International Urogynecology Journal (2019) 30:1575–1580
https://doi.org/10.1007/s00192-018-3802-4

ORIGINAL ARTICLE

Validation study of ultrasound bladder wall thickness measurements

Victoria Asfour[1] · Kayleigh Gibbs[1] · Ana Sophia DaSilva[1] · Ruwan Fernando[1] · Guiseppe Alessandro Digesu[1] · Vik Khullar[1]

Received: 26 July 2018 / Accepted: 19 October 2018 / Published online: 28 November 2018

Abstract

Introduction and hypothesis The aim was to validate ultrasound bladder wall thickness measurements. We scanned at three frequencies (5 MHz, 7 MHz and 9 MHz), using two techniques described in clinical practice and compared them with direct micrometre calliper measurements.

Methods Bladder dome cadaver specimens were dissected from male and female cadavers. The direct micrometre calliper measurement under direct vision was used as the gold standard. We imaged using a Voluson E8 ultrasound scanner at three frequencies, using three probes: AB27D (2-7 MHz), RAB25D (2-5 MHz) and RIC50D (5–9 MHz). The specimens were scanned on two different occasions for intra-observer variability. A second operator, measured the samples again independently for the interobserver agreement. The specimens were gently placed onto a sheathed and gelled probe to avoid deformation. The method of scanning was the same for all the specimens, probes and operators.

Results Twenty-five bladder dome specimens were assessed. The correlation of the ultrasound measurement to the direct measurement improved at higher ultrasound frequencies. Measuring from the inside of the serosal hyperechogenicity also increased the accuracy correlation with the direct measurement for all the frequencies tested.

Conclusions This is the first study validating BWT ultrasound measurements against cadaveric bladder wall calliper measurements. Technology and technique affect accuracy, which is important in clinical practice. The use of 5-MHz probes is not recommended. The most accurate measurement was obtained using high-frequency ultrasound, where the measurement did not include the serosal brightness. These data suggest that high-frequency ultrasound should be used to assess BWT.

Keywords Bladder wall thickness · Ultrasound · Accuracy · Validation

Introduction

Bladder wall thickness (BWT) is a reliable and non-invasive clinical tool that can be used to determine the likelihood of a patient having detrusor overactivity (DO) [1]. BWT can be measured using ultrasound via the transabdominal, transperineal, and transvaginal routes. If the BWT is more than 5 mm, the patient is likely to have DO [2]. There is a difference demonstrated in BWT between urodynamic stress incontinence (USI) and DO. In a study scanning with a 7.5-MHz TVS probe, the BWT in USI was a mean 4.7 mm (95% CI 4.4–5.1), and in DO, the mean BWT was 5.6 mm (95% CI 5.4–5.9). In mixed urinary incontinence (MUI), the mean BWT was 5.4 mm (95% CI 5.2–5.6) with a range that overlaps both USI and DO [2]. It is thought that in DO, repeated detrusor contractions against a closed urethra can cause muscle hypertrophy.

In a study of ambulatory urodynamics, a cut-off at 6 mm was found to be highly suggestive of detrusor overactivity in the absence of stress incontinence, in women with equivocal (or normal) urodynamics [3]. The use of ultrasound has been proposed as a tool to reduce the need for urodynamic investigations in the diagnosis of the cause of urinary incontinence [4].

Bladder wall thickness has been shown to change in patients treated with anti-cholinergic medication [5, 6]. A double-blind placebo-controlled trial showed that treatment with Tolterodine and Solifenacin for women with overactive bladder syndrome resulted in a significant decrease in BWT after 12 weeks of therapy.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00192-018-3802-4) contains supplementary material, which is available to authorized users

✉ Victoria Asfour
Mdb99va@doctors.org.uk

1 St Mary's Hospital, Imperial College Healthcare NHS Trust, Paddington W2 1NY, UK

Studies have used a variety of probes to assess BWT over the years, ranging from 2 MHz to 8 MHz, with some studies not even stating the frequency of probe that was used [7]. It is possible that the frequency of the probe used might affect the measurements of the BWT.

