely included a consistently large and homogeneous white European cohort of primary infertile men (restricted to noninterracial infertile couples), several larger studies across different centers and populations will be needed to substantiate our findings. As a matter of fact, the ASRM guidelines are, perhaps implicitly, designed for U.S. practitioners, and hence are improper for the European population. Second, the analyses were implemented cross-sectionally, and a comparison with a same-race, age-matched cohort of fertile individuals is lacking. Third, the analyses offer no data regarding the potential molecular alterations in spermatogenesis, which might be of importance in investigating possible markers of decreased tT values. Fourth, ASRM recommendations are rather ambiguous, with no clear-cut indication of screening for hypogonadism and therefore a difficult standard to compare with; despite this, we explored this range of ambiguity and found out that neither the 10 million spermatozoa/mL criterion nor our model was capable of diagnosing

hypogonadism, thus corroborating the concept of a baseline endocrine screening in every infertile man. Fifth, we focused on tT values and a tT-derived definition of hypogonadism; in this context, it is of major importance that hypogonadism should be regarded as a clinical syndrome rather than a biochemical diagnosis; however, until male factor infertility–related hypogonadism gains its own nosological dignity, this issue remains unfortunately unaddressed.

CONCLUSIONS

Hypogonadism diagnosis and management are challenging tasks within the field of male factor infertility. Unfortunately, available clinical studies and recommendations offer little support to clinicians. In our study, we reported how ASRM recommendations, besides being supported by scant published evidence, have an unsatisfactory clinical performance when applied to a white European cohort of infertile men. Owing to this, we developed a novel predictive tool aimed at identifying patients at risk of having tT $<$ 3 ng/mL; our nomogram performed consistently better than the ASRM guidelines in identifying hypogonadal men, despite not having a clear-cut impacting accuracy. Actually, owing to [1] the intrinsic difficulty underlying a priori identification of hypogonadal patients, [2] the major importance of a proper hormonal workup, and [3] the limited social cost of a basic infertility hormonal workup, consideration of a hormone assessment to be conducted for every infertile patient might be an appropriate clinical option.

Acknowledgments: The authors thank Dana Kuefner, Ph.D., for reviewing the language in this manuscript.

REFERENCES

1. Jungwirth A, Giwercman A, Tournaye H, Diemer T, Kopa Z, Dohle G, et al. European Association of Urology guidelines on male infertility: the 2012 update. Eur Urol 2012;62:324–32.
2. Pierik FH, Van Ginneken AM, Dohle GR, Vreeburg JT, Weber RF. The advantages of standardized evaluation of male infertility. Int J Androl 2000;23:340–6.
3. Practice Committee of the American Society for Reproductive Medicine. Diagnostic evaluation of the infertile male: a committee opinion. Fertil Steril 2015;103:e18–25.
4. World Health Organization web chapter on couple's infertility. Available at: http://www.who.int/reproductivehealth/topics/infertility/definitions/en/. Accessed April 1, 2015.
5. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373–83.
6. World Health Organisation. Laboratory manual for the examination and processing of human semen. 5th ed. Geneva: World Health Organisation; 2010.
7. Vermeulen A, Verdonck L, Kaufman JM. A critical evaluation of simple methods for the estimation of free testosterone in serum. J Clin Endocrinol Metab 1999;84:3666–72.
8. Bhasin S, Cunningham GR, Hayes FJ, Matsumoto AM, Snyder PJ, Swerdloff RS, et al. Testosterone therapy in men with androgen deficiency syndrome: an Endocrine Society clinical practice guidelines. J Clin Endocrinol Metab 2010;95:2536–59.
9. Cappelleri JC, Rosen RC, Smith MD, Mishra A, Osterloh IH. Diagnostic evaluation of the erectile function domain of the International Index of Erectile Function. Urology 1999;54:346–51.
10. Serefoglu EC, McMahon CG, Waldinger MD, Althof SE, Shindel A, Adaikan G, et al. An evidence-based unified definition of lifelong and acquired premature ejaculation: report of the second International Society for Sexual Medicine Ad Hoc Committee for the Definition of Premature Ejaculation. J Sex Med 2014;11:1423–41.
11. Salonia A, Clementi MC, Ventimiglia E, Colicchia M, Capogrosso P, Castiglione F, et al. Prevalence and predictors of concomitant low sexual desire/interest and new-onset erectile dysfunction—a picture from the everyday clinical practice. Andrology 2014;2:702–8.
12. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–45.
13. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26:565–74.
14. Andrology. In: Nieschlag E, Behre HM, Nieschlag S, editors. Male reproductive health and dysfunction. Berlin: Springer Verlag; 2010.
15. Whitten SJ, Nangia AK, Kolettis PN. Select patients with hypogonadotropic hypogonadism may respond to treatment with clomiphene citrate. Fertil Steril 2006;86:1664–8.
16. Martin-du-Pan RC, Bischof P. Increased follicle stimulating hormone in infertile men. Is increased plasma FSH always due to damaged germinal epithelium? Hum Reprod 1995;10:1940–5.
17. Hall SA, Esche GR, Araujo AB, Travison TG, Clark RV, Williams RE, et al. Correlates of low testosterone and symptomatic androgen deficiency in a population-based sample. J Clin Endocrinol Metab 2008;93:3870–7.
18. Bhasin S, Woodhouse L, Casaburi R, Singh AB, Bhasin D, Berman N, et al. Testosterone dose-response relationships in healthy young men. Am J Physiol Endocrinol Metab 2001;281:E1172–81.
19. Wu FC, Tajar A, Beynon JM, Pye SR, Silman AJ, Finn JD, et al. Identification of late-onset hypogonadism in middle-aged and elderly men. N Engl J Med 2010;363:123–35.
20. Salonia A, Matloob R, Saccà A, Ferrari M, Gallina A, Castiglione F, et al. Are Caucasian-European men delaying fatherhood? Results of a 7 year observational study of infertile couples with male factor infertility. Int J Androl 2012;35:125–32.
21. European Association of Urology (EAU) guidelines on male hypogonadism. Available at: http://uroweb.org/wp-content/uploads/18-Male-Hypogonadism_LR1.pdf. Accessed June 1, 2016.
22. Ervin RB. Prevalence of metabolic syndrome among adults 20 years of age and over, by sex, age, race and ethnicity, and body mass index: United States, 2003–2006. Natl Health Stat Rep 2009;13:1–7.
23. Corona G, Lee DM, Forti G, O'Connor DB, Maggi M, O'Neill TW, et al. Age-related changes in general and sexual health in middle-aged and older men: results from the European Male Ageing Study (EMAS). J Sex Med 2010;7:1362–80.
24. Michalakis K, Mintziori G, Kaprara A, Tarlatzis BC, Goulis DG. The complex interaction between obesity, metabolic syndrome and reproductive axis: a narrative review. Metabolism 2013;62:457–78.
25. Scovell JM, Ramasamy R, Wilken N, Kovac JR, Lipshultz LI. Hypogonadal symptoms in young men are associated with a serum total testosterone threshold of 400 ng/dL. BJU Int 2015;116:142–6.

SUPPLEMENTAL TABLE 1

Specificity and sensitivity for our nomogram according to the several probability thresholds for predicting hypogonadism analyzed.

Threshold probability	Specificity	Sensitivity
1%	0.00	1.00
2%	0.00	1.00
5%	0.01	1.00
10%	0.29	0.90

Note: Although the 10% cutoff would allow for a generally acceptable sensitivity, our question would be whether it is worth losing 10% of men at screening considering the relatively low social cost of tT testing.

Ventimiglia. Validation of the ASRM recommendations. Fertil Steril 2016.