ORIGINAL ARTICLE

WILEY

Validity of web-based self-assessment of pubertal development against pediatrician assessments

Xiaoxia Peng[1*] | Yaguang Peng[1*] | Yuchuan Li[2] | Xiaolu Nie[1] | Chunxiu Gong[2] | Di Wu[2] | Xin Ni[1,3]

[1]Center for Clinical Epidemiology and Evidence-based Medicine, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing, China.
[2]Department of Endocrinology, Genetics and Metabolism, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing, China.
[3]Beijing Key Laboratory for Pediatric Diseases of Otolaryngology, Head and Neck Surgery, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing, China

Correspondence
Di Wu, Department of Endocrinology, Genetics and Metabolism, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing 100045, China.
Email: woody0928@sina.com
Xin Ni, Beijing Key Laboratory for Pediatric Diseases of Otolaryngology, Head and Neck Surgery, Beijing Children's Hospital, Capital Medical University, National Center for Children Health, Beijing 100045, China.
Email: nixin@bch.com.cn

*These authors contributed equally to the study.

Funding source
Bureau of Medical Administration and Supervision, National Health and Family Planning Commission of China, only as a unique financial support for this project, has no interest and role in the design and performance, or in the data collection, analysis and interpretation of the results, or in the preparation and approval of this manuscript.

Received: 22 February, 2018
Accepted: 29 July, 2018

ABSTRACT

Importance: A web-based instrument for self-assessment of puberty could be convenient and feasible for large-scale multicenter population-based epidemiological studies for Tanner stages evaluation.

Objective: To validate web-based self-assessment of pubertal development against assessment by a pediatrician.

Methods: Outpatients aged 8–18 years were consecutively recruited in the endocrinology department of Beijing Children's Hospital from October 2016 to August 2017. A web-based self-assessment instrument for pubertal development was introduced to participants by an appointed pediatrician. Tanner stage of puberty was self-assessed by participants in a private environment. Participants were then examined by a senior pediatrician underwent blinded assessment. Weighted kappa and Spearman correlation analyses were conducted to evaluate agreement. The accuracy of the web-based instrument for self-assessment of pubertal onset was evaluated according to sensitivity, specificity, positive predictive value and negative predictive value.

Results: A total of 174 participants (including 82 girls and 92 boys) were assessed consecutively. Correlation coefficients were 0.872 for pubic hair and 0.933 for testicular volume (P<0.001) among boys; a similar result was obtained for the weighted *kappa* value (0.825). For girls, the correlation coefficient and weighted *kappa* for pubic hair was 0.785 and 0.878, respectively. However, breast self-assessment had a medium level of agreement with pediatrician assessment (weighted *kappa*, 0.495; correlation coefficient, 0.643). Moreover, the accuracy of self-assessment in children aged 10 years or above was better than that in children aged less than 10 years.

Interpretation: Assessment of pubertal development using a web-based self-assessment instrument could be less accurate among children aged less than 10 years, especially for girls' breast assessment. Therefore, self-assessment of pubertal development, especially for breast development, should be interpreted cautiously.

KEYWORDS
Puberty, Tanner stage

DOI:10.1002/ped4.12050

INTRODUCTION

Puberty is the specific stage of a child's life when the body starts to become physically mature and the metabolism begins to physiologically change. As children grow and undergo development during puberty, there is considerable turmoil and instability owing to increased concentrations of reproductive hormones.[1] As in the first month after birth, most biochemical markers and hematological markers demonstrate substantial age-specific dynamics at the onset of puberty. For this reason, age partitioning is a critical step in establishing of pediatric reference intervals, according to guidelines for establishing and verifying reference intervals in the clinical laboratory (C28-A3).[2] It has been reported that some hematological markers, such as hemoglobin concentration, red blood cell count, and serum creatinine concentration, undergo an obvious rise in boys after puberty.[3] The concentration of alkaline phosphatase fluctuates downward with the age of pubertal development.[4] For this reason, similar to age partitioning, sexual maturation during this period should be considered when establishing reference intervals for children, especially for some susceptible biomarkers in adolescents. Thus, scientific and accurate assessment of sexual maturation for children or teenagers in pediatric epidemiologic studies is an essential procedure, especially in studies aiming to establish pediatric reference intervals.