We identified three probes, scanning at three different frequencies, that are in use in routine clinical practice. The RAB25D, is a wide band convex volume, curved array transducer that scans at a maximum bandwidth of 5 MHz (range 1–5 MHz); 98° wide angle. The RAB25D is a 3D probe used for trans-abdominal 3D ultrasound in obstetrics. In urogynaecology, it is used for trans-perineal ultrasound for the assessment of the pelvic floor, including levator ani injuries. The AB27D is a wide band convex transducer, that scans at a mean of 7 MHz and a maximum of 8 MHz (range 2–8 MHz), 107° wide angle. It is a probe in common use in trans-abdominal ultrasound. The RIC50D is a wide band convex volume, endocavity transducer that scans at a maximum of 9 MHz (range 4–9 MHz), 179° wide angle. The RIC50D is the most commonly used probe for transvaginal scanning in general gynaecology and early pregnancy.

There have been two methods of measuring BWT described in clinical practice. In urogynaecology, we measure the full thickness of the bladder wall, including the serosal brightness, which is often seen, particularly at the dome. In urology, papers describe measurement of the detrusor muscle thickness only [8].

The aim was to investigate the most accurate method of scanning the BWT. We aim to validate the BWT measurements by scanning at three frequencies and using two methods of measurement. All specimens were scanned with three probes at 5 MHz, 7 MHz and 9 MHz. Each measurement was taken using two methods of measurement: including and not including the serosal echo of the bladder. The measurements obtained were compared with the micrometre calliper measurements taken under direct vision. The serosa of the bladder is very thin and appears constant in appearance visually across all bladders; therefore, one full thickness measurement of the bladder was obtained with the callipers.

Materials and methods

Male and female bladders were obtained from cadavers. The individuals had consented to their body to be used for teaching and research, in accordance with the Human Tissue Act of 2004. They were fixed using a modified solution of 7% phenol, 7% formalin, 25% isopropyl alcohol and 61% water. The

