Asian Journal of Andrology (2021) 23, 501–509
www.asiaandro.com; www.ajandrology.com

Open Access

ORIGINAL ARTICLE

Male Infertility

Testicular volume in infertile versus fertile white-European men: a case-control investigation in the real-life setting

Luca Boeri[1,2], Paolo Capogrosso[1], Eugenio Ventimiglia[1], Walter Cazzaniga[1,3], Edoardo Pozzi[1], Federico Belladelli[1,3], Filippo Pederzoli[1,3], Massimo Alfano[1], Costantino Abbate[1], Emanuele Montanari[2], Luca Valsecchi[4], Enrico Papaleo[4,5], Paola Viganò[4,5], Patrizia Rovere-Querini[6], Suks Minhas[7], Francesco Montorsi[1,3], Andrea Salonia[1,3]

Testicular volume (TV) is considered a good clinical marker of hormonal and spermatogenic function. Accurate reference values for TV measures in infertile and fertile men are lacking. We aimed to assess references values for TV in white-European infertile men and fertile controls. We analyzed clinical and laboratory data from 1940 (95.0%) infertile men and 102 (5.0%) fertile controls. Groups were matched by age using propensity score weighting. TV was assessed using a Prader orchidometer (PO). Circulating hormones and semen parameters were investigated in every male. Descriptive statistics, Spearman's correlation, and logistic regression models tested potential associations between PO-estimated TV values and clinical variables. Receiver operating characteristic (ROC) curves were used to find TV value cutoffs for oligoasthenoteratozoospermia (OAT) and nonobstructive azoospermia (NOA) status in infertile men. The median testicular volume was smaller in infertile than that of fertile men (15.0 ml *vs* 22.5 ml; $P < 0.001$). TV positively correlated with total testosterone, sperm concentration, and progressive sperm motility (all $P \leq 0.001$) in infertile men. At multivariable logistic regression analysis, infertile status ($P < 0.001$) and the presence of left varicocele ($P < 0.001$) were associated with TV < 15 ml. Testicular volume thresholds of 15 ml and 12 ml had a good predictive ability for detecting OAT and NOA status, respectively. In conclusion, infertile men have smaller testicular volume than fertile controls. TV positively correlated with total testosterone, sperm concentration, and progressive motility in infertile men, which was not the case in the age-matched fertile counterparts.

Asian Journal of Andrology (2021) **23**, 501–509; doi: 10.4103/aja.aja_93_20; published online: 12 March 2021

Keywords: hypogonadism; male infertility; orchidometer; Prader; semen analysis; testicular volume

INTRODUCTION

The European Association of Urology (EAU) guidelines for Male Sexual and Reproductive health outline a strong recommendation to simultaneously investigate both partners belonging to any infertile couple, in order to categorize the cause of infertility.[1] Likewise, EAU guidelines strongly recommend examining all men seeking medical help for fertility problems, including men with abnormal semen parameters for urogenital abnormalities.[1] Therefore, a focused diagnostic workup of the male patient must always be undertaken and should include a medical and reproductive history, a focused physical examination, and a detailed semen analysis, with strict adherence to World Health Organization (WHO) reference values for human semen characteristics.[2,3]

As for the physical examination of the infertile male, it has to include a comprehensive evaluation of the volume, texture, and consistency of the testes. Although it is well established that scrotal ultrasound (US), being noninvasive, safe, and inexpensive,[4–6] may allow to precisely measuring testicular volume (TV), assessing testicular anatomy and testicular structure, as well as allowing to find testis tumors and indirect signs of obstruction (*e.g.*, dilatation of rete testis, enlarged epididymis with cystic lesions, or absent vas deferens), in clinical practice, TV is firstly assessed by means of an estimate using the Prader's orchidometer (PO).[7]

Overall, reduced TV is typically associated with different conditions of male-factor infertility (MFI), such as endocrinopathies, primary testicular failure, chromosomal disorders (*e.g.*, Klinefelter's syndrome), cryptorchidism and varicocele, and the consequent poor semen parameters.[4,8] In particular, TV is considered to be a good clinical marker of hormonal and spermatogenic function.[3,9,10] Conversely, its

[1]Division of Experimental Oncology/Unit of Urology, URI, IRCCS Ospedale San Raffaele, Via Olgettina 60, Milan 20131, Italy; [2]Department of Urology, Foundation IRCCS Ca' Granda – Ospedale Maggiore Policlinico, University of Milan, Via Della Commenda 15, Milan 20122, Italy; [3]University Vita-Salute San Raffaele, Via Olgettina 60, Milan 20131, Italy; [4]Department of Obstetrics and Gynaecology, IRCCS Ospedale San Raffaele, Via Olgettina 60, Milan 20131, Italy; [5]Division of Genetics and Cell Biology, Reproductive Sciences Laboratory, IRCCS Ospedale San Raffaele, Via Olgettina 60, Milan 20131, Italy; [6]Division of Immunology, Transplantation and Infectious Diseases Department of Medicine, IRCCS Ospedale San Raffaele, Via Olgettina 60, Milan 20131, Italy; [7]Department of Urology, Imperial Healthcare NHS Trust, Charing Cross Hospital, Fulham Palace Road, Hammersmith, London W6 8RF, UK.
Correspondence: Dr. A Salonia (salonia.andrea@hsr.it)
Received: 19 June 2020; Accepted: 02 December 2020

value in predicting positive versus negative sperm recovery in men with nonobstructive azoospermia (NOA) has not been clarified yet.[4,11] Therefore, an accurate examination of TV exerts important clinical and prognostic implications in infertile men.

As a whole, PO estimation and testicular US investigation are common modalities to measure TV.[8,12] Despite PO may overestimate testis size when compared with US assessment, PO-derived TV has been considered a reliable surrogate of US-measured TV in clinical practice, which is easy to perform and cost-effective.[3,6,9] Notwithstanding, a number of studies have reported PO-derived TV in men, no commonly accepted uniform reference values have been published yet, mostly due to differences in the nature of the populations studied (*e.g.*, geographic area, nourishment, ethnicity, and environmental factors).[4,13,14] So far, in Europe, the reported mean (standard deviation) PO-derived TV was 20.0 (5.0) ml in the general population,[3,4,8] compared to 18.0 (5.0) ml in infertile men.[4,15,16] However, there is a lack of studies that specifically compare TV in homogenous cohorts of infertile versus fertile men in the real-life setting. Thereof, we sought to perform a real-life investigation of PO-derived TV in a homogeneous cohort of white-European men seeking medical help for couple's infertility, and to compare their values to those from a cohort of same-ethnicity, age-matched fertile controls.

PARTICIPANTS AND METHODS

Study design

The analyses of this case-control study were based on a cohort of 2065 consecutive white-European men assessed at a single academic center (San Raffaele Hospital, Milan, Italy) for couple's infertility (noninterracial infertile couples only) between September 2006 and September 2019. According to the WHO criteria, infertility was defined as not conceiving a pregnancy after at least 12 months of unprotected intercourses regardless of whether or not a pregnancy ultimately occurs.[17] Primary infertility was defined when a couple was never able to conceive; secondary infertility was defined according to the inability to conceive following a previous pregnancy.[17] Patients were only enrolled if they were ≥18 years old and ≤60 years old and had either MFI or mixed-factor infertility. MFI was defined after a comprehensive diagnostic evaluation of all the female partners.

