Original article

Arch Argent Pediatr 2023;121(2):e202202717

Tanner stages and Prader orchidometry in male adolescents. A descriptive, cross-sectional study

Gonzalo Agüero[a], *Enrique Berner*[a]

ABSTRACT

Introduction. The Tanner scale and the Prader orchidometer are the instruments most commonly used to assess pubertal development in children. The assessment of puberty in the clinic is only useful if recent and reliable references in the same population are available for comparison. Objective: to assess the correlation between Tanner stages and testicular volume (TV) in Argentine adolescents.

Population and methods. Study with a descriptive, cross-sectional design conducted in healthy boys aged 9–20 years. Male children and adolescents with urogenital conditions and disorders affecting testicular growth were excluded. The correlation between Tanner stages and TV was assessed using non-parametric tests.

Results. A total of 367 male adolescents with an average age of 13.8 ± 2.5 years were assessed. TV increased in correlation to Tanner stages (Spearman: 0.943, $p < 0.001$) with significantly different volumes, except in the early genital 1-2 stages (p 0.343) and pubic hair 1-2 stages (p 0.447). Among peripubertal boys, 16% (95% confidence interval: 9.6–24.4%, n = 17/106) were wrongly classified based on Tanner stages.

Conclusions. During male puberty, TV increased in correlation to Tanner stages, but no significant differences were observed between Tanner stages 1 and 2. Using the Prader orchidometer is critical to establish the onset of puberty in boys.

Key words: *puberty; adolescent; adolescent growth; adolescent medicine; testis.*

doi: http://dx.doi.org/10.5546/aap.2022-02717.eng

To cite: Agüero G, Berner E. Tanner stages and Prader orchidometry in male adolescents. A descriptive, cross-sectional study. *Arch Argent Pediatr* 2023;121(2):e202202717.

[a] *Department of Adolescence, Hospital General de Agudos Dr. Cosme Argerich, City of Buenos Aires, Argentina.*

Correspondence to *Gonzalo Agüero: aguerogonzalo@gmail.com*

Funding*: Annual research fellowship granted by the* Sociedad Argentina de Pediatría *(period: June 2018–May 2019). The protocol continued until February 2020.*

Conflict of interest: *None.*

Received*: 5-13-2022*
Accepted*: 7-21-2022*

INTRODUCTION

The assessment of male puberty is based primarily on the examination of the external genitalia.[1,2] The gold standard method for assessing pubertal development is the Tanner scale,[3] while that for the measurement of testicular size is the Prader orchidometry;[4] both methods must be performed by adequately trained health care providers.[2,5]

Due to the great variability in the onset, speed, and magnitude of pubertal changes, reference values for pubertal developmental stages and testicular volume (TV) obtained solely on the basis of age vary very widely.[6] For example, healthy boys between 13 and 14 years of age could be in almost any stage of genital development (G1 to G5) and have a TV between prepubertal and adult (1–25 mL).[3,6,7]

In Argentina, to date, there are no updated publications on testicular growth in the context of male puberty. We therefore decided to assess the relationship between Tanner stages and TV in healthy male adolescents, as they are used as developmental indicators. A developmental indicator[8,9] is any measurable milestone associated with the biological process of development. It must be universal, appear sequentially and in the same order in all children, and have the ability to discriminate between different stages of maturation. It should be easy to obtain and measure, with minimally invasive, robust, and sensitive techniques that can be adapted to different situations (clinical care, research, different observers). In this regard, the most accessible developmental indicators are those that are evident in the body's anatomy.

OBJECTIVE

To describe the correlation between Tanner stages and TV as measured with the Prader orchidometer in Argentine adolescents.

POPULATION AND METHODS

This was a descriptive, cross-sectional study conducted between June 1st, 2018 and February 28th, 2020 at the Department of Adolescence of a public hospital in the City of Buenos Aires. The target population was male adolescents. The study was approved by the hospital's Research Ethics Committee. Parents and adolescents older than 14 years were asked to sign an informed consent. Participation was voluntary, anonymous, and confidential, and maintaining the adolescents' privacy.

Inclusion criteria: Argentine male adolescents aged 9 to 20 years who attended the hospital for a health check-up. Exclusion criteria: Pubertal alterations, hypogonadism, congenital urogenital anomalies, scrotal abnormalities, chronic conditions, ongoing acute illness, and refusal to participate in the study.