As a standard for assessment of development in puberty, Tanner stages of maturity are commonly applied commonly by pediatricians in clinical practice. However, it is not feasible for endocrinologists in epidemiologic studies with large sample sizes to conduct Tanner staging assessment. For large-scale multicenter population-based studies, it would be challenging and impractical for Tanner stages to be assessed in every participant by a single pediatrician, or by several pediatricians with good inter-observer agreement. Moreover, some children and parents might be reluctant to agree to such assessment owing to religious or privacy concerns. Several validation studies have found reasonable agreement between self-reported and pediatrician-assessed Tanner stages using a series of pictures and diagrams.[5, 6]

In this study, a web-based instrument for self-assessment of puberty was developed based on a data capture platform. This platform was prepared for a nationwide large-scale multicenter study aiming to establish pediatric reference intervals in China, the Pediatric Reference Interval of China Project (PRINCE). To evaluate the validity of this web-based puberty self-assessment instrument, we conducted a cross-sectional study to assess the agreement between self-assessed Tanner staging and Tanner staging assessed by a professional pediatric endocrinologist.

METHODS

Participants

Participants were consecutively recruited among outpatients in the endocrinology department of Beijing Children's Hospital during the period from October 2016 to August 2017. In a test for agreement between two raters using the kappa statistic, a sample size of 164 participants achieved 85% power to detect a true kappa value of 0.70, based on a significance level of 0.05 when the frequency of children in pubertal onset equals 65%.

Children aged 8–18 years were recruited. Based on the following criteria, we excluded children 1) who had an emergency or critical, acute disease; 2) whose comprehension was restricted due to disease or illness such that self-assessment could not be completed; 3) who were reluctant to take part in this assessment (either the children or their parents).

The study protocol was approved by the Ethics Committee of Beijing Children's Hospital, Capital Medical University, China. All participants and their parents provided their signed informed consent after receiving a detailed interpretation of the assessment made by a pediatrician.

Web-based self-assessment of pubertal development

The web-based instrument for self-assessment of pubertal development used in the study was developed based on the Tanner staging system. Participants could log in to the web-based pubertal development self-assessment module, which is a function module that is part of the electronic data capture platform for PRINCE (access: http://1.202.139.123/). The online self-assessment module provides participants with images, which have a fixed pixel size, of Tanner stages I to V, for boys and girls separately. For boys, the online image shows the five stages of pubic hair development and illustrations of testicular volume, with proportional scales and descriptions of reference objects close to the correct testicular size. For girls, images are provided the five stages of development of both pubic hair and breasts. The images are shown in Figure 1.

An appointed and trained nurse or pediatrician in the endocrinology outpatient department of Beijing Children's Hospital helped participants to log in to the website and explained the content and procedure of the self-assessment instrument. Children assessed their pubertal development by themselves or with the help of a parent or guardian. The participants were required to choose the self-perceived stage of pubertal development by comparing their development situation with the illustrations and descriptions. After completion, they could submit their assessment online.

