



Open Access

ORIGINAL ARTICLE

Validity of self-reported male sexual function scales in a young Chinese population: a comparative study with clinician-assisted evaluation

Hui Zhang^{1,2,3}, Elena Colonnello^{2,4}, Erika Limoncin², Tommaso B Jannini⁵, Xu-Chong Tu^{1,3}, Andrea Sansone², Emmanuele A Jannini², Yan Zhang¹

Psychometric scales, commonly used to gauge sexual function, can sometimes be influenced by response biases. In our research from June 2020 to April 2021, we examined the accuracy of self-reported sexual function scales. We invited patients from the Department of Infertility and Sexual Medicine at the Third Affiliated Hospital of Sun Yat-sen University (Guangzhou, China), who have male sexual dysfunction, to participate by filling out a self-reported version of a specific questionnaire. In addition, they went through a clinician-assisted version of this questionnaire, encompassing tools such as the Premature Ejaculation Diagnostic Tool (PEDT), the 6-item International Index of Erectile Function (IIEF-6), the Erection Hardness Scale (EHS), and the Masturbation Erection Index (MEI). Using the clinician-assisted version as a reference, we categorized patients and applied various statistical methods, such as the Chi-square test, intraclass correlation coefficient (ICC), logistic regression, and the Bland–Altman plot, to gauge reliability. In our study with 322 participants, we found that while there were no notable discrepancies in error rates based on our categorization, certain scales showed significant differences in terms of overestimation and underestimation, with the exception of the PEDT. The positive diagnosis rate consistency between the self-reported and clinician-assisted versions was observed. High ICC values between the two versions across the scales were indicative of remarkable reliability. Our findings show that the self-reported versions of tools such as EHS, IIEF-6, MEI, and PEDT are credible and hold clinical reliability. However, employing a dual-diagnosis approach might be more prudent to circumvent potential misdiagnoses.

Asian Journal of Andrology (2024) 26, 321–327; doi: 10.4103/aja202364; published online: 22 December 2023

Keywords: psychometry; scale; sexual function; sexual health

INTRODUCTION

Psychometric scales are scientifically and clinically relevant tools used by physicians and psychologists to quantitatively assess patients' sexual health^{1–3} by means of patient-reported outcomes (PROs).⁴ Male sexual function can be assessed by several psychometric scales: the 6-item International Index of Erectile Function (IIEF-6), the Erection Hardness Scale (EHS), and the Masturbation Erection Index (MEI) can be used to assess aspects related to erectile function; moreover, the Premature Ejaculation Diagnostic Tool (PEDT) can be used to assess the presence of premature ejaculation (PE) and its clinical consequences.^{5–10} Several studies have confirmed the diagnostic accuracy of the EHS and IIEF-6.^{11–15} The MEI is an adaptation of the IIEF-6 that is applied to masturbation; this scale is particularly useful for erectile dysfunction (ED) patients who have not engaged in penetrative sexual intercourse in the past 4 weeks.⁷ Conversely, the PEDT is the most commonly used 6-item psychometric tool to evaluate PE.¹⁶

All these psychometric scales appear to be easy-to-administer epidemiological tools that are useful for collecting large amounts

of data via web-based surveys,^{17–19} establishing the efficacy of drugs acting on male sexual function,^{20–22} and performing self-assessments in daily clinical practice.²³ During the validation process, all tests were demonstrated to possess good psychometric characteristics, such as reliability and validity. However, the quality of the collected data largely depends on the degree of the individual's understanding of each item, which can be influenced by the education level,²⁴ age,²⁵ occupation,²⁶ social desirability (*i.e.*, the tendency for people to present themselves in a generally favorable fashion),^{27,28} and other factors.^{29–31}

More targeted studies are needed to confirm the validity of these scales when they are administered online without the support of a physician or a psychologist as well as to determine the prevalence of misdiagnoses. Indeed, the validity of self-reporting has been previously examined for cancer,³² cardiovascular disease,^{33,34} and other diseases,^{35–37} but not for sexual function. However, perhaps more than in the previously mentioned fields, the topic of sexuality may specifically lead to desirability bias, *i.e.*, the tendency of survey respondents to

¹Department of Infertility and Sexual Medicine, Third Affiliated Hospital of Sun Yat-sen University, Guangzhou 510000, China; ²Chair of Endocrinology and Medical Sexology (ENDOSEX), Department of Experimental Medicine, University of Rome Tor Vergata, Rome 00133, Italy; ³Department of Urology, First Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou 510000, China; ⁴Section of Medical Pathophysiology, Food Science and Endocrinology, Sapienza University of Rome, Rome 00133, Italy; ⁵School of Psychiatry, Department of Systems Medicine, University of Rome Tor Vergata, Rome 00133, Italy.

Correspondence: Dr. Y Zhang (zhxm@sina.com)

Received: 11 May 2023; Accepted: 12 October 2023

answer questions in a manner that will be viewed favorably by peers or supervisors, which further advocates for specific studies.³⁸

Based on the potential bias of patient self-administered questionnaires, we further clarified the validity of patient self-administered questionnaires. We aimed to perform a comparative study between patient self-reported sexual function scales (self-reported version [SRV]) and the version filled out with the help of professional caregivers (clinician-assisted version [CAV]).