Complete data from 102 same-ethnicity, age-matched, fertile controls (*i.e.*, men who had fathered at least one child, spontaneously conceived, with a time to pregnancy within 12 months, as for WHO criteria[17]) were also collected. According to our research protocol, fertile men were recruited via their partners who had been expectant and new mothers at the Department of Obstetrics and Gynaecology (San Raffaele Hospital) and underwent the same comprehensive assessment of the infertile counterpart.

All participants were homogenously assessed by the same expert academic urologist (AS), with a thorough medical history and a complete physical examination (including breast, abdomen, and external genitalia). The Charlson Comorbidity Index (CCI) was used to score health-significant comorbidities, coded using the International Classification of Diseases, 9th revision.[18] Calculated body mass index (BMI) was obtained for each participant, further treated as a categorical variable using the National Institute of Health (NIH) definitions of "normal" (from 18.5 kg m^{-2} to 24.9 kg m^{-2}), "overweight" (from 25 kg m^{-2} to 29.9 kg m^{-2}), and "obese" (≥30 kg m^{-2}).[19] TV was assessed in all cases using PO estimation by the same urologist;[8] for the specific purpose of this study, we recorded the volume of each testicle and the mean value between the two sides. Varicocele was also clinically assessed in every patient.[12]

Venous blood samples were drawn from each patient between 7 a.m. and 11 a.m. after an overnight fast. Follicle-stimulating hormone (FSH), luteinizing hormone (LH), total testosterone (tT), prolactin, thyroid-stimulating hormone (TSH), and sex hormone-binding globulin (SHBG) levels were measured for every individual. Hypogonadism was defined as tT ≤3.03 ng ml^{-1}.[20] Chromosomal analysis and genetic testing were performed in every infertile male (karyotype analysis and tests for Y-chromosome microdeletions and cystic fibrosis mutations).[21]

Participants underwent at least two consecutive semen analyses.[2] For the specific purposes of this study, we considered semen volume, sperm concentration, and progressive sperm motility and morphology. The same laboratory was used for analyses of all parameters.

Data collection followed the principles outlined in the Declaration of Helsinki. All men signed an informed consent agreeing to share their own anonymous information for future studies. The study was approved by the IRCCS San Raffaele Hospital Ethical Committee, Milan, Italy (Prot. 2014 – Pazienti Ambulatoriali).

Statistical methods

Distribution of data was tested with the Shapiro–Wilk test. Data were presented as medians (interquartile range [IQR]) or frequencies (proportions). A 95% confidence interval (CI) was estimated for the association of categorical parameters. Groups were matched by age using propensity score weighting. After matching, the final cohort consisted of 1940 (95.0%) infertile and 102 (5.0%) fertile participants. First, the clinical and demographic characteristics, hormonal values, and semen parameters were compared between infertile and fertile men with the Mann–Whitney U test and the Chi-square test. Similarly, we applied descriptive statistics to compare primary and secondary infertile men. Second, descriptive statistics tested the associations between clinical characteristics, laboratory values, and semen parameters according to a further segregation of primary infertile men only into four androgenic conditions,[22,23] as follows: eugonadal (normal tT [≥3.0 ng ml^{-1}] and normal LH [≤9.4 mUI ml^{-1}]); secondary hypogonadism (low tT [<3.0 ng ml^{-1}] and low/normal LH [≤9.4 mUI ml^{-1}]); primary hypogonadism (low tT [<3.0 ng ml^{-1}] and elevated LH [>9.4 mUI ml^{-1}]); and compensated hypogonadism (normal tT [≥3.0 ng ml^{-1}] and elevated LH [>9.4 mUI ml^{-1}]). Spearman's correlation coefficients were used to depict the association between TV values and different variables. Univariable (UVA) and multivariable (MVA) logistic regression models were used to identify variables associated with TV <15 ml[4,24] in the whole cohort and in infertile men only. Finally, receiver operating characteristic (ROC) curves were generated to find TV value cutoffs (defined as Youden J Index) to predict either oligoasthenoteratozoospermia (OAT) or NOA status in infertile men.[2] Statistical analyses were performed using SPSS version 26 (IBM Corp., Armonk, NY, USA). All tests were two sided, and statistical significance level was determined at $P < 0.05$.

RESULTS

The descriptive statistics of the entire cohort of participants as segregated according to fertility status before and after matching groups by age were presented (**Table 1**). After matching, groups were comparable in terms of age and BMI; conversely, infertile men had higher burden of health significant comorbidities compared to that of fertile controls ($P \leq 0.04$). Overall, the median (IQR) PO-derived TV was 15.0 (11–20) ml versus 22.5 (20–25) ml in infertile and fertile men, respectively ($P < 0.001$); both right and left testicles depicted lower TV in infertile than that of fertile men (both $P < 0.001$). In both cohorts,

Table 1: Descriptive statistics of the participants according to fertility status