Physical examination and pediatrician-assessed pubertal development

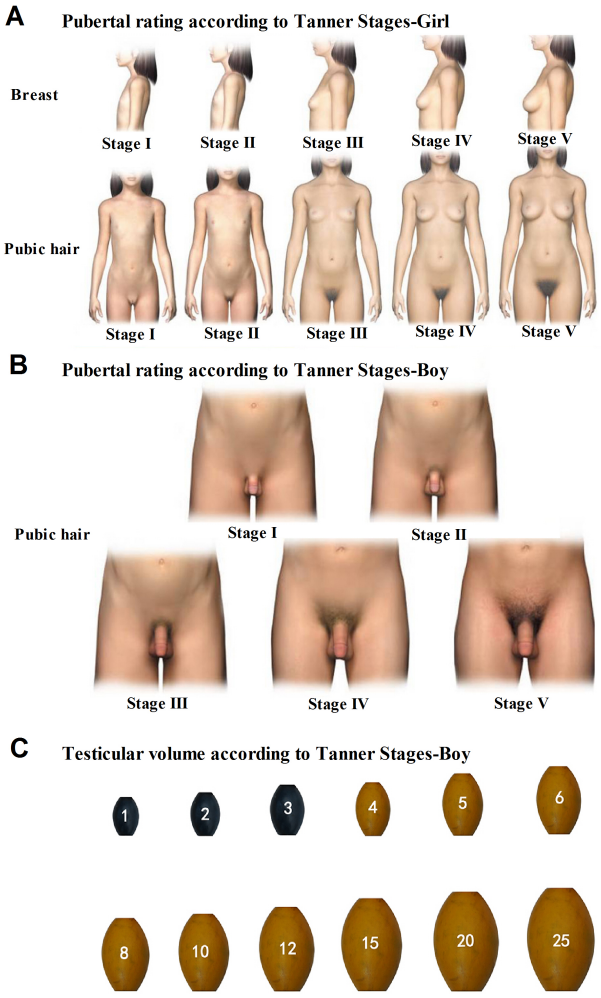

FIGURE 1 Images of Tanner stages for girls and boys in the web-based self-assessment. A, Tanner stages for girls. B, Tanner stages for boys. C, Testicular volume self-assessment illustration for boys.

One blinded senior pediatric endocrinologist assessed Tanner staging, based on examination of public hair and breast development in girls and pubic hair and testicular development in boys. Testicular volume was measured by the physician using a standard model of testicular size. Children were examined by the pediatric endocrinologist according to a standard operating procedure, focusing on consolidated rules of assessment, including differentiation of subcutaneous fat and breast tissue in overweight girls and inquiring about pubic hair removal, which could bias assessment of the Tanner stage. In situations of asymmetric breast sizes in girls or asymmetric testicular volumes in boys, the Tanner stage was defined according to the higher stage. The onset of puberty was defined as when Tanner stage of at least 2 for breast development and testicular volume of 4 mL or above. The pediatric endocrinologist was blinded to the self-assessment results and communicated their assessment to the nurse or pediatrician who had helped the participant with the web-based assessment.

Statistical analysis

Continuous variables are presented as mean and standard deviation (SD). Dichotomous and ordinal variables are presented as N (%). The agreement between self-assessed and pediatrician-assessed pubertal development status was evaluated using Spearman correlation and the kappa statistic. Weighted kappa values were computed because the outcome is an ordinal variable. For testicular volume, the paired *t*-test was used to evaluate the agreement between self-assessment and pediatrician assessment. For comparison with pediatrician-assessed pubertal onset, the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of self-assessed pubertal onset were computed. In addition, participants were analyzed by age-stratification (aged <10 years and ≥10 years) because one recent study indicated that children in southern China have puberty onset at age 9–10 years.[7] Statistical analysis was conducted using SAS 9.4 (SAS Institute Inc., Cary, NC, USA).

RESULTS

After data cleaning, a total of 174 participants, including 82 girls and 92 boys, had complete data (these 174 participants were assessed by themselves and by a pediatrician). The mean age at assessment was 12.1 (2.3) years, age 11.8 (2.2) years for girls and 12.4 (2.3) years for boys. The age and sex distribution of participants is shown in Table 1.

TABLE 1 Age distribution of the participants

Age (years)	N	Ratio of girl to boy
8	7	3:4
9	18	11:7
10	23	12:11
11	22	12:10
12	28	12:16
13	26	14:12
14	22	9:13
15–18	18	9:19
Total	174	82:92

Table 2 shows the proportion of Tanner stages among the 174 participants assessed by themselves or by a pediatrician. Spearman correlation analysis indicated that there was high agreement between self-assessment and pediatrician assessment: correlation coefficients were 0.872 for pubic hair and 0.933 for testicular volume ($P < 0.001$).