PARTICIPANTS AND METHODS

Participants

This study retrospectively examined participants who visited the Department of Infertility and Sexual Medicine at the Third Affiliated Hospital of Sun Yat-sen University (Guangzhou, China) for sexual dysfunctions such as PE, ED, or infertility between June 1, 2020 and April 1, 2021. Participants included in the study completed self-reported sexual function questionnaires and were assessed by professional caregivers. All participants had engaged in or attempted sexual intercourse and masturbation within the past 4 weeks. In addition, all study participants are identified as heterosexual.

Procedure

In accordance with standard procedures of the Department of Infertility and Sexual Medicine at the Third Affiliated Hospital of Sun Yat-sen University, participants completed the EHS, IIEF-6, MEI, and PEDT on a mobile phone platform with the assistance of an outpatient terminal (Mozi 3.0, Shenzhen Aomila Technology Co., Ltd., Shenzhen, China). The report was then submitted to the server and saved as the SRV. We used the Chinese versions of all four scales. Among them, the Chinese version of the PEDT has been validated.³⁹ However, even though the EHS and IIEF-6 are widely used, the Chinese versions of these scales have not been formally validated. On the other hand, the MEI is not commonly used in the Chinese language context. Therefore, to minimize bias, we chose the linguistically validated Chinese version of the PEDT. For the EHS, IIEF-6, and MEI, we referred to published studies^{15,40} and contacted the original authors.⁷ This engagement provided us with essential insights into the interpretation and application of these scales in a Chinese context. Although we did not develop official Chinese versions, our approach, including collaboration with language experts and andrology specialists, ensured that our adaptations of these scales were as close to the original in intent and meaning as possible within the cultural context.

Subsequently, a trained physician individually explained the meaning of each item in the scale and asked participants about their medical histories in detail, thereby assisting participants in examining each item. Based on the information provided by the participants, the physician helped the participants reselect previously incorrect ones and kept the answers that did not require changes. To avoid potential bias, during the checking process, the terminal was only visible to the physician, and participants could not see their previous answers. This report was then submitted to the server as the CAV.

With the CAV as the control, the number of errors was counted according to each of the five variables, including age, occupation, education level, and scores on the PEDT and IIEF-6 (for instance, if a patient was above a clinical threshold on the IIEF-6 in his SRV but not his CAV, he received a positive misdiagnosis score of 1 for the IIEF-6). We then subtracted the SRV score from the CAV score for each patient in each scale. If the result was negative, meaning that the SRV score was higher than the CAV score, considering the latter as

the standard, the patient likely overestimated their sexual function in the self-report. On the other hand, a positive difference implied an underestimation. However, since the PEDT is a reverse-scored scale, a negative difference was interpreted as an underestimation of sexual function, and a positive difference was interpreted as an overestimation. The rates of over- and underestimation of the four self-reported scales were calculated. EHS scores ≤ 2 ,⁴¹ IIEF-6 scores ≤ 25 ,⁴² MEI scores ≤ 25 ,⁷ and PEDT scores ≥ 11 ⁸ were regarded as positive diagnosis results. Cases with positive SRV and negative CAV were identified as false positives, whereas those with negative SRV and positive CAV were identified as false negatives. These rates of the SRV questionnaires were calculated. To ensure consistency of CAV, all participants were recruited by the same physician (HZ).

Statistical analyses

The error rates were compared by Chi-square or Fisher's exact test for each of the five variables. Logistic regression was used to explore the relationship between the above variables and the error rate. The rates of over- and underestimation of the four self-reported scales were calculated. The Chi-square test or Fisher's exact test was used to determine significant differences between the overestimated and underestimated rates for each scale. These rates of the SRV questionnaires were calculated. The intraclass correlation coefficient (ICC) of the two questionnaires was also calculated. The above analysis was conducted using SPSS 20.0 statistical software (SPSS, Inc., Chicago, IL, USA). The consistency of both the SRV and CAV of the EHS, IIEF-6, MEI, and PEDT was evaluated by Bland-Altman plots (MedCalc Software Ltd., Ostend, Belgium).⁴³ These plots show the difference between the SRV and CAV over the average of the two measures. The limits of agreement (LoA) were computed as the mean difference $\pm 1.96 \times$ standard deviation (s.d.) and represented the extent of under- and overreporting of self-reported values compared to auxiliary-reported values.

Ethical statement

The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The trial was conducted in accordance with the Declaration of Helsinki (as revised in 2013). All data from participants included in this study have been de-identified. The study was approved by the Institutional Ethics Board of the Third Affiliated Hospital of Sun Yat-sen University (Approval No. [2021] 02-249-01). The Ethics Committee exempted the study from informed consent requirements.

RESULTS

This study included data from 322 participants (**Table 1**). Their primary reasons for seeking treatment were ED, PE, infertility, or other sexual dysfunctions, such as low libido. With the CAV as the control, we found that the overall error rate of the SRV was 14.0% ($n = 45$), and the error rate was relatively high among those older than 40 years ($\geq 30\%$). Regarding occupation, physicians and students (bachelor's degree and above) had the lowest error rates (0), whereas workers, those in the service industry, and civil servants had the highest, reaching above 20%. Regarding education, participants with master's degrees had the lowest error rate (9.5%), and junior high school diploma holders had the highest (18.9%). Participants with overt PE (PEDT ≥ 11) had similar error rates as those in whom PE was suspected (PEDT ≤ 10) and those with normal ejaculatory control (PEDT ≤ 8). There was a similar trend among the severity of ED (mild, moderate, and severe) as measured by the IIEF-6 scores

as well. However, there were no significant differences within the groups ($P > 0.05$; **Table 2**).