All adolescents were assessed by the principal investigator in a warm setting (≥ 20 °C) to avoid the cremasteric reflex. Firstly, the Tanner stage was established. Secondly, the scrotal contents were examined by palpation to rule out urogenital conditions. Finally, TV was assessed starting with the right testis and using the same orchidometer for all measurements. To achieve a reliable measurement, the surrounding scrotal skin was reduced as much as possible without compressing or deforming the testis. The statistical analysis was blinded to make up for the lack of blinding in the TV measurements.

Instruments

Tanner scale:[3] It allows to visually determine morphological changes in external genitalia (G stages) and pubic hair (PH stages), classifying them into 5 stages. Stage 1 corresponds to prepuberty; stages 2–4, puberty in progress; and stage 5, full development.

Prader orchidometer:[4] It allows to make a testicular measurement by comparative palpation with an ellipsoid model of known volume (1–25 mL). TV ≥ 4 mL indicated the onset of puberty. When testicular size was between 2 ellipsoids, the intermediate volume was estimated (e.g., between 12 mL and 15 mL, volume was 13.5 mL); if > 25 mL, the volume was recorded at 25 mL.

Sampling. Assuming that up to 19 years of age, TV in males has a standard deviation of 8 mL,[10] and seeking a 95% confidence level with an estimation accuracy of 4 mL, the sample size was established at 50 boys per Tanner stage, for a total of 250 male adolescents. Patients were included sequentially by simple random sampling.

Statistical analysis. Since the TV did not adjust to a normal distribution (Kolmogorov-Smirnov test), quantitative variables were described using median and interquartile range values; and frequencies, using a 95% confidence interval (CI). The correlation was assessed using Spearman's coefficient. The comparisons were performed using the Wilcoxon signed-rank test (right/left testis) and the Kruskal-Wallis test (TV/Tanner stage). The finding of testicular asymmetry was assessed based on the

asymmetry percentage formula: [(right TV - left TV)/right TV] × 100]. The significance level was established at $p < 0.05$. Data were processed using the SPSS Statistics® 26 software (IBM, 2019).

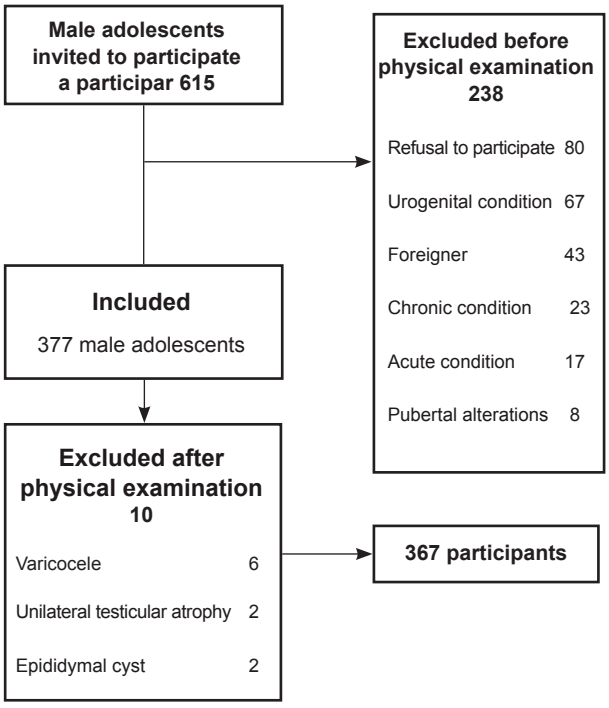
RESULTS

A total of 615 boys were invited to participate; 238 were excluded based on their history and 10, based on their physical examination (*Figure 1*). Finally, 367 healthy male adolescents with an average age of 13.8 ± 2.5 years were assessed.

A high correlation was observed between both testes and G stages (Spearman: 0.943, $p < 0.001$); also between PH stages and the right (Spearman: 0.893, $p < 0.001$) and left (Spearman: 0.889, $p < 0.001$) testes.

Testicular asymmetry was found in 82/367 adolescents (22.3%, 95% CI: 18.2–27%) (*Table 1*). The percentage of asymmetry was ≤ 20% in 68/367 adolescents (18.5%, 95% CI: 14.7–23%); in those patients in whom the percentage of asymmetry was > 20%, the difference was not clinically significant (e.g.: RT

Figure 1. Flowchart of adolescents in the study



The most common urogenital condition was varicocele, with a total of 24 cases. Male adolescents with a urogenital condition were referred for medical follow-up.