Variable	*Before propensity score weighting*			*After propensity score weighting*		
	Fertile	*Infertile*	[a]*P*	*Fertile*	*Infertile*	[a]*P*
Participants, *n* (%)	102 (4.7)	2065 (95.3)		102 (5.0)	1940 (95.0)	
Age (year)			0.04			0.7
Median (IQR)	36.0 (33–39)	37.0 (33–41)		36.0 (33–39)	36.0 (33–39)	
Range	19–48	19–48		19–48	19–48	
BMI (kg m^{-2})			0.8			0.9
Median (IQR)	24.5 (23.1–27.9)	24.9 (23.2–27.1)		24.5 (23.1–27.9)	24.6 (23.1–26.9)	
Range	18.5–37.9	18.5–45.7		18.5–37.9	18.5–45.7	
BMI categorized (kg m^{-2}), *n* (%)			0.4			0.5
18.5–24.9	55 (53.9)	1040 (50.4)		55 (53.9)	1010 (52.1)	
25.0–29.9	34 (33.3)	832 (40.3)		34 (33.3)	737 (38.0)	
≥30	13 (12.8)	193 (9.3)		13 (12.8)	193 (9.9)	
CCI (score)			0.03			0.04
Median (IQR)	0 (0)	0 (0)		0 (0)	0 (0)	
Mean (s.d.)	0.2 (0.2)	0.9 (0.4)		0.2 (0.2)	0.8 (0.3)	
Range	0–2	0–8		0–2	0–8	
Cryptorchidism, *n* (%)		174 (8.4)			170 (8.7)	
Karyotype abnormalities, *n* (%)		62 (3.0)			58 (2.9)	
Chromosome Y deletions, *n* (%)		14 (0.6)			14 (0.7)	
Mean testicular volume (Prader's estimation; ml)			<0.001			<0.001
Median (IQR)	22.5 (20–25)	15.0 (12–20)		22.5 (20–25)	15.0 (11–20)	
Range	10–30	2–25		10–30	2–25	
Testicular volume <15 ml, *n* (%)	12 (11.7)	1061 (51.4)	<0.001	12 (11.7)	993 (51.2)	<0.001
Left testicular volume (Prader's estimation; ml)			<0.001			<0.001
Median (IQR)	20.0 (20–25)[b]	15.0 (12–20)[b]		20.0 (20–25)[b]	15.0 (12–20)[b]	
Range	10–30	2–25		10–30	2–25	
Right testicular volume (Prader's estimation; ml)			<0.001			<0.001
Median (IQR)	25.0 (20–25)	15.0 (12–20)		25.0 (20–25)	15.0 (12–20)	
Range	10–30	2–25		10–30	2–25	
Varicocele, *n* (%)	15 (14.7)	946 (46.7)	<0.001	15 (14.7)	885 (45.6)	<0.001
tT (ng ml^{-1})			0.02			0.02
Median (IQR)	4.9 (4.0–5.9)	4.5 (3.4–5.7)		4.9 (4.0–5.9)	4.5 (3.3–5.5)	
Range	2.4–9.3	0.1–28.4		2.4–9.3	0.1–28.4	
tT <3 ng ml^{-1}, *n* (%)	8 (7.8)	330 (16.0)	0.03	8 (7.8)	304 (15.7)	0.03
FSH (mUI ml^{-1})			<0.001			<0.001
Median (IQR)	4.1 (3.0–5.6)	5.7 (3.4–11.8)		4.1 (3.0–5.6)	5.7 (3.3–11.5)	
Range	1.4–12.6	0.1–98.2		1.4–12.6	0.1–98.2	
LH (mUI ml^{-1})			0.3			0.6
Median (IQR)	4.4 (3.5–5.6)	4.9 (2.9–6.2)		4.4 (3.5–5.6)	4.4 (2.9–6.1)	
Range	1.5–10.4	0.2–67.0		1.5–10.4	0.2–57.0	
Prolactin (ng ml^{-1})			0.7			0.7
Median (IQR)	8.8 (6.7–11.5)	8.5 (6.2–12.0)		8.8 (6.7–11.5)	8.7 (6.6–12.0)	
Range	1.0–67.0	0.8–58.3		1.0–67.0	1.0–58.3	
SHBG (nmol l^{-1})			0.4			0.5
Median (IQR)	33.5 (26–46)	34.0 (25–44)		33.5 (26–46)	34.0 (25–46)	
Range	15–75	20–140		15–75	20–140	
TSH (mUI l^{-1})			0.6			0.3
Median (IQR)	1.7 (1.4–2.1)	1.7 (1.1–2.3)		1.7 (1.4–2.1)	1.9 (1.3–3.6)	
Range	0.7–10.7	0.3–9.7		0.7–10.7	0.7–9.7	
Semen volume (ml)			0.5			0.5
Median (IQR)	3.0 (2.0–4.0)	3.0 (2.0–4.0)		3.0 (2.0–4.0)	3.0 (2.0–4.0)	
Range	0.1–9.0	0.1–10.0		0.1–9.0	0.1–10.0	
Sperm concentration (×10^6 ml^{-1})			<0.001			<0.001
Median (IQR)	50.0 (26.7–70.0)	18.0 (4.9–32.0)		50.0 (26.7–70.0)	18.3 (4.9–45.0)	
Range	1.0–150.0	0.5–455.3		1.0–150.0	0.5–455.3	
Concentration ≤15 ×10^6 ml^{-1}, *n* (%)	9 (8.8)	776 (45.6)	<0.001	9 (8.8)	683 (43.2)	<0.001
Progressive sperm motility (%)			<0.001			<0.001
Median (IQR)	46.0 (35.0–57.2)	25.0 (10.0–39.0)		46.0 (35.0–57.2)	25.0 (10.0–40.0)	
Range	15.0–80.0	0–96.0		15.0–80.0	0–96.0	

Contd...

Table 1: Contd...

Variable	*Before propensity score weighting*			*After propensity score weighting*		
	Fertile	*Infertile*	[a]*P*	*Fertile*	*Infertile*	[a]*P*
Progressive sperm motility ≤32%, *n* (%)	19 (18.6)	1113 (65.4)	<0.001	19 (18.6)	1017 (64.3)	<0.001
Normal sperm morphology (%)			0.7			0.7
Median (IQR)	3.0 (1.0–10.0)	3.0 (1.0–10.0)		3.0 (1.0–10.0)	3.0 (1.0–10.0)	
Range	1.0–49.0	0–100.0		1.0–49.0	0–100.0	
Normal sperm morphology ≤4%, *n* (%)	59 (57.8)	944 (55.4)	0.6	59 (57.8)	863 (54.5)	0.5
Azoospermia cases, *n* (%)		363 (17.6)			357 (18.4)	

[a]*P* value according to the Mann–Whitney U test and Chi-square test; [b]*P*<0.001, left testicle versus right testicle of the same group, according to the Wilcoxon signed-rank test. IQR: interquartile range; BMI: body mass index; CCI: Charlson Comorbidity Index; tT: total testosterone; SHBG: sex hormone-binding globulin; TSH: thyroid-stimulating hormone; s.d.: standard deviation

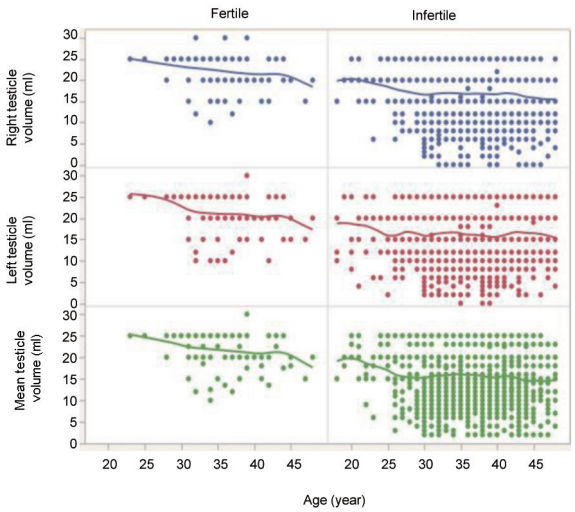


Figure 1: Testicular volume distribution in fertile and infertile participants by age.

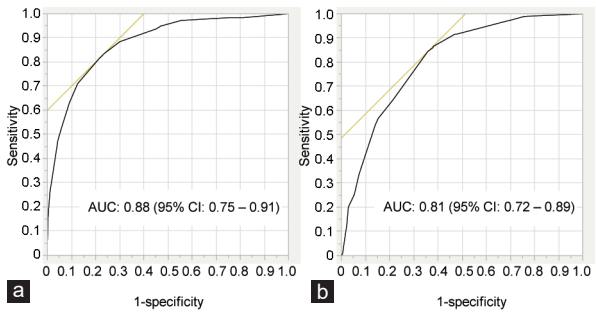


Figure 2: ROC analysis demonstrating the sensitivity and specificity of testicular volume in detecting (**a**) nonobstructive azoospermia and (**b**) oligoasthenoteratozoospermia status in infertile men. ROC: receiver operating characteristic; AUC: area under the curve; CI: confidence interval.

the left testis was smaller than the contralateral (both $P < 0.001$). TV distribution according to age in fertile and infertile men was presented (**Figure 1**). TV did not correlate with age in both groups. Left varicocele was significantly more prevalent in infertile compared to fertile men ($P < 0.001$).

TV distribution in men with and without varicocele is shown in **Supplementary Figure 1**. In infertile patients, the mean TV and left TV were lower in men with left varicocele than those without varicocele (both $P < 0.01$), but this was not the case for fertile controls. Testicular volumes were significantly lower in infertile men with a history of cryptorchidism and in men with Klinefelter's syndrome than those without a history of undescended testes or karyotype alterations (all $P < 0.001$; **Supplementary Figure 2**).

As expected, gonadal androgenic status was poorer in infertile compared to fertile men ($P \leq 0.03$); likewise, sperm concentration and sperm progressive motility were lower in infertile than fertile individuals (both $P < 0.001$). Morphology data did not differ between infertile and fertile men.

Descriptive statistics of infertile men as segregated according to primary versus secondary infertility are shown in **Table 2**. Primary infertile men were younger, had a lower CCI score, and had smaller TV than those with secondary infertility (all $P \leq 0.02$).