The underestimation rates of the EHS, IIEF-6, MEI, and PEDT in CAV were 3.1% ($n = 10$), 5.3% ($n = 17$), 6.8% ($n = 22$), and 2.5% ($n = 8$), respectively; the overestimation rates were 0.6% ($n = 2$), 1.2% ($n = 4$), 0.3% ($n = 1$), and 1.6% ($n = 5$), respectively. There was a significant difference between the two groups in overestimation and underestimation rates for the different scales (all $P < 0.05$) except for the PEDT ($P = 0.401$; **Table 2**). The logistic regression results showed that the error rate was not affected by age, occupation, education, IIEF-6 score, or PEDT score (all $P > 0.05$; **Table 3**).

Overall, there was no significant difference in the positive diagnosis rate between the SRV and CAV groups ($P > 0.05$). By comparing the CAV, the false-positive and false-negative rates of the four scales in the SRV questionnaire were calculated. The false-positive rates of the EHS, IIEF-6, MEI, and PEDT were 2.8% ($n = 9$), 1.6% ($n = 5$), 4.0% ($n = 13$), and 0.3% ($n = 1$), respectively; the false-negative rates were 0 ($n = 0$), 0 ($n = 0$), 0.3% ($n = 1$), and 1.2% ($n = 4$), respectively (**Table 4**).

The ICC evaluates the reproducibility of the SRV and CAV. The ICC for the EHS, IIEF-6, MEI, and PEDT was 0.899 ($P < 0.001$), 0.966 ($P < 0.001$), 0.931 ($P < 0.001$), and 0.990 ($P < 0.001$), respectively, which shows excellent reproducibility (**Table 5**).

In the Bland–Altman plots of the EHS, IIEF-6, MEI, and PEDT, 1.9% ($n = 6$), 2.8% ($n = 9$), 4.3% ($n = 14$), and 2.8% ($n = 9$) were outside 95% of the LoA, respectively. Their differences were mainly within the consistency boundary and approximately 0. Therefore, the SRV and CAV are considered to have a good level of agreement (**Figure 1**).

DISCUSSION

As quantitative evaluation tools, psychometric scales are widely used in various health-related fields. Sexual medicine has a particular debt to psychometry. Epidemiology and clinical definitions of sexual dysfunctions as well as the validation of drugs used for sexual health have been produced based on popular sexual inventories such as the EHS, IIEF-6, MEI, or PEDT.^{5–8} Moreover, standardized psychometric tools can surely facilitate diagnosis and therapeutic follow-up by physicians not particularly trained in sexual medicine. Finally, self-filled questionnaires may help patients to express their

needs without the shame of having to talk about sexually sensitive topics.

However, in front of so many lights, sexual psychometry may have a number of shadows. Individuals use “Doctor Google” to produce self-diagnoses and, in the best case, to find a specialist using the web;⁴⁴ but also to search for counterfeit drugs on the same Internet.⁴⁵ Although seeing a specialist is the best option, in many cases, patients may, in fact, feel the need to make a preliminary self-assessment, and this often occurs via the web.^{13,46}

In this study, the EHS, IIEF-6, MEI, and PEDTs were utilized to compare the validity of self-reported scales both individually and overall. For individuals, we analyzed the potential cognitive biases based on age, education, occupation, PEDT, and IIEF-6 scores.^{29–31} We found that there was no association between age and error rate. Age is an important factor affecting cognition.²⁵ In general, the more age increases, the more cognitive decline occurs.^{25,47} However, perhaps since the age range of our included population was between 19 years and 59 years, no significant age-related differences in error rates⁴⁷ were found in our study. Regarding occupational classification, the different occupations did not show significant differences in error rates either. Interestingly, there was no significant difference between the groups based on education level, and all were within 20%. This may suggest that the wording of the psychometric tools examined here is relatively well understood independent of education level. The classification analysis of the PEDT and IIEF-6 also indicates that the different scores, meaning the presence or absence of either PE or ED, will not affect the error rate. Together with the results of logistic regression, it can be inferred that age, level of academic qualifications, occupation, PEDT, and IIEF-6 did not significantly affect the self-reported scores.

To further investigate whether participants would overestimate or underestimate their sexual function, we calculated the proportion of overestimation and underestimation of the participant’s sexual function in SRV separately, using CAV as a standard. The results showed that, with the exception of the PEDT, all results of the EHS, IIEF-6, and MEI indicated that participants were more inclined to underestimate their sexual function. This could be due to the particular ability to explore the intimacy of the sexual medicine

Table 1: Baseline data of all patients included in the study

Variable	Value
Age (year), median (IQR)	30 (27–34)
Marital status, n (%)	
Married	127 (39.4)
Unmarried	183 (56.8)
Divorced	12 (3.7)
BMI (kg m^{-2}), median (IQR)	22.7 (20.6–24.5)
Main complaint, n (%)	
ED	157 (48.8)
PE	221 (68.6)
Infertility	12 (3.7)
Others	9 (2.8)
IIEF-6, median (IQR)	25 (20–28)
PEDT, median (IQR)	12 (8–15)
EHS, median (IQR)	3 (3–4)
MEI, median (IQR)	26 (21–29)