Agreement was also high between self-assessed testicular volume and pediatrician-assessed testicular volume, as shown in Figure 2. For boys, the mean testicular volume was 6.90 by self-assessment and 7.08 by pediatrician assessment, with no statistically significant difference ($t =$

TABLE 2 Agreement between web-based self-assessment and physician assessment for Tanner staging (N)

Tanner stage	Pediatrician-assessed for sexual development						Weighted Kappa Value (95% *CI*)
	I	II	III	IV	V	Total	
Girls: Tanner staging for breast development							
I	1	2	1	0	0	4	
II	3	20	12	2	0	37	
III	2	1	19	8	0	30	0.495
IV	0	0	1	7	0	8	(0.359–0.631)
V	0	0	1	1	1	3	
Total	6	23	34	18	1	82	
Girls: Tanner staging for pubic hair development							
I	26	1	2	0	0	29	
II	1	21	5	2	0	29	
III	0	0	12	3	0	15	0.785
IV	0	0	1	5	2	8	(0.690–0.880)
V	0	0	0	0	1	1	
Total	27	22	20	10	3	82	
Boys: Tanner staging for pubic hair development							
I	54	1	0	0	0	55	
II	8	11	0	0	0	19	
III	0	3	4	1	0	8	0.825
IV	0	1	2	1	0	4	(0.737–0.912)
V	0	0	0	0	6	6	
Total	62	16	6	2	6	92	

CI, confidence interval.

−1.127, $P = 0.263$).

Table 2 shows the weighted kappa values between self-assessment and pediatrician assessment for pubic hair and breast development. The agreement rate of pubertal assessment between the children and pediatrician was 58.5% for breasts and 79.3% for pubic hair in girls. For boys, the agreement rate of pubertal assessment between the children and pediatrician was 82.6% based on pubic hair. Weighted kappa values demonstrated similar agreement (49.5%, 78.5%, and 82.5% for breast and pubic hair in girls and pubic hair in boys, respectively).

Pairwise correlation analysis results between self-assessed and pediatrician-assessed pubertal development are shown in Table 3. There was intermediate or high correlation between self-assessment and pediatrician assessment for indexes of puberty assessment. However, the correlation coefficients in children aged 9 years and under were lower than those in children aged 10 years and above. Sensitivity analysis for testicular volume assessment was also conducted by omitting self-assessment of 15 mL or higher; the correlation coefficient between self-assessment and physician-assessment was 0.923 using linear Pearson correlation and 0.909 using Spearman correlation.

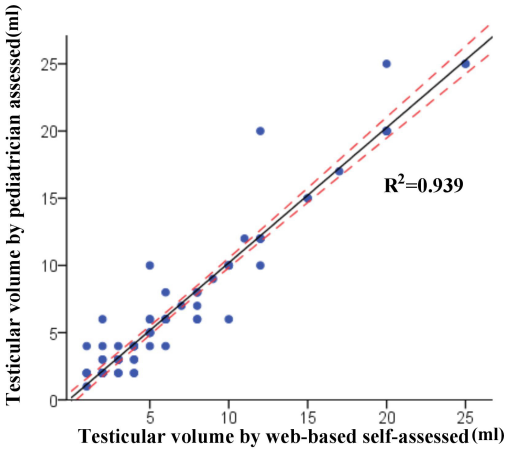

FIGURE 2 Scatter plot for assessment of testicular volume in boys, by web-based self-assessment and pediatrician assessment. The solid line showed the fitted curve of scatters and the dashed line showed the confidence interval of the estimated values. The solid points, which potentially could be overlapped by the values, represented each assessment via two methods.

TABLE 3 Correlation analysis between web-based self-assessment and pediatrician assessment for Tanner staging by age group

Characteristics	Spearman Correlation Coefficient	*P*
Girls		
Breast		
Total	0.643	<0.001
< 10 years	0.439	0.117
≥ 10 years	0.672	<0.001
Pubic hair		
Total	0.878	<0.001
< 10 years	0.782	0.001
≥ 10 years	0.848	<0.001
Boys		
Pubic Hair		
Total	0.872	<0.001
< 10 years	0.442	0.173
≥ 10 years	0.901	<0.001
Testicular volume		
Total	0.933	<0.001
< 10 years	0.579	0.062
≥ 10 years	0.946	<0.001

As for the determination of pubertal onset, the sensitivity of web-based self-assessment was more than 90% based on breast development, pubic hair, or testicular volume (Table 4). However, the specificity of web-based self-assessed breast development was lower than other indexes (17%, 95% *CI*: 8%–25%).

DISCUSSION

Reproductive hormones fluctuate during pubertal development, which can potentially influence the level of some common clinical laboratory biomarkers. For this reason, sexual maturation during this period should be considered when establishing reference intervals for children, especially for some susceptible biomarkers in adolescents.