The characteristics and the descriptive statistics of primary infertile men according to their gonadal status are presented in **Table 3**. Of all, eugonadism and primary, secondary, and compensated hypogonadism were found in 1350 (78.1%) and 44 (2.5%), 227 (13.1%), and 107 (6.2%) men, respectively. The median age did not vary markedly with gonadal status. The median BMI was significantly different across groups, with primary and secondary hypogonadal men having the highest values (all $P < 0.001$, compared to that of eugonadal group). Men with primary and compensated hypogonadism had lower TV compared to that of eugonadal individuals (both $P < 0.001$). Moreover, primary and compensated hypogonadal men depicted the lowest values of sperm concentration and progressive sperm motility (all $P < 0.03$).

Spearmen's correlation revealed a positive association between TV and tT (rho = 0.16; $P \leq 0.001$), sperm concentration (rho = 0.46; $P \leq 0.001$), and progressive sperm motility (rho = 0.21; $P \leq 0.001$) in infertile men. Conversely, a negative correlation was found between TV and FSH (rho = −0.16; $P \leq 0.001$) and LH (rho = −0.45; $P \leq 0.001$) in infertile men. No significant correlations were found between TV and clinical variables in fertile men.

Logistic regression models predicting TV <15 ml in the whole cohort of participants and in the subcohort of infertile men are shown in **Table 4**. Overall, at MVA logistic regression analysis, infertile status (odds ratio [OR] = 7.2; $P < 0.001$) and the presence of varicocele (OR = 2.1; $P < 0.001$) were associated with TV <15 ml. In the cohort of infertile men only, the presence of varicocele (OR = 1.7; $P < 0.001$), azoospermia (OR = 3.2; $P < 0.001$), a primary hypogonadism (OR = 7.2; $P < 0.001$), and a compensated hypogonadal status (OR = 1.8; $P = 0.04$) were independently associated with TV <15 ml, after accounting for a history of undescended testes and karyotype abnormalities.

ROC analysis showed that TV had a good predictive ability for NOA status in infertile men (area under the curve [AUC]: 0.88; 95% CI: 0.70–0.89; **Figure 2a**). A TV cutoff value of 12 ml could diagnose NOA with 81.2% sensitivity and 78.2% specificity. Similarly, a TV cutoff value of 15 ml could diagnose OAT with 84.4% sensitivity and 74.1% specificity (**Figure 2b**).

Table 2: Descriptive statistics of infertile patients according to primary versus secondary infertility after matching (*n*=1940)

Variable	*Primary*	*Secondary*	[a]*P*
Patients, *n* (%)	1728 (89.1)	212 (10.9)	
Age (year)			0.02
Median (IQR)	36.0 (33–39)	38.0 (35–41)	
Range	19–48	19–48	
BMI (kg m^{-2})			0.7
Median (IQR)	24.9 (23.2–27.8)	24.6 (23.1–27.1)	
Range	18.5–45.7	18.5–41.6	
CCI (score)			0.02
Median (IQR)	0 (0)	0 (0)	
Mean (s.d.)	0.5 (0.2)	0.7 (0.4)	
Range	0–7	0–8	
Mean testicular volume (Prader's estimation; ml)			<0.001
Median (IQR)	15.0 (12–20)	18.0 (13–20)	
Range	2–25	4–25	
Testicular volume <15 ml, *n* (%)	909 (52.6)	84 (39.6)	<0.01
Left testicular volume (Prader's estimation; ml)			<0.001
Median (IQR)	15.0 (12–20)	20.0 (13–24)	
Range	2–25	2–25	
Right testicular volume (Prader's estimation; ml)			<0.001
Median (IQR)	15.0 (12–20)	20.0 (12–22)	
Range	2–25	2–25	
Varicocele, *n* (%)	786 (45.5)	99 (46.6)	0.3
tT (ng ml^{-1})			0.5
Median (IQR)	4.5 (3.4–5.6)	4.4 (3.3–5.8)	
Range	0.1–28.4	0.1–10.2	
tT <3 ng ml^{-1}, *n* (%)	271 (15.7)	33 (15.5)	0.7
FSH (mUI ml^{-1})			<0.001
Median (IQR)	5.9 (3.1–10.8)	4.1 (2.6–8.7)	
Range	0.1–98.2	0.3–64.1	
LH (mUI ml^{-1})			<0.001
Median (IQR)	4.2 (1.9–5.9)	3.0 (2.3–5.1)	
Range	0.2–57.0	0.2–21.3	
Prolactin (ng ml^{-1})			0.7
Median (IQR)	8.6 (6.2–12.1)	8.7 (5.9–10.4)	
Range	1.0–58.3	2.9–32.3	
SHBG (nmol l^{-1})			0.7
Median (IQR)	32.4 (24–42)	32.1 (23–42)	
Range	20–140	20–95	
TSH (mUI l^{-1})			0.5
Median (IQR)	1.6 (1.2–2.3)	1.6 (1.2–2.9)	
Range	0.7–9.7	0.6–9.2	
Semen volume (ml)			0.3
Median (IQR)	3.0 (2.0–4.0)	3.0 (2.0–4.0)	
Range	0.1–9.0	0.1–10.0	
Sperm concentration (×10^6 ml^{-1})			<0.001
Median (IQR)	18.3 (4.0–44.0)	25.0 (9.6–53.0)	
Range	0.5–455.3	0.5–167.9	
Progressive sperm motility (%)			0.1
Median (IQR)	25.0 (10–39)	26.0 (9.6–53)	
Range	0–96.0	0.5–167.9	
Normal sperm morphology (%)			0.5
Median (IQR)	3.0 (1.0–10.0)	3.0 (1.0–10.0)	
Range	0–100.0	0–93.0	
Azoospermia cases, *n* (%)	335 (19.4)	22 (10.3)	0.01

[a]*P* value according to the Mann–Whitney U test and Chi-square test. IQR: interquartile range; BMI: body mass index; CCI: Charlson Comorbidity Index; tT: total testosterone; SHBG: sex hormone-binding globulin; TSH: thyroid-stimulating hormone; s.d.: standard deviation

DISCUSSION

Testicular volume assessment is a relevant part of the diagnostic workup of every infertile man in the real-life setting; indeed, a reduced TV has been related to poor semen parameters, hormonal abnormalities, and pregnancy outcomes.[3,4,9,25] Despite a number of epidemiological studies have reported TV measures in infertile and fertile men, no practical normal reference values for TV are available in men presenting for couple's infertility, mostly because of relevant differences

Table 3: Descriptive statistics of primary infertile patients according to hypogonadism status after matching (*n*=1728)