IQR: interquartile range; BMI: body mass index; ED: erectile dysfunction; PE: premature ejaculation; IIEF-6: 6-item International Index of Erectile Function; PEDT: Premature Ejaculation Diagnostic Tool; EHS: Erection Hardness Scale; MEI: Masturbation Erection Index

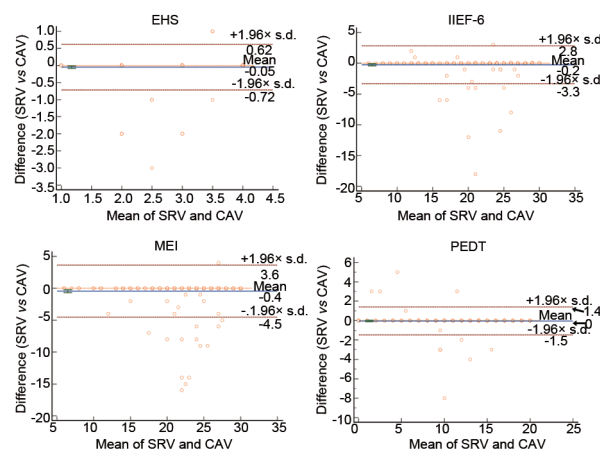


Figure 1: Bland–Altman plots for the EHS, IIEF-6, MEI, and PEDT. The solid line indicates the mean difference, and the dashed lines indicate the limits of agreement (mean difference \pm 1.96 \times standard deviation). EHS: Erection Hardness Scale; IIEF-6: 6-item International Index of Erectile Function-6; MEI: Masturbation Erection Index; PEDT: Premature Ejaculation Diagnostic Tool; SRV: self-reported version; CAV: clinician-assisted version; s.d.: standard deviation.

Table 2: Demographic characteristics and scoring trends of each scale under different classification factors

Classification factor	Participant (n)	Misdeclaration		EHS, n (%)		IIEF-6, n (%)		MEI, n (%)		PEDT, n (%)	
		Participant, n (%)	^a P	-	+	-	+	-	+	-	+
Age (year)			0.190								
>20	1	0 (0)		0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
20–29	143	19 (13.3)		4 (2.8)	0 (0)	8 (5.6)	3 (2.1)	9 (6.3)	1 (0.7)	2 (1.4)	2 (1.4)
30–39	159	20 (12.6)		4 (2.5)	1 (0.6)	8 (5.0)	1 (0.6)	10 (6.3)	0 (0)	6 (3.8)	3 (1.9)
40–49	16	5 (31.3)		1 (6.3)	1 (6.3)	1 (6.3)	0 (0)	3 (18.8)	0 (0)	0 (0)	0 (0)
50–59	3	1 (33.3)		1 (33.3)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Occupation			0.743								
Physician	6	0 (0)		0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Student	8	0 (0)		0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Company employee	65	7 (10.8)		1 (1.54)	0 (0)	5 (7.7)	0 (0)	4 (6.2)	0 (0)	1 (1.5)	1 (1.5)
Unemployed	9	1 (11.1)		0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (11.1)	0 (0)	0 (0)
Individual household	35	4 (11.4)		1 (2.9)	0 (0)	0 (0)	1 (2.9)	3 (8.6)	0 (0)	2 (5.7)	0 (0)
Information technology	31	4 (12.9)		2 (6.5)	1 (3.2)	0 (0)	1 (3.2)	0 (0)	0 (0)	1 (3.2)	0 (0)
Salesman	44	6 (13.6)		2 (4.6)	0 (0)	2 (4.6)	0 (0)	1 (2.3)	0 (0)	1 (2.3)	2 (4.6)
Freelance work	27	5 (18.5)		1 (3.7)	0 (0)	3 (11.1)	1 (3.7)	4 (14.8)	0 (0)	1 (3.7)	1 (3.7)
Service industry	23	5 (21.7)		0 (0)	1 (4.4)	2 (8.7)	0 (0)	1 (4.4)	0 (0)	2 (8.7)	0 (0)
Worker	18	4 (22.2)		1 (5.6)	0 (0)	1 (5.6)	0 (0)	3 (16.7)	0 (0)	0 (0)	0 (0)
Civil servant	14	4 (28.6)		1 (7.1)	0 (0)	2 (14.3)	0 (0)	3 (21.4)	0 (0)	1 (7.1)	1 (7.1)
Others	42	5 (11.9)		1 (2.4)	0 (0)	2 (4.8)	0 (0)	3 (7.1)	0 (0)	1 (2.4)	0 (0)
Education			0.744								
Primary school	6	1 (16.7)		1 (16.7)	0 (0)	0 (0)	0 (0)	1 (16.7)	0 (0)	0 (0)	0 (0)
Junior high school diploma	53	10 (18.9)		2 (3.8)	0 (0)	5 (9.4)	2 (3.8)	6 (11.3)	0 (0)	3 (5.7)	1 (1.9)
High school diploma	49	8 (16.3)		2 (4.1)	0 (0)	2 (4.1)	1 (2.0)	5 (10.2)	0 (0)	1 (2.0)	3 (6.1)
College degree	83	12 (14.5)		2 (2.4)	2 (2.4)	4 (4.8)	1 (1.2)	3 (3.6)	0 (0)	0 (0)	0 (0)
Bachelor's degree	104	11 (10.6)		0 (0)	0 (0)	6 (5.8)	0 (0)	7 (6.7)	1 (1.0)	1 (1.0)	0 (0)
Master's degree	21	2 (9.5)		0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (4.8)	1 (4.8)
PhD	6	1 (16.7)		0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (16.7)	0 (0)
PEDT			0.751								
≤8	91	11 (12.1)		4 (4.4)	0 (0)	5 (5.5)	1 (1.1)	5 (5.5)	1 (1.1)	0 (0)	4 (4.4)
9–10	30	5 (16.7)		1 (3.3)	1 (3.3)	0 (0)	0 (0)	1 (3.3)	0 (0)	1 (3.3)	1 (3.3)
≥11	201	29 (14.4)		5 (2.5)	1 (0.5)	12 (6)	3 (1.5)	16 (8)	0 (0)	7 (3.5)	0 (0)
IIEF-6			0.591								
6–10	13	2 (15.4)		1 (7.7)	0 (0)	0 (0)	0 (0)	1 (7.7)	0 (0)	0 (0)	1 (7.7)
11–16	35	6 (17.4)		1 (2.9)	1 (2.9)	0 (0)	2 (5.7)	4 (11.4)	0 (0)	1 (2.9)	0 (0)
17–21	56	9 (16.1)		2 (3.6)	0 (0)	4 (7.1)	1 (1.8)	4 (7.1)	1 (1.8)	1 (1.8)	2 (3.6)
22–25	63	11 (17.5)		4 (6.4)	1 (1.6)	6 (9.5)	1 (1.6)	4 (6.4)	0 (0)	0 (0)	2 (3.2)
26–30	155	17 (11)		2 (1.3)	0 (0)	7 (4.5)	0 (0)	9 (5.8)	0 (0)	6 (3.9)	0 (0)
Total	322	45 (14)		10 (3.1)	2 (0.6)	17 (5.3)	4 (1.2)	22 (6.8)	1 (0.3)	8 (2.5)	5 (1.6)
^a P					0.020 ^b		0.004 ^b		<0.001 ^b		0.401 ^b