Table 1. Comparison between left and right testicular volumes in 367 male adolescents

Left testis (LT) – Right testis (RT).	Negative ranks (LT < RT)	74
	Positive ranks (LT > RT)	8
Wilcoxon signed-rank test (Z -6.803, $p < 0.001$).	Matched (LT = RT)	285
	Total	367

Comparison of median left and right testicular volumes. Negative ranks indicate that the right testes were larger. Positive ranks indicate that the left testes were larger.

8 mL and LT 6 mL, or RT 8 mL and LT 10 mL). In 10/367 postpubertal boys (2.7%, 95% CI: 1.3–5%), testes with a volume > 25 mL were found. In 7/367 peripubertal boys (1.9%, 95% CI: 0.8–3.9%), pubic hair with PH2 characteristics (thin, long, and pigmented) was observed, but with predominantly scrotal distribution and not at the base of the penis, as in the original description.[3] These patients were classified as PH2.

The Kruskal-Wallis test showed that the distribution of testicular size was different across the Tanner stages, except for G1-G2 and PH1-PH2 (*Table 2*). Among the 106 peripubertal boys in G1-G2 stages, 5/106 (4.7%, 95% CI: 1.6–10.7%) were classified as prepubertal (G1) although their TV was ≥ 4 mL, and 12/106 (11.7%, 95% CI: 6–18.9%) were classified as pubertal (G2) although their TV was < 4 mL.

Table 3 describes TV corresponding to the different pubertal stages.

DISCUSSION

The Tanner scale and the Prader orchidometry are the gold standard methods for the assessment of male puberty in the absence of other easily objectifiable pubertal milestones; however, it is common to find different methodological approaches that hinder the comparison between studies.[2,5,9,11]

Most population-based studies on male puberty described the pubertal timing or the secular trend based on the description of pubertal events according to age of onset.[7,10,12–23] However, given the great variability in the onset and end of puberty, it is necessary to assess certain developmental indicators according to the Tanner stages, as in the case of hypothalamic-pituitary-gonadal axis hormones[24] and, most likely, orchidometry.

Few studies have analyzed TV measured with an orchidometer based on Tanner stages. Joustra et al.,[21] assessed 769 boys and developed reference tables for ultrasound and orchidometric TV based on age. In their analysis, only PH stages were used and right and left TV were averaged. According to published graphs, the distribution of testicular size is excessively ample (1–25 mL), particularly in PH1, PH2,

Table 2. Distribution of testicular volume across Tanner stages in 367 male adolescents. Kruskal-Wallis test for independent samples

Right testicular volume (mL)			Left testicular volume (mL)		
Sample 1 – Sample 2	**Statistics**	**Significance**	**Sample 1 – Sample 2**	**Statistics**	**Significance**
G1 – G2	-43.440	0.343	G1 – G2	-44.516	0.303
G1 – G3	-109.167	< 0.001	G1 – G3	-109.291	< 0.001
G1 – G4	-192.960	< 0.001	G1 – G4	-188.550	< 0.001
G1 – G5	-266.827	< 0.001	G1 – G5	-268.372	< 0.001
G2 – G3	-65.727	0.006	G2 – G3	-64.775	0.007
G2 – G4	-149.520	< 0.001	G2 – G4	-144.034	< 0.001
G2 – G5	-223.387	< 0.001	G2 – G5	-223.856	< 0.001
G3 – G4	-83.793	< 0.001	G3 – G4	-79.259	< 0.001
G3 – G5	-157.660	< 0.001	G3 – G5	-159.081	< 0.001
G4 – G5	-73.867	< 0.001	G4 – G5	-79.822	< 0.001
Sample 1 – Sample 2	**Statistics**	**Significance**	**Sample 1 – Sample 2**	**Statistics**	**Significance**
PH1 – PH2	-40.809	0.385	PH1 – PH2	-39.624	0.447
PH1 – PH3	-111.116	< 0.001	PH1 – PH3	-111.165	< 0.001
PH1 – PH4	-171.758	< 0.001	PH1 – PH4	-170.066	< 0.001
PH1 – PH5	-245.298	< 0.001	PH1 – PH5	-244.454	< 0.001
PH2 – PH3	-70.307	0.012	PH2 – PH3	-71.541	0.010
PH2 – PH4	-130.949	< 0.001	PH2 – PH4	-130.442	< 0.001
PH2 – PH5	-204.489	< 0.001	PH2 – PH5	-204.830	< 0.001
PH3 – PH4	-60.642	0.017	PH3 – PH4	-58.901	0.023
PH3 – PH5	-134.182	< 0.001	PH3 – PH5	-133.289	< 0.001
PH4 – PH5	-73.540	< 0.001	PH4 – PH5	-74.388	< 0.001