Variable	*Eugonadism*	*Primary hypogonadism*	*Secondary hypogonadism*	*Compensated hypogonadism*	[a]*P*
Patients, *n* (%)	1350 (78.1)	44 (2.5)	227 (13.1)	107 (6.2)	
Age (year)					0.6
Median (IQR)	36.0 (33–40)	37.0 (33–41)	38.0 (34–41)	37.0 (32–40)	
Range	19–48	20–48	19–48	19–48	
BMI (kg m^{-2})					<0.001
Median (IQR)	24.7 (23.1–26.8)	26.8 (25.1–31.0)[b]	26.6 (24.7–29.7)[b]	24.9 (22.7–26.8)	
Range	18.5–31.4	19.3–40.1	20.4–45.7	18.8–39.1	
CCI (score)					0.03
Median (IQR)	0 (0)	0 (0)	0 (0)[b]	0 (0)[b]	
Mean (s.d.)	0.1 (0.6)	0.3 (0.5)	0.6 (0.3)	0.9 (0.2)	
Range	0–5	0–6	0–7	0–7	
Mean testicular volume (Prader's estimation; ml)					<0.001
Median (IQR)	15.0 (12–20)	7.0 (4–10)[b]	15.0 (11–20)	10.0 (7–13)[b]	
Range	5–25	2–25	4–25	2–25	
Testicular volume <15 ml, *n* (%)	647 (47.9)	39 (89.4)	129 (56.8)	94 (87.5)	<0.001
Left testicular volume (Prader's estimation; ml)					<0.001
Median (IQR)	15.0 (12–20)	6.0 (4–10)[b]	15.0 (10–20)	10.0 (6–12)[b]	
Range	2–25	2–25	2–25	2–25	
Right testicular volume (Prader's estimation; ml)					<0.001
Median (IQR)	15.0 (12–20)	8.0 (4–10)[b]	15.0 (12–20)	10.0 (6–15)[b]	
Range	2–25	2–25	5–25	2–25	
tT (ng ml^{-1})					<0.001
Median (IQR)	4.9 (3.9–6.0)	2.2 (1.7–2.7)	2.5 (2.2–2.8)	4.7 (3.7–5.7)	
Range	3.0–28.4	0.5–2.9	0.5–2.9	3.1–21.1	
FSH (mUI ml^{-1})					<0.001
Median (IQR)	5.4 (3.2–9.5)	26.4 (23.4–39.4)	5.8 (3.4–11.5)	24.0 (15.0–32.6)	
Range	0.1–40.3	11.0–74.0	1.0–59.4	2.1–98.2	
LH (mUI ml^{-1})					<0.001
Median (IQR)	4.1 (2.9–5.6)	14.4 (12.6–19.7)	3.6 (2.4–5.2)	12.0 (10.0–15.0)	
Range	0.2–9.4	9.7–57.0	1.0–9.4	9.5–43.4	
Prolactin (ng ml^{-1})					<0.001
Median (IQR)	8.5 (6.2–11.9)	10.8 (7.1–16.9)	8.5 (6.0–12.1)	10.8 (7.1–18.3)	
Range	1.0–57.0	4.5–58.3	2.0–58.0	3.2–50.0	
SHBG (nmol l^{-1})					<0.001
Median (IQR)	34.0 (26–43)	28.5 (22–34)	22.0 (17–29)	35.0 (26–45)	
Range	20–75	20–42	20–53	20–140	
TSH (mUI l^{-1})					0.9
Median (IQR)	1.7 (1.2–2.3)	1.6 (1.2–2.2)	1.5 (1.2–2.4)	1.6 (1.2–2.4)	
Range	0.7–9.7	0.8–6.0	0.7–9.6	0.7–9.0	
Semen volume (ml)					0.2
Median (IQR)	3.0 (1.0–4.0)	3.0 (2.0–4.0)	3.0 (1.0–4.0)	3.0 (2.0–4.0)	
Range	0.1–9.0	0.1–8.0	0.2–9.0	0.1–9.0	
Sperm concentration (×10^6 ml^{-1})					
Median (IQR)	13.8 (3.1–38.7)	0.4 (0.1–5.8)[b]	10.0 (2.0–32.0)	2.2 (0.5–11.3)[b]	<0.001
Range	1.0–455.3	0.5–19.0	0.5–155.5	0.5–70.0	
Progressive sperm motility (%)					0.03
Median (IQR)	24.0 (9.0–37.0)	12.0 (2.5–28.0)[b]	14.0 (5.0–31.0)	10.0 (0–25.0)[b]	
Range	0–96.0	1.0–30.0	0–72.0	0–65.0	
Normal sperm morphology (%)					0.6
Median (IQR)	3.0 (1.0–10.0)	1.0 (0–9.0)	2.0 (1.0–9.0)	2.0 (1.0–9.0)	
Range	0–100.0	0–12.0	0–83.0	0–20.0	

[a]*P* value according to the Kruskal–Wallis test and Fisher's test; [b]*P*<0.001, the selected group versus the eugonadal group. IQR: interquartile range; BMI: body mass index; CCI: Charlson Comorbidity Index; tT: total testosterone; SHBG: sex hormone-binding globulin; TSH: thyroid-stimulating hormone; s.d.: standard deviation

in terms of modalities to assess (*i.e.*, US-based *vs* PO-derived) and the characteristics of the studied populations (*i.e.*, ethnicity, age at presentation).[4,13,14] More in depth, there is a lack of reliable studies comparing TVs, as obtained with a standardized methodology, between cohorts of same-ethnicity, age-matched fertile and infertile individuals in the real-life setting.

Current case–control findings depicted median (IQR) PO-derived TVs of 15.0 ml (11–20 ml) and 22.5 ml (20–25 ml) in infertile and fertile men, respectively ($P < 0.001$). Moreover, we observed that both right and left testicles depicted lower TV in infertile than fertile men. In both cohorts, the left testis was consistently smaller than the contralateral, and it was the case in both fertile (mean: 21 ml *vs* 23 ml, left testis *vs*

Table 4: Logistic regression models predicting testicular volume <15 ml in the whole cohort and in infertile men after matching

Variable	*Whole cohort*		*Infertile men*	
	UVA model (OR; P [95% CI])	*MVA model (OR; P [95% CI])*	*UVA model (OR; P [95% CI])*	*MVA model (OR; P [95% CI])*
Age	1.01; 0.99 [0.57–1.72]	NA	1.01; 0.78 [0.53–1.61]	NA
BMI	0.44; 0.15 [0.14–1.35]	NA	0.41; 0.12 [0.13–1.28]	NA
CCI	1.37; 0.07 [0.87–5.11]	NA	2.95; 0.11 [0.76–5.56]	NA
Varicocele	2.41; <0.001 [2.01–2.86]	2.1; <0.001 [1.69–2.54]	2.26; <0.001 [1.89–2.71]	1.66; <0.001 [1.28–2.14]
Infertile status	7.92; <0.001 [4.31–14.56]	7.22; <0.001 [3.76–13.32]	NA	NA
Azoospermia	NA	NA	3.39; <0.001 [2.63–4.36]	3.21; <0.001 [2.43–6.16]
Cryptorchidism	NA	NA	2.68; <0.001 [1.90–3.78]	1.31; 0.31 [0.76–2.22]
Karyotype abnormalities	NA	NA	3.69; <0.001 [1.99–6.84]	1.21; 0.88 [0.65–2.63]
Hypogonadism status	NA	NA		
Eugonadism	NA	NA	Ref	Ref
Primary hypogonadism	NA	NA	11.3; <0.001 [3.41–10.67]	7.21; <0.001 [2.12–10.54]
Secondary hypogonadism	NA	NA	1.31; 0.06 [0.98–1.76]	1.01; 0.98 [0.68–1.44]
Compensated hypogonadism	NA	NA	8.43; <0.001 [4.31–9.65]	1.75; 0.04 [1.11–4.21]

UVA: univariate model; MVA: multivariate model, BMI: body mass index; CCI: Charlson Comorbidity Index; CI: confidence interval; OR: odds ratio; NA: not applicable; Ref: reference value

right testis, respectively) and infertile men (mean: 15 ml *vs* 16.7 ml, left testis *vs* right testis, respectively). These findings also have been confirmed throughout the age-matched comparisons.