^aP, Fisher's exact test; ^bstatistical values between overestimations and underestimations on different scales. – represents an underestimated score, and + represents an overestimated score. EHS: Erection Hardness Scale; IIEF-6: 6-item International Index for Erectile Function; MEI: Masturbation Erection Index; PEDT: Premature Ejaculation Diagnostic Tool

psychometric tools used herein and explained considering the impact of social desirability.⁴⁸ Social desirability, which refers to individuals' tendency to respond in a manner that is viewed favorably by others, is a common issue when conducting questionnaires and patient consultations. This issue is particularly relevant to the degree of privacy afforded in these interactions.⁴⁹ During the course of patient consultation, patients could exaggerate their condition to receive more attention from their physicians.⁵⁰ In contrast, in some cases, patients can choose a specific response not because it pertains to their experience but because it is supposed to be the most socially accepted response or a response that is supposed by the patient to be the "preferred response" for the physician. In addition, it is necessary to also consider the influence of different cultural backgrounds on the interpretation of the results, which may produce different levels

of social desirability. This may alter the results obtained by sexual psychometry. However, the results of this study show that this risk is statistically quite low.

In fact, comparing the positive diagnostic rates of each scale in CAV and SRV and the repeatability and consistency between the CAV and SRV, evaluated by calculating the ICC and using the Bland–Altman plot, demonstrated excellent repeatability of the four questionnaires evaluated here. In particular, since the Bland–Altman index was <5%, we can see that although there were individual differences between the CAV and SRV, they generally had a good level of agreement. In other words, data obtained from SRV of the EHS, IIEF-6, MEI, and PEDT reflect the real status of the patients as assessed by their physicians. Indeed, even without a specialist who explains the items and provides guidance through the completion, patients were able to correctly

Table 3: Logistic regression of the risk factors for error rate

Classification factor	OR	95% CI	P
Age	1.041	0.982–1.104	0.177
Occupation			
Information technology	1.000	NA	NA
Physician	0	0	0.999
Student	0.661	0.742–0.196	0.694
Company employee	0.925	1.077–0.231	0.940
Unemployed	0.567	0.665–0.164	0.561
Individual household	0.675	1.355–0.328	0.659
Salesman	0.757	0.688–0.065	0.752
Freelance work	0	0	0.999
Service industry	0.826	0.844–0.186	0.779
Worker	0.398	0.532–0.123	0.394
Civil servant	0.473	0.617–0.166	0.434
Others	0.320	2.431–0.421	0.286
Education			
Primary school	1.000	NA	NA
Junior high school diploma	0.949	0.026–34.926	0.977
High school diploma	1.137	0.064–20.225	0.930
College degree	0.998	0.054–18.551	0.999
Bachelor's degree	0.922	0.053–15.929	0.955
Master's degree	0.630	0.039–10.250	0.745
PhD degree	0.491	0.024–10.064	0.645
PEDT	1.011	0.947–1.079	0.756
IIEF-6	0.980	0.931–1.032	0.429

IIEF-6: 6-item International Index for Erectile Function-6; PEDT: Premature Ejaculation Diagnostic Tool; 95% CI: 95% confidence interval; OR: odds ratio

assess their state of sexual health, meaning that these are reliable psychometric measures.

In this study, we aimed to investigate whether patients can effectively conduct preliminary self-diagnosis using self-report questionnaires. We planned to evaluate the comparability of the results from questionnaires on sexual function filled out by patients with those completed with the assistance of a clinical doctor. This finding holds significant importance, as it may indicate whether patients can use these questionnaires as a tool for preliminary self-diagnosis, or even better, for screening. This could ultimately guide them in deciding whether to seek further medical assistance.