Each row tests the null hypothesis that the distributions of testicular volume between Tanner stages (Sample 1 – Sample 2) are equal. Asymptotic significance values are shown (bilateral tests). The significance level was 0.05. Significance values were adjusted applying the Bonferroni correction for multiple tests. The median differences between G1-G2 and PH1-PH2 were not statistically significant.

Table 3. Testicular volume (median and interquartile range) according to Tanner stages in 367 male adolescents

Tanner genital stage (G)	TV (mL)	Median (IQR)	Tanner pubic hair stage (PH)	TV (mL)	Median (IQR)
G1	RT	2 (1.5–3)	**PH1**	RT	2 (2–4)
	LT	2 (1.5–2.1)		LT	2 (2–4)
G2	RT	5 (4–6)	**PH2**	RT	6 (3–8)
	LT	4 (3.1–6)		LT	6 (3–6.8)
G3	RT	10 (8–10)	**PH3**	RT	10 (8–10)
	LT	10 (8–10)		LT	10 (8–10)
G4	RT	13.5 (12–15)	**PH4**	RT	13.5 (12–15)
	LT	12 (12–15)		LT	12 (13.5–15)
G5	RT	20 (15–25)	**PH5**	RT	20 (15–22.5)
	LT	20 (15–21.9)		LT	17.5 (12–20)

TV: testicular volume. IQR: interquartile range. LT: left testis. RT: right testis.
The differences in TV between the genital and pubic hair stages are due to the fact that the synchrony between these events is not total during puberty. In addition, the asymmetry between the right and left testes is frequently observed during pubertal testicular growth.

and PH3 stages. Such large variability could be due to population differences or an inadequate orchidometric technique. Babani et al.,[25] assessed the average TV corresponding to the G and PH stages in 525 boys aged 10 to 18 years. In their study, the TV in the G1 and PH1 (prepubertal) stages was > 4 mL (pubertal), and this may also be analyzed as a population difference or an inadequate orchidometric technique. In Argentina, Bianculli and Bergadá[26] published the first study on normal male puberty in 1977. The description of TV and the degree of pubertal development were incorporated into the "Comprehensive Care of Adolescents and Young Adults" guidelines, published by the *Sociedad Argentina de Pediatría* (1990),[27] although these data were not used again.

In our study, G and PH stages were classified separately, as has been previously recommended,[3,6,7] and then the volume of both testes was measured without averaging. In the boys studied, TV increased in correlation with Tanner stages. Each Tanner stage showed a characteristic and statistically significant TV, except for G1-G2 and PH1-PH2, which showed differences that were not significant. This may be due to the low variability in peripubertal testicular sizes (1–6 mL) and errors in pubertal staging prior to orchidometry.

Largo and Prader demonstrated that an increased TV is the most reliable indicator of pubertal onset compared to visual inspection.[7] Therefore, if we rely solely on Tanner stages, a percentage of prepubertal boys will be classified as pubertal and vice versa. In our study, 16% of peripubertal boys were wrongly classified on the basis of inspection. Mul[12] stated that it would be appropriate to redefine the G2 stage, as it leaves much room for confusion: "the scrotum and testes have enlarged and there is a change in the texture of the scrotal skin. There is also some reddening of the scrotal skin..." The original description is not pertinent to the question of which of the three criteria is most relevant, and to whether all criteria have to be met or at least one or two of them. In addition, it does not explicitly state the minimum TV necessary for the genital stage to be labeled as G2. Consistent with this perspective and based on our results, we propose to redefine the G2 stage: genital stage with a TV between 4 and 6 mL, with or without minimal changes in scrotal appearance (texture and color). Here we focus on the minimum testicular size to consider the onset of puberty (4 mL) and also on determining the growth immediately after this (5–6 mL). These changes in TV result from gonarche, while the initial changes in scrotal texture and color may be influenced by other factors (room temperature and light, skin type, and pigmentation).