We consider these findings of paramount clinical importance because they can be used as reference values for TVs in white-European fertile and infertile men in the daily clinical diagnostic workup. Indeed, a great variability in TV measures was previously reported in literature. Takihara *et al.*,[14] for instance, reported that the normal range of TV was greater than 14 ml in Japanese men and greater than 17 ml in the USA. Similarly, in a population-based study of healthy young Korean men, the reported threshold for TV was approximately 18 ml.[13] In Europe, epidemiological studies showed that the mean PO-derived TV was 20 ml from the general population and 18 ml in infertile men.[3,8,15,16] Our findings are in line with European data as for the general population, but they clearly detail an even worse condition related to TVs in the subcohort of infertile men. In particular, we showed that infertile men had 8 times greater risk of having TV <15 ml than that of fertile controls. The cause for the ethnic difference in TV reported in the literature is currently unknown, but may be related with differences in average body size, dietary customs, lifestyle, and *in utero* exposure to smoking and so on.[16]

Differences in TV between the right and left sides are frequently found in clinical practice. Some authors found that TV of the right side was larger than that of the contralateral,[26,27] but others failed to find any difference between the two testes.[13] Our study showed that the left testicle was consistently smaller than the contralateral in both fertile and infertile men. This difference can be mostly related to the presence of varicocele, that is more frequently found on the left and has been associated with testicular dysfunction.[28] Accordingly, our findings confirmed that infertile men with varicocele had lower TV than those without varicocele, and the presence of varicocele was associated with a 2-fold higher risk of having TV <15 ml. Of relevance, a negative association between varicocele and TV was also found in men recruited from the general population (*i.e.*, not selected with regard to their fertility status).[29] Moreover, the role of shear-wave elastography in patients with varicocele has been investigated;[30] in the case of varicocele, the testis was stiffer than that of the contralateral, and shear-wave elastography had been considered helpful in assessing testicular pathologic alterations owing to varicocele.

TV is closely related to both exocrine (spermatogenesis) and endocrine (steroidogenesis) functions of the testis. Accordingly, it is known that TV is lower in hypogonadal men.[31] Ruiz-Olvera *et al.*[10] analyzed a cohort of 312 men with either sexual dysfunction or infertility and showed that TV was strongly associated with tT values. Similarly, studies including men from the general population (thus including both fertile and infertile men) revealed that TV was positively associated with tT and inversely correlated with FSH and LH values.[13,29] Our study confirmed both those latter observations, with a positive correlation between TV and tT and a negative correlation between TV and FSH/LH values in infertile men. Conversely, TV was not associated with hormonal parameters in fertile men. Moreover, when we looked at primary infertile men according to different classes of hypogonadism, we found that individuals with primary and compensated hypogonadism had the lowest TV values among all groups. Moreover, primary and compensated hypogonadal men had the highest risk of having TV <15 ml, which has already been associated with a certain degree of spermatogenic dysfunction.[24] Despite normal tT values, this would suggest that infertile men with compensated hypogonadism might have an initial impaired testicular function.[23]

Previous reports have reported the association between a reduced PO-derived TV and poor semen parameters, thus including lower sperm concentration,[13,15,29,32,33] lower sperm motility,[13,32,33] and lower sperm morphology rates as compared with that of reference ranges.[15] Moreover, in a longitudinal study including 4045 subjects with sexual dysfunction, TV was also associated with fatherhood.[25] To better quantify the risk of impaired semen parameters, Sakamoto *et al.*[33] reported that semen profile would have been subnormal in cases with TV <20 ml, and even critically impaired <14 ml.[14] Recently, the role of magnetic resonance imaging (MRI) in the evaluation of infertile men with a specific focus on sperm parameters has been considered. In this context, testicular MRI showed high predictive accuracy in differentiating obstructive azoospermia from NOA[34,35] in infertile men, and diffusion-weighted MRI imaging was found to increase with aging and found to be associated with spermatogenesis hypofunction.[36,37] As a whole, our results confirmed a positive association between TV and sperm concentration and sperm progressive motility in infertile men. Furthermore, current analyses depicted that PO-derived TV of 15 ml and 12 ml had good predictive ability for detecting OAT and NOA status, respectively. These thresholds may be of relevance throughout the diagnostic workup of infertile men as a reliable initial marker of possible severe semen impairment in the real-life setting.

The clinical implication of our study is several-fold. First, we conducted the first case-control investigation of consistent PO-derived TV measurement in a homogenous, same-ethnicity, age-matched cohort of infertile men versus fertile individuals, thus providing reliable "normal" reference values of TV in white-European men. Second, we detailed the importance of TV assessment in infertile subjects relatively to hormonal and seminal outcomes as a key step throughout the daily diagnostic workup of men presenting for couple's infertility.[12] All reported findings are of relevance as compared with previously published observations because we homogenously investigated an age-matched cohort of fertile and infertile patients with a thorough hormonal and semen evaluation, using a really comparable group of fertile men as for WHO definition criteria.[17] Conversely, most of the previous studies have deliberately excluded infertile men[13] or conditions that could have altered TV (*e.g.*, varicocele, cryptorchidism),[10,13] thus potentially limiting the clinical validity of their findings in the real-life setting.

Our study is not devoid of limitations. First, despite the fact that we analyzed a relatively large, homogeneous, same-ethnicity cohort of infertile and age-comparable fertile men, this was a single-center-based study, raising the possibility of selection biases; thereof, larger studies across different centers and cohorts are needed to externally validate our findings. Second, the current findings were based by definition only on PO-derived TV measurements, that might have overestimated the size as compared to formal US-based assessments.[4] However, our assessment follows the rules also detailed by the most recently updated EAU guidelines, which confirm PO-derived TV assessment as a reliable surrogate of US-measured TV in the day-to-day clinical practice. Moreover, at least in this specific cohort, US assessments have been performed by a number of US specialists, thus allowing a potential inter-observer variability; conversely, all PO-derived TV measurements have been performed over time by a single-expert uro-andrologist (AS), thus providing a good reliability in terms of method standardization throughout the study period.

In conclusion, in this case-control study, we detailed median (IQR) PO-derived reference values for TV in both infertile (15.0 [11–20] ml) and fertile (22.5 [20–25] ml) white-European men. Testicular volume positively correlated with tT, sperm concentration, and progressive sperm motility in infertile men, which was not the case in the age-matched fertile counterparts. Of all, in infertile men, PO-derived TV thresholds of 15 ml and 12 ml had good predictive ability for detecting OAT and NOA status, respectively. Primary infertile men with primary and compensated hypogonadism are at higher risk of smaller TV, which is eventually suggestive for the overall impaired testicular function.

AUTHOR CONTRIBUTIONS

LB designed the study, collected data, performed statistical analyses, and drafted the manuscript. AS designed the study, collected data, and coordinated all the steps of the study. PC, EV, WC, EP, FB, MA, FP, CA, LV, EP, PV, PRQ, SM, EM, and FM collected data, participated in coordination, and revised the manuscript. All authors read and approved the final manuscript.