The main limitation of this study is the number of participants, which was relatively small for the stratified analysis of different factors. Calculating the incidence of such a small number of stratified factors may, in fact, not represent the true error rate of the group. In addition, the false-positive rate of self-diagnosis of those who were willing to see a doctor may cause score bias, which is also a deficiency of this study.

We also included infertile patients, in whom occasional failure, for example, in the case of programmed sexual intercourse for reproductive needs, may particularly “weigh”. In these patients, occasional failures are often identified as serious sexual dysfunctions by the patients themselves. Potentially, this may increase the discrepancy between SRV and CAV in this subset of patients. Similarly, infertility is a comprehensive problem with a strong link between sexual function and the same potential to influence the patient’s perception of their sexual function, reflected in the outcome, which is the inconsistency between SRV and CAV scores.⁵¹

Last but not least, it should be considered that strong cultural characteristics shape the way sexuality is lived and perceived;⁵²

Table 4: Positive, false-positive, and false-negative rates of different scales in the two questionnaires

Diagnostic result	EHS, n/total (%)			IIEF-6, n/total (%)			MEI, n/total (%)			PEDT, n/total (%)		
	SRV	CAV	P	SRV	CAV	P	SRV	CAV	P	SRV	CAV	P
Positive diagnosis	54/322 (16.8)	45/322 (14.0)	0.325	172/322 (53.4)	167/322 (51.9)	0.693	168/322 (52.2)	156/322 (48.5)	0.344	198/322 (61.5)	201/322 (62.4)	0.808
False positive	9/322 (2.8)	NA	NA	5/322 (1.6)	NA	NA	13/322 (4.0)	NA	NA	1/322 (0.3)	NA	NA
False negative	0/322 (0)	NA	NA	0/322 (0)	NA	NA	1/322 (0.3)	NA	NA	4/322 (1.2)	NA	NA

EHS: Erection Hardness Scale; IIEF-6: 6-item International Index for Erectile Function; MEI: Masturbation Erection Index; PEDT: Premature Ejaculation Diagnostic Tool; NA: not available; SRV: self-reported version; CAV: clinician-assisted version



Table 5: Intraclass correlation coefficient analysis of the self-reported version and clinician-assisted version

Reliability and significance	EHS	IIEF-6	MEI	PEDT
ICC	0.899	0.966	0.931	0.990
P	<0.001	<0.001	<0.001	<0.001

EHS: Erection Hardness Scale; IIEF-6: 6-item International Index for Erectile Function; MEI: Masturbation Erection Index; PEDT: Premature Ejaculation Diagnostic Tool; ICC: intraclass correlation coefficient

therefore, as this study was performed in a single country (China), it can be assumed that some differences in regard to the present study's outcomes could occur elsewhere.

CONCLUSIONS

The SRV of the EHS, IIEF-6, MEI, and PEDT must be considered as reliable and accurate tools. They can help practitioners, as well as researchers, to assess sexual function in a quick and cost-friendly manner. However, the risk of misdiagnosis of these self-administered surveys should not be underestimated, and a careful diagnosis is suggested, especially in certain subsets of patients, such as those who suffer from infertility.

AUTHOR CONTRIBUTIONS

HZ and EC led the study's conceptualization, design, and manuscript drafting. EL, TBJ, and XCT assisted in data collection and analysis. AS and EAJ significantly contributed by providing deep insights into data interpretation. EAJ, in particular, applied his specialized knowledge to unravel complex patterns within the data, offering innovative perspectives that enriched our understanding of the study's findings. YZ supervised the entire research in research design and manuscript revision. All authors read and approved the final manuscript.

COMPETING INTERESTS

EAJ is or has been a paid speaker or consultant for Bayer, Ibsa, Menarini, Otsuka, Pfizer, Shionogi, and Viatrix. AS has been a paid consultant for Menarini. Other authors declare no competing interests.

ACKNOWLEDGMENTS

EAJ and AS have been partially supported for this study by the Italian Ministry of University PRIN (Grant No. 2017S9KTNE_002). HZ was supported by the Scientific Research Project of the Traditional Chinese Medicine Bureau of Guangdong Province (No. 20221086).

REFERENCES

- Corona G, Jannini E, Maggi M. Inventories for male and female sexual dysfunctions. *Int J Impot Res* 2006; 18: 236–50.
- Jannini TB, Rossi R, Sconci V, Bonanni RL, De Michele F, *et al*. Italian validation of Arizona Sexual Experience Scale (ASEX) on patients suffering from psychotic spectrum disorders. *Riv Psichiatr* 2022; 57: 18–22.
- Ciocca G, Jannini TB, Ribolsi M, Rossi R, Niuoli C, *et al*. Sexuality in ultra-high risk for psychosis and first-episode psychosis: a systematic review of literature. *Front Psychiatry* 2021; 12: 750033.
- Peng J, Fang D, Li H, Tang Y, Yuan Y, *et al*. Efficacy of dapoxetine treatment in Chinese patients with premature ejaculation and possible factors affecting efficacy in the real-world practice. *BMC Urol* 2020; 20: 11.
- Rosen RC, Cappelleri J, Smith M, Lipsky J, Pena BM, *et al*. Development and evaluation of an abridged, 5-item version of the International Index of Erectile Function (IIEF-5) as a diagnostic tool for erectile dysfunction. *Int J Impot Res* 1999; 11: 319–26.
- Ramanathan R, Mulhall J, Rao S, Leung R, Salamanca JI, *et al*. Predictive correlation between the International Index of Erectile Function (IIEF) and Sexual Health Inventory for Men (SHIM): implications for calculating a derived SHIM for clinical use. *J Sex Med* 2007; 4: 1336–44.
- Limoncin E, Gravina GL, Lotti F, Maseroli E, Ciocca G, *et al*. The Masturbation Erection Index (MEI): validation of a new psychometric tool, derived from the six-item version of the International Index of Erectile Function (IIEF-6) and from the Erection Hardness Score (EHS), for measuring erectile function during masturbation.