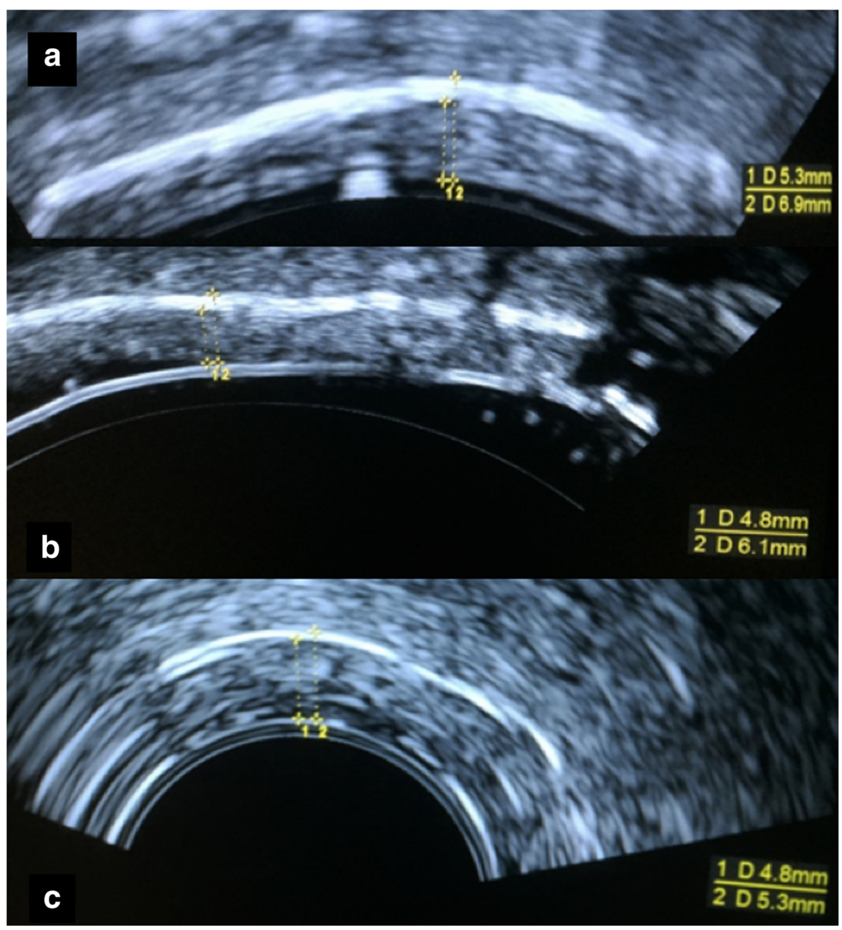

Fig. 1 Bladder wall thickness ultrasound images taken on the same specimen. **a** 7 MHz (AB27D); **b** 5 MHz (RAB25D); **c** 9 MHz (RIC50D)

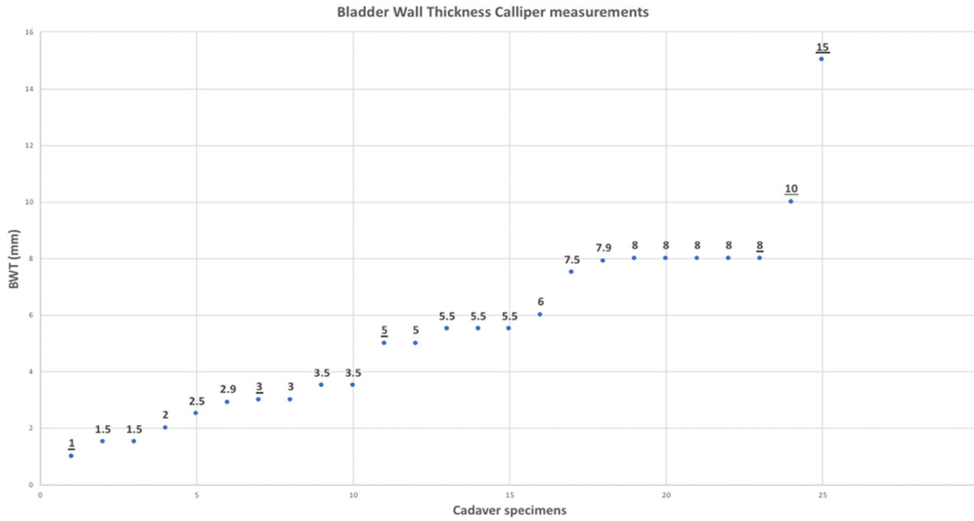

Fig. 2 Graph demonstrating the range of bladder wall thickness encountered in this study. Male specimens bladder wall thickness measurements are *underlined*

bladder was visualised and examined grossly through the open abdominal wall. Four-by-four-centimetre specimens were dissected from the dome of the bladder using a sharp surgical blade. The harvested bladder dome specimens were measured using a dial micrometre calliper (Mono-block Vernier Callipers) for bladder wall thickness and were stored for later ultrasound evaluation. The micrometre calliper measurement was taken to be the gold standard.

The harvested bladder dome specimens were imaged using GE Voluson E8 utilising three different ultrasound probes: RAB25D (5 MHz), AB27D (7 MHz) and RIC50D (9 MHz). The probe was gelled and sheathed. A further layer of gel was applied on top of the sheath. The bladder specimens were gently placed on top of the probe to avoid deformation. Another thick layer of gel was applied on top of the bladder wall.

The ultrasound was repeated until a good-quality image was obtained. Ultrasound measurements of bladder wall thickness were obtained using two different methods for each probe. The first method measured to the inside of the serosal brightness and the second to the outside of the serosal brightness (Fig. 1).

Bladder specimens were scanned blindly on two different occasions by the author (VA). The second operators (AD) or (KG), were trained in measuring the BWT according to the protocol and measured the samples again, blinded to the previous measurements. When the BWT varied in the image obtained, the thickest part of the specimen was scanned. The thickest BWT on the best-quality image was recorded.