COMPETING INTERESTS

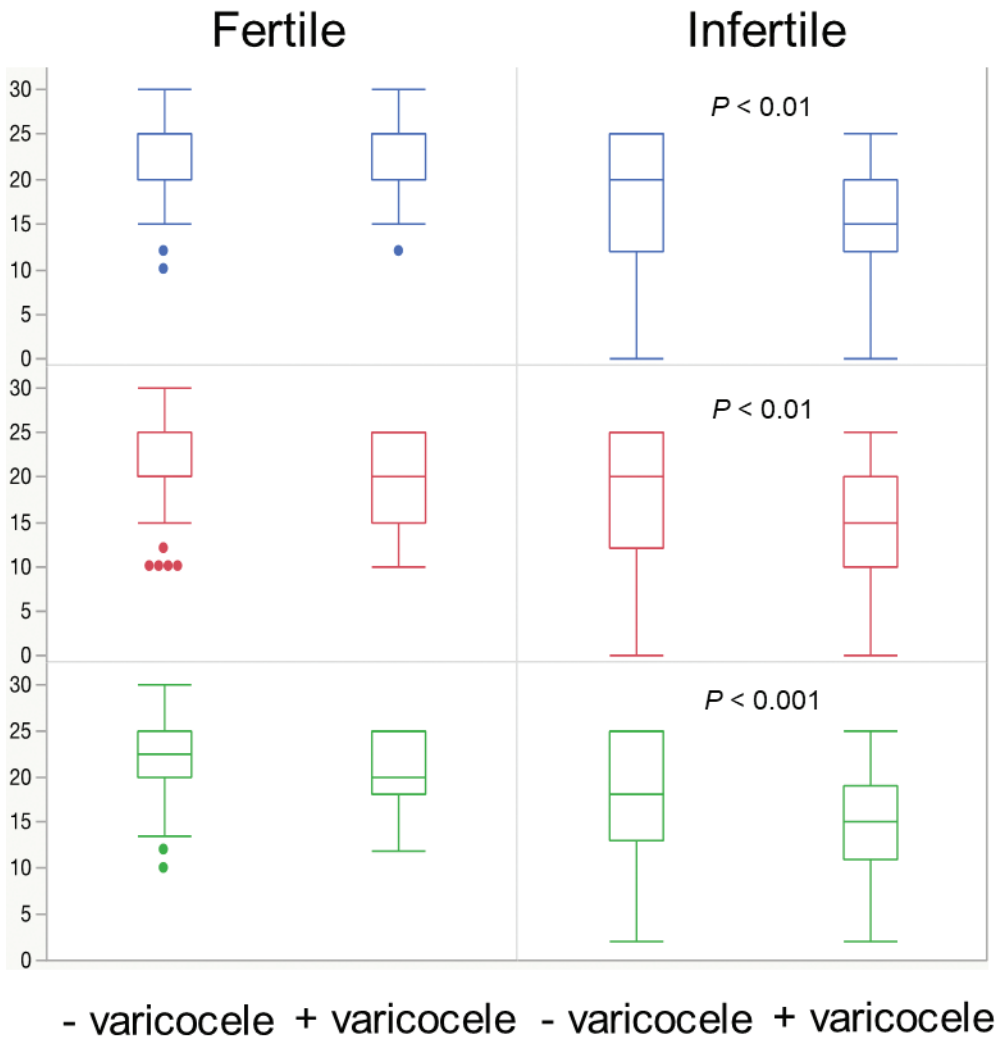
All authors declared no competing interests.

Supplementary Information is linked to the online version of the paper on the *Asian Journal of Andrology* website.