- BJU Int* 2019; 123: 530–7.
- Symonds T, Perelman MA, Althof S, Giuliano F, Martin M, *et al*. Development and validation of a premature ejaculation diagnostic tool. *Eur Urol* 2007; 52: 565–73.
- Hou G, Gao M, Zheng Y, Hou N, Zhang S, *et al*. Nomogram for stratifying patients with lifelong premature ejaculation before using the PHQ-9: an observational study. *Eur J Clin Invest* 2022; 52: e13809.
- Zheng Y, Gao M, Hou G, Hou N, Feng X, *et al*. A prospectively validated nomogram for predicting the risk of PHQ-9 score ≥ 15 in patients with erectile dysfunction: a multi-center study. *Front Public Health* 2022; 10: 836898.
- Matsuda Y, Hisasue Si, Kumamoto Y, Kobayashi K, Hashimoto K, *et al*. Correlation between erection hardness score and nocturnal penile tumescence measurement. *J Sex Med* 2014; 11: 2272–6.
- Parisot J, You R, Salomon L, de la Taille A, Lingombet O, *et al*. Erection hardness score for the evaluation of erectile dysfunction: further psychometric assessment in patients treated by intracavernous prostaglandins injections after radical prostatectomy. *J Sex Med* 2014; 11: 2109–18.
- Capogrosso P, Vertosick EA, Benfante NE, Eastham JA, Scardino PJ, *et al*. Are we improving erectile function recovery after radical prostatectomy? Analysis of patients treated over the last decade. *Eur Urol* 2019; 75: 221–8.
- Xi Y, Colonnello E, Ma G, Limoncin E, Ciocca G, *et al*. Validity of erectile function assessment questionnaires in premature ejaculation patients: a comparative study between the abridged forms of the international index of erectile function and proposal for optimal cutoff redefinition. *J Sex Med* 2021; 18: 440–7.
- Wang C, Zhang H, Liu Z, Tu X, Zhang Y. A modified procedure to diagnose erectile dysfunction using the International Index of Erectile Function (IIEF-6) combined with the Premature Ejaculation Diagnosis Tool (PEDT) via an internet survey. *Sex Med* 2022; 10: 100506.
- Sansone A, Aversa A, Corona G, Fisher A, Isidori A, *et al*. Management of premature ejaculation: a clinical guideline from the Italian Society of Andrology and Sexual Medicine (SIAMS). *J Endocrinol Invest* 2021; 44: 1103–18.
- Bergmann MM, Jacobs EJ, Hoffmann K, Boeing H. Agreement of self-reported medical history: comparison of an in-person interview with a self-administered questionnaire. *Eur J Epidemiol* 2004; 19: 411–6.
- Ekstrm S, Kull I, Nilsson S, Bergstrm A. Web-based self-reported height, weight, and body mass index among Swedish adolescents: a validation study. *J Med Internet Res* 2015; 17: e73.
- Pursey K, Burrows TL, Stanwell P, Collins CE. How accurate is web-based self-reported height, weight, and body mass index in young adults? *J Med Internet Res* 2014; 16: e4.
- Fisher WA, Gruenwald I, Jannini EA, Lev-Sagie A, Lowenstein L, *et al*. Standards for clinical trials in male and female sexual dysfunction: I. phase I to phase IV clinical trial design. *J Sex Med* 2016; 13: 1805–17.
- Fisher WA, Gruenwald I, Jannini EA, Lev-Sagie A, Lowenstein L, *et al*. Standards for clinical trials in male and female sexual dysfunction: II. patient-reported outcome measures. *J Sex Med* 2016; 13: 1818–27.
- Fisher WA, Gruenwald I, Jannini EA, Lev-Sagie A, Lowenstein L, *et al*. Standards for clinical trials in male and female sexual dysfunction: III. unique aspects of clinical trials in male sexual dysfunction. *J Sex Med* 2017; 14: 3–18.
- Fleming CM, Bowden M. Web-based surveys as an alternative to traditional mail methods. *J Environ Manage* 2009; 90: 284–92.
- Lövden M, Fratiglioni L, Glymour MM, Lindenberg U, Tucker-Drob EM. Education and cognitive functioning across the life span. *Psychol Sci Public Interest* 2020; 21: 6–41.
- Murman DL. The impact of age on cognition. *Semin Hear* 2015; 36: 111–21.
- Jorm AF, Rodgers B, Henderson AS, Korten AE, Jacomb PA, *et al*. Occupation type as a predictor of cognitive decline and dementia in old age. *Age Ageing* 1998; 27: 477–83.
- Holden RR, Passy J. Social desirability. In: Leary MR, Hoyle RH, editors. *Handbook of Individual Differences in Social Behavior*. New York: The Guilford Press; 2009. p441–54.
- Leary MR, Hoyle RH. *Handbook of Individual Differences in Social Behavior*. New York: The Guilford Press; 2009.
- Kremen WS, Lachman ME, Pruessner JC, Sliwinski M, Wilson RS. Mechanisms of age-related cognitive change and targets for intervention: social interactions and stress. *J Gerontol A Biol Sci Med Sci* 2012; 67: 760–5.
- Dufouil C, Alperovitch A, Tzourio C. Influence of education on the relationship between white matter lesions and cognition. *Neurology* 2003; 60: 831–6.
- Jefferson AL, Gibbons LE, Rentz DM, Carvalho JO, Manly J, *et al*. A life course model of cognitive activities, socioeconomic status, education, reading ability, and cognition. *J Am Geriatr Soc* 2011; 59: 1403–11.
- Fiederling J, Shams AZ, Haug U. Validity of self-reported family history of cancer: a systematic literature review on selected cancers. *Int J Cancer* 2016; 139: 1449–60.
- Selem SS, Castro MA, César CL, Marchioni DM, Fisberg RM. Validity of self-reported hypertension is inversely associated with the level of education in Brazilian individuals. *Arq Bras Cardiol* 2013; 100: 52–9.
- de Menezes TN, Oliveira EC, de Sousa Fischer MA. Validity and concordance between self-reported and clinical diagnosis of hypertension among elderly residents in