We applied a web-based self-assessment approach, which was one of the modules in an electronic data capture platform for a national multicenter project to establish pediatric reference intervals. Using our web-based instrument, participants can conveniently evaluate Tanner stages of pubertal development online by themselves. To evaluate the validity of this web-based self-assessment approach, this study was preliminarily conducted among endocrinology outpatients of Beijing Children's Hospital. The study findings indicate clinically acceptable agreement of Tanner staging between web-based self-assessment and pediatrician assessment for participants aged 10 years or above. The observed weighted kappa values between the two assessment approaches were comparable with those in previous studies, in which pediatric assessment was regarded as the gold standard for assessment of pubertal development.[8-12] Our study showed that web-based self-assessment of Tanner stages is consistent with assessment by a pediatrician (kappa > 0.8) for boys and pubic hair development is well assessed by girls (kappa > 0.7). These findings suggest that self-assessment of pubertal development using online illustrations and descriptions of each Tanner stage for boys and girls could be a valid alternative to physician assessment, to meet the needs of epidemiologic studies for the evaluation of sexual maturation in practice. Kappa values have differed among some earlier studies, from 0.3[13] to 0.9[14, 15], with an average of about 0.5[16, 17].

In our study, the correlation coefficients of pubic hair and testicular volume for boys and pubic hair for girls were all over 0.85; correlation for breasts in girls was weaker (r = 0.6). Similar trends in sensitivity and specificity have also been observed for distinguishing pubertal maturation with Tanner staging. A recent study in Denmark reported that assessments of pubic hair development were highly correlated in both boys and girls (r = 0.70 and 0.80, respectively), as was girls' assessment of breast development (r = 0.74).[18] However, a Mexican study compared the approaches of self-assessment with physician assessment of hormone levels to determine pubertal maturation; their findings showed that breast stages in girls were well reported (r = 0.89) and superior to physician assessment (r = 0.80).[6] In our study, the lower consistency of assessment for girls' breast development differed from other studies. We found that pediatricians usually assessed at a higher stage of breast development than girls themselves (Table 2). This might be associated with limitations of the web-based instrument itself, in which online images alone might not adequately describe the key differences between Tanner breast stages. Although a nurse or physician provided some instruction and helped participants with the self-assessment, different abilities of comprehension and discernment among participants to make assessments using online images might lead to the observed variation. In addition, the obesity, especially adipose tissue in the chest area, might impact the accuracy of breast development self-assessment. The difference in assessment procedures among studies might affect the agreement between the two assessment approaches. Firstly, there might be little time for self assessment in the outpatient department such that the validity of self-assessment would be underestimated. The kappa value would be improved if self-assessment was planned with sufficient time provided before physician assessment.[5, 19] Secondly, the assessment environment could affect the procedure owing to the psychological and emotional situation of participants. For example, agreement would be underestimated if there were a lack of privacy during

TABLE 4 Sensitivity, specificity, positive predictive value and negative predictive value of self-assessed pubertal development in age-stratified analysis

Self-assessed(Age stratification)	Pediatrician-assessed			Sensitivity	Specificity	PPV	NPV
	Yes	No	Total				
Age < 10 years							
Girls: Tanner stage ≥ 2 for breast development							
Yes	13	0	13	0.93 (0.64–0.99)	–	1.00 (0.72–1.00)	0.00 (0–0.95)
No	1	0	1				
Total	14	0	14				
Girls: Tanner stage ≥ 2 for pubic hair development							
Yes	2	1	3	1.00 (0.20–1.00)	0.92 (0.60–0.99)	0.67 (0.13–0.98)	1.00 (0.68–1.00)
No	0	11	11				
Total	2	12	14				
Boys: Tanner stage ≥ 2 for pubic hair development							
Yes	1	1	2	0.50 (0.03–0.97)	0.89 (0.51–0.99)	0.50 (0.03–0.97)	0.89 (0.51–0.99)
No	1	8	9				
Total	2	9	11				
Boys: Tanner stage ≥ 4 for testicular volume development							
Yes	2	1	3	0.67 (0.13–0.98)	0.88 (0.47–0.99)	0.67 (0.13–0.98)	0.88 (0.47–0.99)
No	1	7	8				
Total	3	8	11				
Age ≥ 10 years							
Girls: Tanner stage ≥ 2 for breast development							
Yes	60	5	65	0.97 (0.88–0.99)	0.17 (0.01–0.64)	0.92 (0.82–0.97)	0.33 (0.02–0.87)
No	2	1	3				
Total	62	6	68				
Girls: Tanner stage ≥ 2 for pubic hair development							
Yes	50	0	50	0.94 (0.83–0.99)	1.00 (0.75–1.00)	1.00 (0.91–1.00)	0.83 (0.58–0.96)
No	3	15	18				
Total	53	15	68				
Boys: Tanner stage ≥ 2 for pubic hair development							
Yes	28	7	35	0.97 (0.80–0.99)	0.87 (0.74–0.94)	0.80 (0.63–0.91)	0.98 (0.87–0.99)
No	1	45	46				
Total	29	52	81				
Boys: Tanner stage ≥ 4 for testicular volume development							
Yes	51	3	54	0.94 (0.84–0.99)	0.88 (0.68–0.97)	0.94 (0.84–0.99)	0.88 (0.68–0.97)
No	3	22	25				
Total	54	25	79				