The two separate investigators were blinded to each other's assessment. Measurements of agreement were calculated using Spearman's correlation test for comparison of the probes with each other. Bland–Altman plots and linear regression was carried out to assess inter-observer and intra-observer variability. A significance level of $p < 0.05$ was used to assess correlations.

In Fig. 1, the same bladder specimen was scanned with three different probes. The method of measurement is demonstrated in Fig. 1. In Fig. 1a measurements were taken with the 5-MHz probe. The first measurement was taken inside the white line delineating the serosal bladder surface. The measurement inside the white line was 5.3 mm and the measurement starting from the outside of the line was 6.9 mm. Figure 1b shows the same measurement performed with the 7 MHz probe. The measurement inside of the white line was 4.8 mm and the measurement starting from the outside of the line was 6.1 mm. Figure 1c shows the same measurement taken with 9 MHz, where the measurement on the inside of the white line is 4.8 mm and the measurement from the outside was 5.3 mm.

A power calculation for sample size was performed for our population using the alpha as 0.05 with a 95% confidence interval. The number required to demonstrate a significant difference is a sample size of 16.

Results

Twenty-five bladder specimens were obtained from 19 female and 6 male cadavers.

Table 1 Comparing bladder wall thickness as measured by different ultrasound probes, and calliper micrometre measurements under direct vision

Calliper	5 MHz		7 MHz		9 MHz	
	Out	In	Out	In	Out	In
ρ value	0.606	0.728	0.760	0.817	0.801	0.867
p value	< 0.001	0.000	0.000	0.000	0.000	0.000

Spearman's rho (ρ) correlation coefficient and significance level (*p* value)

Measurements were taken including the serosal echo (out), and not including the serosal echo (in)

Table 2 Intra–observer correlation of bladder wall thickness measurements

	Mean	Standard deviation	Mean difference	Standard deviation of difference	95% CI of difference	Linear regression significance p value
Calliper	5.5	3.24	−0.88	1.36	−1.44, −0.32	0.7
RAB(3D) Out	8.6	2.78	0.16	1.01	−0.25, 0.58	0.9
RAB(3D) In	5.9	2.44	0.09	0.83	−0.26, 0.43	0.6
AB(2D) Out	7.6	2.56	−0.19	0.47	−0.38, 0.01	0.08
AB(2D) In	5.4	2.27	−0.88	1.35	−1.44, −0.32	0.4
RIC (TVS) Out	6.2	2.3	−0.92	0.38	−0.25, 0.07	0.5
RIC (TVS) In	5.3	2.1	−4.26	14.64	−10.3, 1.78	0.5

The cadavers were all elderly, with a mean age of 86 years (median 86 years; standard deviation 9.3; range 70 to 102 years) at time of death. Cause of death included old age, stroke, ischaemic heart disease, heart failure, pharyngeal cancer, pneumonia, cholangiocarcinoma and Ewing's sarcoma.

The 5-MHz probe gave a less sharp image overall. The 7-MHz probe gave subjectively better-quality images for analysis. The 9-MHz probe gave the thinnest and sharpest bladder outline, as seen in Fig. 1. The BWT varied widely between cadavers, with the male cadavers having the thickest bladders, possibly reflecting underlying bladder outlet obstruction. Mean BWT was 5.4 mm (range 1–15 mm, median 5.5 mm, mode 8 mm; See Fig. 2).

The calliper measurements correlated most closely with the measurements on the inside of the white serosal echo. The measurements including the echo were always wider than the calliper measurements. The difference in width between the in/out measurements decreased proportionately as the frequency of the probe increased (see Supplementary Table 6).

Spearman's rho (ρ) correlation coefficient was used to assess the agreement correlation among the three probes. The agreement for the 5-MHz probe with the calliper measurement was moderate to poor ($\rho = 0.60$–0.72, $p < 0.001$). The agreement for the 7-MHz probe with the calliper measurement was moderate ($\rho = 0.76$–0.81, $p < 0.0001$). The agreement for the 9-MHz probe with the calliper measurement was good ($\rho = 0.80$–0.86, $p < 0.0001$; Table 1).