Some authors, such as Tomova[16] and Wang,[20] assessed puberty by measuring the testes and penis (length and diameter) because this technique is more objective compared to the Tanner G stage. Although subjectivity plays an important role in classifying the Tanner G stage, an advantage is that it is done by inspection. In a previous article, we found an excellent consistence between the Tanner stages and the Prader orchidometry among different

trained observers.[28] While penile measurement considerably increases the assessment time and the manipulation of the genitalia, with a questionable clinical performance in adolescents, an orchidometry takes less time and has a greater clinical benefit when assessing the impact of different testicular disorders, such as varicocele[29] or Klinefelter syndrome.[30]

Testicular asymmetry was observed in 22% of the participants. No ancillary tests were required because it was not associated with urogenital conditions, and this finding was deemed as asynchronous testicular growth during puberty. Vaganée et al.,[31] previously reported that testicular asymmetry is common in healthy adolescents without varicocele.

Implications for practice

This study provides updated TV values in relation to the Tanner stages for our country. Taking into account the limits of pubertal onset (9–14 years old), it is convenient to measure both testes and assess their volume according to Tanner G and PH stages because they increase correspondingly. This will allow to monitor testicular growth and detect any alteration (micro- or macroorchidism).

Since a genital examination is not sufficient to discriminate between G1 and G2, it is critical to use the Prader orchidometer to determine the onset of puberty. In this regard, we propose to redefine the G2 stage, which could reduce diagnostic error in peripubertal boys, either in the pediatric office or in clinical research. We also describe a possible variant of the PH2 stage (same characteristics, but with scrotal distribution); if this finding is confirmed in other studies, its correct definition should be assessed.

A small percentage of healthy postpubertal boys have a TV > 25 mL and cannot be accurately assessed with the Prader orchidometer.[21,28] Similarly, testicular asymmetry may be found in boys without urogenital conditions.[31] This difference should not exceed 1 orchidometer ellipsoid, particularly beyond 6 mL, when orchidometer volumes become discontinuous.

Limitations

This study had a cross-sectional design; and participants were selected from a public hospital in the City of Buenos Aires. Caution is advised when extrapolating these results. It is necessary to compare these findings with large-scale longitudinal studies.

CONCLUSIONS

During male puberty, testicular volume increases in correlation with Tanner stages. Each pubertal stage has a characteristic testicular volume, although this difference is not statistically significant in the early stages (G1-G2 and PH1-PH2). It is critical to use the Prader orchidometer to detect pubertal onset in male adolescents. ■

Acknowledgments

We would like to thank everyone working at the hospital's Department of Adolescence.

REFERENCES

1. Abreu A, Kaiser U. Pubertal development and regulation. *Lancet Diabetes Endocrinol*. 2016; 4(3):254-64.
2. Dorn L, Dahl R, Woodward H, Biro F. Defining the boundaries of early adolescence: a user's guide to assessing pubertal status and pubertal timing in research with adolescents. *Appl Dev Sci*. 2006; 10(1):30-56.
3. Marshall W, Tanner J. Variations in the pattern of pubertal changes in boys. *Arch Dis Child*. 1970; 45(239):13-23.
4. Prader A. Orchiometry. In: Donald RA (ed). Endocrine Disorders: a guide to diagnosis. New York: Marcel Dekker; 1984.Pages.33-6.
5. Dorn L, Biro F. Puberty and its measurement: a decade review. *J Res Adolesc*. 2011; 21(1):180-95.
6. Sotos J, Tokar N. Appraisal of testicular volumes: volumes matching ultrasound values references to stages of genital development. *Int J Pediatr Endocrinol*. 2017; 2017:7.
7. Largo R, Prader A. Pubertal development in Swiss boys. *Helv Paediatr Acta*. 1983; 38(3):211-28.
8. Rockett J, Lynch C, Buck G. Biomarkers for assessing reproductive development and health: Part 1 – Pubertal development. *Environ Health Perspect*. 2004; 12(1):105-12.
9. Cameron N. Assessment of maturation. In Cameron N, Bogin B (eds). Human Growth and Development. 2nd ed. London: Elsevier; 2012.Pages.515-35.
10. Goede J, Hack W, Sijstermans K, Voort Doedens L, et al. Normative values for testicular volume measured by ultrasonography in a normal population from infancy to adolescence. *Horm Res Paediatr*. 2011; 76(1):56-64.
11. Deardorff J, Hoyt L, Carter R, Shirtcliff E. Next steps in puberty research: broadening the lens toward understudied populations. *J Res Adolesc*. 2019; 29(1):133-54.
12. Mul D, Fredriks M, van Buuren S, Oostdijk W, et al. Pubertal development in the Netherlands 1965-1997. *Pediatr Res*. 2001; 50(4):479-86.
13. Juul A, Teilmann G, Scheike T, Hertel N, et al. Pubertal development in Danish children: comparison of recent European and US data. *Int J Androl*. 2006; 29(1):247-55.
14. Bundak R, Darendeliler F, Gunoz H, Bas F, et al. Analysis of puberty and pubertal growth in healthy boys. *Eur J Pediatr*. 2007; 166(6):595-600.
15. García Cuartero B, González Vergaz A, Frías García E, Arana Cañete C, et al. Valoración de la tendencia secular de la pubertad en niños y niñas. *An Pediatr (Barc)*. 2010; 73(6):320-6.
16. Tomova A, Deepinder F, Robeva R, Lalabonova H, et al. Growth and development of male external genitalia: a cross-sectional study of 6200 males aged 0 to 19 years. *Arch Pediatr Adolesc Med*. 2010; 164(12):1152-7.
17. Sun Y, Tao F, Su PY, China Puberty Research Collaboration. National estimates of pubertal milestones among urban and