PPV, positive predictive value; NPV, negative predictive value.

the assessment procedure.[10] Enough time should be given to children for self-assessment to obtain better agreement with pediatrician assessment. Another way to improve the assessment environment is to provide participants with a mirror for more accurate assessment.[19] In addition, using the web-based platform on a tablet or mobile phone could solve the above problem, making it possible for participants to conduct self-assessment in their own home at any time.

Pediatrician assessment of Tanner stages is regarded as the gold standard for evaluation of pubertal maturation in clinical practice, under the assumption that it is free of error. In fact, like self-assessment, pediatrician assessment may also be influenced by subjectivity. There is evidence that assessment by a pediatrician is subject to considerable within-observer and between-observer variability.[6] Some studies have demonstrated that the reliability of pediatrician assessment is overestimated, especially for assessment of testicular volume,[20, 21] and breast development,[22] and that the validity is highly depended on appropriate training.[23, 24]

The comprehensive ability and cognitive competence of young participants will affect the validity of any self-assessment procedure. From subgroup analysis by age (Table 3), we observed that for participants younger than 10 years of age, the correlation coefficients between the two approaches were lower than for children aged 10 years or above, in both boys and girls. It has been reported that girls with excessive BMI have a significantly higher prevalence of thelarche from ages 8.0 to 9.6 years and pubarche from ages 8.0 to 10.2 years.[25] A Chinese population study obtained similar results,[26] indicating an earlier median age of thelarche and pubarche onset among girls with central obesity. For overweight or obese girls, the presence of excess fat tissue may result in inaccurate assessment of breast Tanner stage in both self-assessed and pediatrician assessed procedures.[27-29] Errors in self-assessment among younger girls and boys are likely dependent on weight, height stature, and other factors.[9, 10, 18] However, difference in pediatrician assessment may depend on difference in experience and standardization.[6]

The present web-based self-assessment module was mainly intended to assess the stage of puberty among adolescent candidates for PRINCE. The Canadian Laboratory Initiative in Pediatric Reference Intervals (CALIPER) has recommended that age and sex partitions should be required, to cover changes in adolescence.[30] Our previous study[31] also revealed that some factors, such as red blood cells and hemoglobin, are distributed differently by age, especially when partitioned at 14 years. To establish the pediatric reference interval of common laboratory items in a scientific manner, Tanner staging would be helpful in deciding the age partition and analyzing differences during this special period among boys and girls. This web-based self-assessment approach makes epidemiologic investigation more feasible and acceptable for reference individuals. In the pilot stage of the PRINCE project, 81.2% (216/266) of eligible individuals (age ≥ 8 years) completed the web-based self-assessment. The web-based self-assessment was accepted by children, their parents, pediatricians, and investigators at each center.