The Spearman's rho agreement of the 7-MHz probe with the 9-MHz probe was good ($\rho = 0.78$–0.80, $p = 0.000$). The Spearman's rho agreement of the 5-MHz probe with the 7-MHz probe was moderate ($\rho = 0.60$–0.72, $p < 0.001$). The Spearman's rho agreement of the 5-MHz probe with the 9-MHz probe was poor (($\rho = 0.57$–0.69, $p < 0.001$).

Bland–Altman plots were performed to assess the inter-observer and intra-observer variability, and the linear regression significance test did not reveal any significant difference in measurements (Tables 2, 3). There was no significant difference between the values obtained when measuring inside and outside the serosal white line reflection (Table 4, Supplementary Table 5).

We observed that the 9-MHz (TVS) probe allowed visualisation of trabeculations that were seen on observation of the bladder specimens. Trabeculations could not be seen when scanning at lower frequencies (7 MHz, 5 MHz). Thickness varied owing to the presence of trabeculations, which led to a few outliers in our data. Interestingly, the thickest specimens had less variation in thickness. It was the average-sized specimens that had the widest variation of thickness at the sites of trabeculations and in between the trabeculations.

Table 3 Inter-observer correlation of bladder wall thickness measurements. Mean and standard deviation of AD/KG measurements

	Mean	Standard deviation	Mean Difference mean	Standard deviation of difference	95% CI of difference	Linear regression significance p value
Calliper	6.3	3.2	−0.88	1.36	−1.4, −0.3	0.8
RAB (3D) out	8.3	2.7	0.3	1.31	−0.24, 0.84	0.8
RAB (3D) in	5.9	2.4	0.08	1.05	−0.36, 0.51	0.7
AB(2D) out	7.7	2.3	−0.03	0.93	−0.41, 0.36	0.1
AB (2D) in	5.7	2.1	−0.26	0.75	−0.57, 0.05	0.3
RIC (TVS) out	6.4	2.4	−0.22	0.42	−0.39, −0.05	0.4
RIC (TVS) in	5.5	2.4	−0.14	0.56	−0.38, 0.09	0.07

Table 4 Bladder wall thickness measurements inside and outside the peritoneal line

	Mean difference	Standard deviation of difference	95% CI of difference	Linear regression significance *p* value
RAB (3D) out/in	2.66	1.12	2.1, 3.1	0.1
AB (2D) out/in	2.16	1.08	1.7, 2.6	0.2
RIC (TVS) out/in	0.87	0.87	0.6, 1.1	0.3

Discussion

This is the first study validating the effect of technology and technique in the assessment of BWT. The cadaver bladder walls evaluated varied widely in size, with calliper measurements ranging from 1 mm to 15 mm. There were different correlations across the range of probes from 5- to 9-MHz frequencies. When compared with the 7-MHz and 9-MHz probes, the agreement for the 5-MHz probe is poor. The higher frequency probes had a better correlation with the calliper measurements. The correlation of the ultrasound measurements was directly proportional with the probe frequency.

The serosal surface caused a bright echo with the curved array (5-MHz, 7-MHz) probes, but less so with the transvaginal probe. The TVS (9-MHz) probe was also higher frequency and endo-array technology, delineating more tissue detail. For example, it is possible to observe trabeculations with a 9-MHz probe with better spatial resolution, that are not seen with the 5-MHz or the 7-MHz probes. This could be because the transvaginal probe is higher frequency (9 MHz), or because of the technology of the array. Urology papers scanning BWT in men and children, abdominally, specify that they measure the detrusor muscle without this echo [8]. The 5-MHz measurement taken from the outside of the serosal brightness yielded the thickest measurements.

In women, when using transperineal scanning with curved linear probes, the echo can vary in brightness and degree of prominence. It is seen mostly on the dome of the bladder and less commonly on the trigone. This effect is less evident when scanning with the transvaginal probe. It would be useful if future studies on BWT could comment on the frequency of the probe used to scan, the type of probe technology, and the method used to assess the BWT.