rural Chinese boys. *Ann Hum Biol*. 2012; 39(6):461-7.
18. Papadimitriou A, Douros K, Kleanthous K, Papadimitriou D, et al. Pubertal maturation of contemporary Greek boys: no evidence of secular trends. *J Adol Health*. 2011; 49/4):434-6.
19. Gaete X, Garcia R, Riquelme J, Codner E. La pubertad en niños chilenos muestra un adelantamiento en el inicio del crecimiento testicular. *Rev Med Chile*. 2015; 143(3):297-303.
20. Wang Y, Zeng Q, Xiong F, Zeng Y. Male externa genitalia growth curves and charts for children and adolescents aged 0 to 17 years in Chongqing, China. *Asian J Androl.* 2018; 20(6):567-71.
21. Joustra S, van der Plas E, Goede J, Oostdijk W, et al. New reference charts for testicular volume in Dutch children and adolescent allow the calculation of standard deviation scores. *Acta Paediatr.* 2015; 104(6):e271-8.
22. Oehme NHB, Roelants M, Særvold Bruserud I, Madsen A, et al. Reference data for testicular volume measured with ultrasound and pubic hair in Norwegian boys are comparable with Northern European populations. *Acta Paediatr*. 2020; 109(8):1612-9.
23. Liu C, Liu X, Zhang X, Yang B, et al. Referential values of testicular volume measured by ultrasonography in normal children and adolescents: Z-score establishment. *Front Pediatr*. 2021; 9:648711.
24. Rey RA. Biomarcadores de hipogonadismo masculino en la infancia y la adolescencia. *Adv Lab Med*. 2020; 1(2):20190043.
25. Babani T, Goyal D, Sareen D. Assessment of mean testicular volume in adolescent school boys of Udaipur district in 10–18 years at different stages of pubertal development. *Indian J Child Health*. 2020; 7(2):63-5.
26. Bianculli C, Bergadá C. Pubertad masculina normal. Un análisis clínico, antropométrico y hormonal en nuestra población. *Rev Hosp Niños B Aires*. 1977; 19(73):21-9.
27. Bianculli C, Garcia S. Semiología de la Pubertad. In Atención integral de adolescentes y jóvenes. Criterios diagnósticos y tratamiento. Ciudad de Buenos Aires: Sociedad Argentina de Pediatría; 1990.Page.148.
28. Agüero G, Berner E. Concordancia en la medición testicular de varones adolescentes con 3 métodos de orquidometría. *Arch Argent Pediatr.* 2021; 119(4):251-8.
29. Sack B, Schäfer M, Kurtz M. The dilemma of adolescent varicoceles: Do they really have to be repaired? *Curr Urol Rep*. 2017; 18(5):38.
30. Ferlin A. Strategies to improve early diagnosis of Klinefelter syndrome. *Expert Rev Endocrinol Metab*. 2020; 15(6):375-8.
31. Vaganée D, Daems F, Aerts W, Dewaide R, et al. Testicular asymmetry in healthy adolescent boys. *BJU Int*. 2018; 122849:654-66.