- northeastern Brazil. *Am J Hypertens* 2014; 27: 215–21.
- 35 Matsui D, Yamamoto T, Nishigaki M, Miyatani F, Watanabe I, *et al*. Validity of self-reported number of teeth and oral health variables. *BMC Oral Health* 2017; 17: 1–8.
- 36 Reiniger AP, Londero AB, Ferreira TD, da Rocha JM, Moreira CH, *et al*. Validity of self-reported measures for periodontitis surveillance in a rural sample. *J Periodontol* 2020; 91: 617–27.
- 37 Imai T, Kuwahara K, Miyamoto T, Okazaki H, Nishihara A, *et al*. Validity and reproducibility of self-reported working hours among Japanese male employees. *J Occup Health* 2016; 58: 340–6.
- 38 Burgio G, Giammusso B, Calogero AE, Mollaioli D, Condorelli RA, *et al*. Evaluation of the mistakes in self-diagnosis of sexual dysfunctions in 11,000 male outpatients: a real-life study in an andrology clinic. *J Clin Med* 2019; 8: 1679.
- 39 Huang YP, Chen B, Ping P, Wang HX, Hu K, *et al*. The premature ejaculation diagnostic tool (PEDT): linguistic validity of the Chinese version. *J Sex Med* 2014; 11: 2232–8.
- 40 Zhang K, Xu B, Liu D, Wang X, Zhu J, *et al*. Sildenafil improves erectile hardness in Chinese men with erectile dysfunction: a real-life study analyzed on age stratification. *Urology* 2014; 83: 831–6.
- 41 Mulhall JP, Goldstein I, Bushmakina AG, Cappelleri JC, Hvidsten K. Validation of the erection hardness score. *J Sex Med* 2007; 4: 1626–34.
- 42 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.
- 43 Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 327: 307–10.
- 44 Salem J, Paffenholz P, Bolenz C, von Brandenstein M, Cebulla A, *et al*. Websites on bladder cancer: an appropriate source of patient information? *J Cancer Educ* 2019; 34: 381–7.
- 45 Sansone A, Cuzin B, Jannini EA. Facing counterfeit medications in sexual medicine. A systematic scoping review on social strategies and technological solutions. *Sex Med* 2021; 9: 100437.
- 46 Kimura M, Shimura S, Tai T, Kobayashi H, Baba S, *et al*. A web-based survey of erection hardness score and its relationship to aging, sexual behavior, confidence, and risk factors in Japan. *Sex Med* 2013; 1: 76–86.
- 47 Small BJ, Dixon RA, McArdle JJ. Tracking cognition-health changes from 55 to 95 years of age. *J Gerontol B Psychol Sci Soc Sci* 2011; 66 Suppl 1: i153–61.
- 48 King BM. The influence of social desirability on sexual behavior surveys: a review. *Arch Sex Behav* 2022; 51: 1495–501.
- 49 Dodou D, de Winter JC. Social desirability is the same in offline, online, and paper surveys: a meta-analysis. *Comput Human Behav* 2014; 36: 487–95.
- 50 Wagner G, Miller LG. Is the influence of social desirability on patients' self-reported adherence overrated? *J Acquir Immune Defic Syndr* 2004; 35: 203–4.
- 51 Luca G, Parrettini S, Sansone A, Calafiore R, Jannini E. The Inferto-Sex Syndrome (ISS): sexual dysfunction in fertility care setting and assisted reproduction. *J Endocrinol Invest* 2021; 44: 2071–102.
- 52 Colonnello E, Jannini EA. Impact of Chinese traditional culture and related social norms on current Chinese sexuality and on the future of Chinese sexual medicine. In: Rowland DL, Jannini EA, editors. *Cultural Differences and the Practice of Sexual Medicine: A Guide for Sexual Health Practitioners*. Springer; 2020. p95–113.

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

©The Author(s)(2023)