The strength of this study is the ability to measure the same bladder specimen directly with a micrometre calliper in its full thickness and at three different ultrasound frequencies. To our knowledge, this has never been done before. We also validated the technique of scanning bladder wall thickness. We evaluated and proved that high-frequency ultrasound technology is of additional benefit in the accurate measurement of BWT. There is only a 0.9-mm difference between the mean thickness for DO and USI, which makes the accuracy of measurements extremely important [2].

For example, a recent study investigating BWT and urodynamic diagnosis on 125 patients found BWT to be a predictor of response to treatment [9]. The study was done with a 5-MHz TVS probe. Their data did not reach statistical significance for demonstrating the difference in BWT between the diagnostic groups of SUI, DO and MUI [9]. Our data suggest that if they had used a higher frequency ultrasound, utilising the technique of not including the serosal echo, their data would have been clearer in showing a difference between the groups.

A limitation of this study is that the bladders were scanned ex-vivo (not in their anatomical position). In clinical scanning, images may not be as clear owing to interposing tissues. Death and tissue fixation with formalin causes changes in the elasticity of soft tissues. The tensile strength and elasticity were not studied. It is possible that the cadaveric bladder echogenicity or BWT may have changed; however, this model provided good-quality images for the study of accuracy between the measurement methods.

Further work could investigate the difference in BWT measurements using the methods outlined in this paper for a clinical study. We need to delineate if these differences also apply in a clinical setting.

Conclusion

This study has proven that technology and technique are important in assessing BWT accurately. Scanning with a 5-MHz probe is not recommended. The best correlation with the gold standard measurement (direct measurement of the cadaveric tissue) was with the 9-MHz probe, measuring from the inside of the hyperechogenic serosal echo of the bladder. High-frequency ultrasound is more accurate in assessing BWT.

Compliance with ethical standards

Conflicts of interest None.

References

1. Khullar V, Salvatore L, Cardozo L, Bourne TH, Abbott D, Kelleher C. A novel technique for measuring bladder wall thickness in women using transvaginal ultrasound. Ultrasound Obstet Gynecol. 1994;4:220–3.
2. Panayi DC, Tekkis P, Fernando R, Hendricken C, Khullar V. Ultrasound measurement of bladder wall thickness is associated with the overactive bladder syndrome. Neurourol Urodyn. 2010;29:1295–8.
3. Robinson D, Anders K, Cardozo L, Bidmead J, Toozs-Hobson P, Khullar V. Can ultrasound replace ambulatory urodynamics when investigating women with irritative urinary symptoms? BJOG. 2002;109:145–8.

4. Latthe P, Middleton L, Rachaneni S, McCooty S, Daniels J, Coomarasamy A, et al. Ultrasound bladder wall thickness and detrusor overactivity: a multicentre test accuracy study. BJOG. 2017;124:1422–9.
5. Bray R, Cartwright R, Cardozo L, Hill S, Guan Z, Khullar V. Tolterodine ER reduced increased bladder wall thickness in women with overactive bladder. A randomised, placebo-controlled, double-blind, parallel group study. Neurourol Urodyn. 2018;37:237–43.
6. Robinson D, Oelke M, Khullar V, Wijkstra H, Tretter R, Stow B, et al. Bladder wall thickness in women with symptoms of overactive bladder and detrusor overactivity: results from the randomised, placebo-controlled shrink study. Neurourol Urodyn. 2016;35:819–25.
7. Oelke M, Khullar V, Wijkstra H. Review on ultrasound measurement of bladder or detrusor wall thickness in women: techniques, diagnostic utility, and use in clinical trials. World J Urol. 2013;31:1093–104.
8. Muller L, Bergstrom T, Hellstrom M, Svensson E, Jacobsson B. Standardized ultrasound method for assessing detrusor muscle thickness in children. J Urol. 2000;164:134–8.
9. Akselim B, Doganay M, Ozcan N, Akselim S, Cavkaytar S. Correlation of bladder wall thickness and treatment success in types of urinary incontinence. Int Urogynecol J. 2017;28:417–22.