A limitation of the present study is that most participants were hospital-based. This might introduce potential selection bias as most enrolled patients asked for pubertal assessment because of potential abnormal Tanner staging. However, this study aimed to validate the consistency between the two approaches. Another limitation was the lack of a "gold standard" to compare with the two assessment approaches, although all pediatrician assessment was carried out by one senior endocrinologist. As mentioned, the relative consistency between web-based self-assessment and pediatrician assessment could be determined, with the assumption that judgment of Tanner staging by a pediatrician is accurate. Another important limitation was regarding the web-based assessment instrument itself. The images of pubic hair, as well as those for development stages of breasts (for girls) and genitals (for boys) were frontal images, which might affect comprehension and discernment for accurate assessment. The difference between stages of testicular volume was not obvious in the images, which may lead to visual misjudgment. BMI data were not collected in this study, which restricted further analysis of the impact of BMI on agreement between participant and physician ratings. In addition, the participants received instruction from nurses when taking self-assessment in this study, so their results might be more accurate than those of individuals who perform self-assessment without face-to-face instruction, such as participants in a large epidemiologic study. Greater attention should be focused on this issue when using the web-based self-assessment in large population study, to make pubertal development assessment feasible. All of the above are potential factors to consider when applying this web-based approach in future epidemiologic studies, as well as providing detailed instruction for participants and, clear images with no interfering elements, the need for textual presentation, and the multimedia approach taken, to improve applicability of the web-based self-assessment instrument.

The web-based instrument for self-assessment of pubertal development in this study is easy for children to grasp intuitively, making self-assessment of puberty more convenient for wider use, by protecting children's privacy. However, self-assessment by children less than 10 years old could be less accurate, especially for girls' assessment of breast development. Therefore, self-assessment of pubertal development, especially for breasts, should be interpreted cautiously.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

REFERENCES

1. Tahmasebi H, Higgins V, Fung AWS, Truong D, White-Al Habeeb NMA, Adeli K. Pediatric reference intervals for biochemical markers: Gaps and challenges, recent national initiatives and future perspectives. *EJIFCC*. 2017;28:43-63.
2. IFCC, CLSI, EP28-A3C document, Defining, Establishing