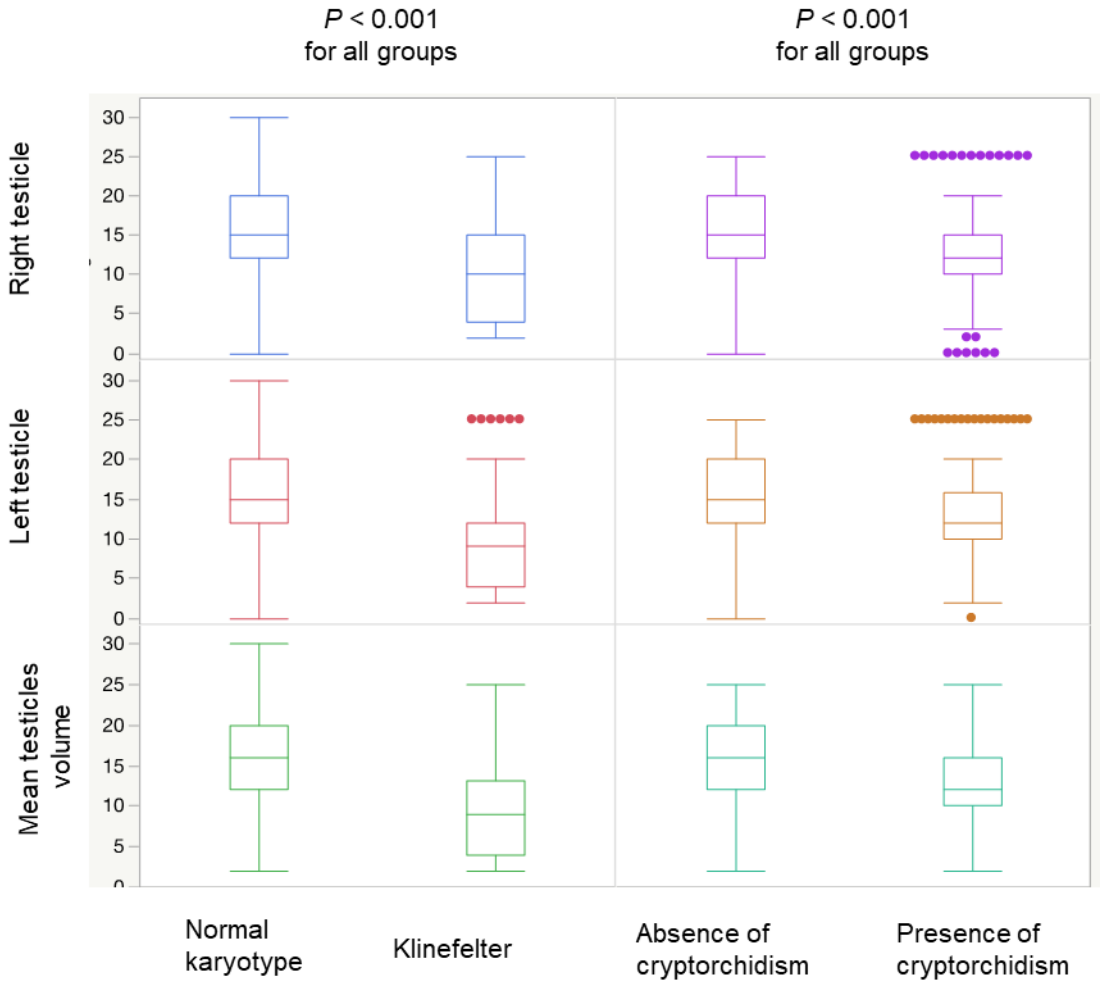
REFERENCES

1 Tharakan T, Bettocchi C, Carvalho J, Corona G, Joensen UN, *et al.* Male sexual and reproductive health-does the urologist have a role in addressing gender inequality in life expectancy? *Eur Urol Focus* 2020; 6: 791–800.
2 Cooper TG, Noonan E, von Eckardstein S, Auger J, Baker HW, *et al.* World Health Organization reference values for human semen characteristics. *Hum Reprod Update* 2010; 16: 231–45.
3 Lotti F, Frizza F, Balercia G, Barbonetti A, Behre HM, *et al.* The European Academy of Andrology (EAA) ultrasound study on healthy, fertile men: clinical, seminal and biochemical characteristics. *Andrology* 2020; 8: 1005–20.
4 Lotti F, Maggi M. Ultrasound of the male genital tract in relation to male reproductive health. *Hum Reprod Update* 2015; 21: 56–83.
5 Mittal PK, Little B, Harri PA, Miller FH, Alexander LF, *et al.* Role of imaging in the evaluation of male infertility. *Radiographics* 2017; 37: 837–54.
6 Jurewicz M, Gilbert BR. Imaging and angiography in male factor infertility. *Fertil Steril* 2016; 105: 1432–42.
7 Behre HM, Nashan D, Nieschlag E. Objective measurement of testicular volume by ultrasonography: evaluation of the technique and comparison with orchidometer estimates. *Int J Androl* 1989; 12: 395–403.
8 Nieschlag E, Behre HM. Anamnesis and physical examination. In: Nieschlag E, Behre HM, Nieschlag S, editors. Andrology: Male Reproductive Health and Dysfunction. 3rd ed. Berlin: Springer-Verlag Berlin Heidelberg; 2010. p93–100.
9 Rastrelli G, Corona G, Lotti F, Boddi V, Mannucci E, *et al.* Relationship of testis size and LH levels with incidence of major adverse cardiovascular events in older men with sexual dysfunction. *J Sex Med* 2013; 10: 2761–73.
10 Ruiz-Olvera SF, Rajmil O, Sanchez-Curbelo JR, Vinay J, Rodriguez-Espinosa J, *et al.* Association of serum testosterone levels and testicular volume in adult patients. *Andrologia* 2018;50: e12933.
11 Bernie AM, Mata DA, Ramasamy R, Schlegel PN. Comparison of microdissection testicular sperm extraction, conventional testicular sperm extraction, and testicular sperm aspiration for nonobstructive azoospermia: a systematic review and meta-analysis. *Fertil Steril* 2015; 104: 1099–103.e1–3.
12 Salonia A, Bettocchi C, Carvalho J, Corona G, Jones TH, *et al.* EAU guidelines on sexual and reproductive health. Available from: https://uroweb.org/wp-content/uploads/EAU-Guidelines-on-Sexual-and-Reproductive-Health-2020.pdf. [Last assessed on 2020 April 05].
13 Bahk JY, Jung JH, Jin LM, Min SK. Cut-off value of testes volume in young adults and correlation among testes volume, body mass index, hormonal level, and seminal profiles. *Urology* 2010; 75: 1318–23.
14 Takihara H, Cosentino MJ, Sakatoku J, Cockett AT. Significance of testicular size measurement in andrology: II. Correlation of testicular size with testicular function. *J Urol* 1987; 137: 416–9.
15 Jørgensen N, Carlsen E, Nermoen I, Punab M, Suominen J, *et al.* East-West gradient in semen quality in the Nordic-Baltic area: a study of men from the general population in Denmark, Norway, Estonia and Finland. *Hum Reprod* 2002; 17: 2199–208.
16 Jensen TK, Jørgensen N, Punab M, Haugen TB, Suominen J, *et al.* Association of in utero exposure to maternal smoking with reduced semen quality and testis size in adulthood: a cross-sectional study of 1,770 young men from the general population in five European countries. *Am J Epidemiol* 2004; 159: 49–58.
17 World Health Organization. Infertility definitions and terminology. Available from:http://www.who.int/reproductivehealth/topics/infertility/definitions. [Last accessed on 2019 May 02].
18 Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40: 373–83.
19 World Health Organization. Body mass index – BMI nutrition. Available from: http://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi. [Last accessed on 2010 May 05].
20 Bhasin S, Cunningham GR, Hayes FJ, Matsumoto AM, Snyder PJ, *et al.* Testosterone therapy in men with androgen deficiency syndromes: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* 2010; 95: 2536–59.
21 Ventimiglia E, Capogrosso P, Boeri L, Pederzoli F, Cazzaniga W, *et al.* When to perform karyotype analysis in infertile men? Validation of the European association of urology guidelines with the proposal of a new predictive model. *Eur Urol* 2016; 70: 920–3.
22 Tajar A, Forti G, O'Neill TW, Lee DM, Silman AJ, *et al.* Characteristics of secondary, primary, and compensated hypogonadism in aging men: evidence from the European Male Ageing Study. *J Clin Endocrinol Metab* 2010; 95: 1810–8.
23 Ventimiglia E, Ippolito S, Capogrosso P, Pederzoli F, Cazzaniga W, *et al.* Primary, secondary and compensated hypogonadism: a novel risk stratification for infertile men. *Andrology* 2017; 5: 505–10.
24 Lipshultz L, Howards S, Niederberger C. Infertility in the Male. Cambridge: Cambridge University Press; 2009. p57–9.
25 Fisher AD, Rastrelli G, Bandini E, Corona G, Balzi D, *et al.* Metabolic and cardiovascular outcomes of fatherhood: results from a cohort of study in subjects with sexual dysfunction. *J Sex Med* 2012; 9: 2785–94.
26 Lenz S, Giwercman A, Elsborg A, Cohr KH, Jelnes JE, *et al.* Ultrasonic testicular

texture and size in 444 men from the general population: correlation to semen quality. *Eur Urol* 1993; 24: 231–8.
27 Pilatz A, Rusz A, Wagenlehner F, Weidner W, Altinkilic B. Reference values for testicular volume, epididymal head size and peak systolic velocity of the testicular artery in adult males measured by ultrasonography. *Ultraschall Med* 2013; 34: 349–54.
28 Baazeem A, Belzile E, Ciampi A, Dohle G, Jarvi K, *et al*. Varicocele and male factor infertility treatment: a new meta-analysis and review of the role of varicocele repair. *Eur Urol* 2011; 60: 796–808.
29 Hart RJ, Doherty DA, McLachlan RI, Walls ML, Keelan JA, *et al*. Testicular function in a birth cohort of young men. *Hum Reprod* 2015; 30: 2713–24.
30 Turna O, Aybar MD. Testicular stiffness in varicocele: evaluation with shear wave elastography. *Ultrasonography* 2020; 39: 350–5.
31 Salonia A, Rastrelli G, Hackett G, Seminara SB, Huhtaniemi IT, *et al*. Paediatric and adult-onset male hypogonadism. *Nat Rev Dis Primers* 2019; 5: 38.
32 Arai T, Kitahara S, Horiuchi S, Sumi S, Yoshida K. Relationship of testicular volume to semen profiles and serum hormone concentrations in infertile Japanese males. *Int J Fertil Womens Med* 1998; 43: 40–7.
33 Sakamoto H, Ogawa Y, Yoshida H. Relationship between testicular volume and testicular function: comparison of the Prader orchidometric and ultrasonographic measurements in patients with infertility. *Asian J Androl* 2008; 10: 319–24.
34 Han BH, Park SB, Seo JT, Chun YK. Usefulness of testicular volume, apparent diffusion coefficient, and normalized apparent diffusion coefficient in the MRI evaluation of infertile men with azoospermia. *AJR Am J Roentgenol* 2018; 210: 543–8.
35 Regent B, Skrobisz K, Kozak O, Matuszewski M, Studniarek M. MRI in the evaluation of the azoospermic male. *Diagn Interv Radiol* 2020; 26: 271–6.
36 Wang H, Guan J, Lin J, Zhang Z, Li S, *et al*. Diffusion-weighted and magnetization transfer imaging in testicular spermatogenic function evaluation: preliminary results. *J Magn Reson Imaging* 2018; 47: 186–90.
37 Tsili AC, Giannakis D, Sylakos A, Ntorkou A, Astrakas LG, *et al*. Apparent diffusion coefficient values of normal testis and variations with age. *Asian J Androl* 2014; 16: 493–7.



Supplementary Figure 1: Testicular volume distribution in fertile and infertile subjects with and without varicocele. *P* value according to the Mann–Whitney U test.



Supplementary Figure 2: Testicular volume distribution in infertile men according to the presence of karyotype alterations or a history of undescended testes. *P* value according to the Mann–Whitney U test.