and Verifying Reference Intervals in the Clinical Laboratory: Approved Guideline, third ed., 28, 2010 (No. 30).
3. Zierk J, Arzideh F, Rechenauer T, et al. Age- and sex-specific dynamics in 22 hematologic and biochemical analytes from birth to adolescence. *Clin Chem.* 2015;61:964-973.
4. Shaw JLV, Cohen A, Konforte D, Binesh-Marvasti T, Colantonio DA, Adeli K. Validity of establishing pediatric reference intervals based on hospital patient data: a comparison of the modified Hoffmann approach to CALIPER reference intervals obtained in healthy children. *Clin Biochem.* 2014;47:166-172.
5. Norris SA, Richter LM. Usefulness and reliability of Tanner pubertal self-rating to urban black adolescents in South Africa. *J Res Adolesc.* 2005;15:609-624.
6. Chavarro JE, Watkins DJ, Afeiche MC, et al. Validity of self-assessed sexual maturation against physician assessments and hormone levels. *J Pediatr.* 2017;186:172-178.
7. Li YH, Ma HM, Chen HS, Su Z, Gu YF, Du ML. Longitudinal study of the pattern of pubertal development in Cantonese schoolgirls. *Chin J Pediatr.* 2009;47:410-415. (In Chinese)
8. Brooks-Gunn J, Warren MP, Rosso J, Gargiulo J. Validity of self-report measures of girls' pubertal status. *Child Dev.* 1987;58:829-841.
9. Jaruratanasirikul S, Kreetapirom P, Tassanakijpanich N, Sriplung H. Reliability of pubertal maturation self-assessment in a school-based survey. *J Pediatr Endocrinol Metab.* 2015;28:367-374.
10. Schlossberger NM, Turner RA, Irwin CE Jr. Validity of self-report of pubertal maturation in early adolescents. *J Adolesc Health.* 1992;13:109-113.
11. Chan NP, Sung RY, Nelson EA, So HK, Tse YK, Kong AP. Measurement of pubertal status with a Chinese self-report pubertal development scale. *Matern Child Health J.* 2010;14:466-473.
12. Li D, Shi HJ, Zhang Y, Wang W, Tan H, Wang ZW. Applicability of self-rated pubertal development scale among urban Chinese adolescents. *Chin J Epidemiol.* 2012;33:580-583. (In Chinese)
13. Morris NM, Udry JR. Validation of a self-administered instrument to assess stage of adolescent development. *J Youth Adolesc.* 1980;9:271-280.
14. Boas SR, Falsetti D, Murphy TD, Orenstein DM. Validity of self-assessment of sexual maturation in adolescent male patients with cystic fibrosis. *J Adolesc Health.* 1995;17:42-45.
15. Williams RL, Cheyne KL, Houtkooper LK, Lohman TG. Adolescent self-assessment of sexual maturation. Effects of fatness classification and actual sexual maturation stage. *J Adolesc Health Care.* 1988;9:480-482.
16. Coleman L, Coleman J. The measurement of puberty: A review. *J Adoles.* 2002;25:535-550.
17. Desmangles JC, Lappe JM, Lipaczewski G, Haynatzki G. Accuracy of pubertal Tanner staging self-reporting. *J Pediatr Endocrinol Metab.* 2006;19:213-221.
18. Rasmussen AR, Wohlfahrt-Veje C, Tefre de Renzy-Martin K, et al. Validity of self-assessment of pubertal maturation. *Pediatrics.* 2015;135:86-93.
19. Wacharasindhu S, Pri-Ngam P, Kongchonrak T. Self-assessment of sexual maturation in Thai children by Tanner photograph. *J Med Assoc Thai.* 2002;85:308-319.
20. Carlsen E, Andersen AG, Buchreitz L, et al. Inter-observer variation in the results of the clinical andrological examination including estimation of testicular size. *Int J Androl.* 2000;23:248-253.
21. Tatsunami S, Matsumiya K, Tsujimura A, et al. Inter/intra investigator variation in orchidometric measurements of testicular volume by ten investigators from five institutions. *Asian J Androl.* 2006;8:373-378.
22. Hergenroeder AC, Hill RB, Wong WW, Sangi-Haghpeykar H, Taylor W. Validity of self-assessment of pubertal maturation in African American and European American adolescents. *J Adolesc Health.* 1999;24:201-205.
23. Slora EJ, Bocian AB, Herman-Giddens ME, et al. Assessing inter-rater reliability (IRR) of Tanner staging and orchidometer use with boys: a study from PROS. *J Pediatr Endocrinol Metab.* 2009;22:291-299.
24. Slough JM, Hennrikus W, Chang Y. Reliability of Tanner staging performed by orthopedic sports medicine surgeons. *Med Sci Sports Exerc.* 2013;45:1229-1234.
25. Rosenfield RL, Lipton RB, Drum ML. Thelarche, pubarche, and menarche attainment in children with normal and elevated body mass index. *Pediatrics.* 2009;123:84-88.
26. Chen C, Zhang Y, Sun W, et al. Investigating the relationship between precocious puberty and obesity: a cross-sectional study in Shanghai, China. *BMJ Open.* 2017;7:e014004.
27. Bonat S, Pathomvanich A, Keil MF, Field AE, Yanovski JA. Self-assessment of pubertal stage in overweight children. *Pediatrics.* 2002;110:743-747.
28. Lee K, Valeria B, Kochman C, Lenders CM. Self-assessment of height, weight, and sexual maturation: validity in overweight children and adolescents. *J Adolesc Health.* 2006;39:346-352.
29. Sun Y, Tao FB, Su PY. Self-assessment of pubertal Tanner stage by realistic colour images in representative Chinese obese and non-obese children and adolescents. *Acta Paediatr.* 2012;101:e163-166.
30. Konforte D, Shea JL, Kyriakopoulou L, et al. Complex biological pattern of fertility hormones in children and adolescents: a study of healthy children from the CALIPER cohort and establishment of pediatric reference intervals. *Clin Chem.* 2013;59:1215-1227.
31. Lv Y, Feng G, Ni X, Song W, Peng X. The critical gap for pediatric reference intervals of complete blood count in China. *Clin Chim Acta.* 2017;469:22-25.

How to cite this article: Peng X, Peng Y, Li Y, Nie X, Gong C, Wu D, Ni X. Validity of web-based self-assessment of pubertal development against pediatrician assessments. *Pediatr Invest.* 2018;2:141-148. https://doi.org/10.1002/ped4.12050