Prospective evaluation of interobserver diagnostic agreement for focal urinary bladder wall abnormalities detected by ultrasound

G. Francica ^{a,*}, F. Scarano ^a, S.A. Bellini ^b, A. Miragliuolo ^b

^a Diagnostic and Interventional Ultrasound Unit, Camilliani Hospital "S. Maria della Pietà", Casoria, Naples, Italy

^b Urology Department, Camilliani Hospital "S. Maria della Pietà", Casoria, Naples, Italy

KEYWORDS

Urinary bladder;
Ultrasound;
Bladder cancer;
Interobserver
variability.

Abstract *Introduction:* The interobserver variability of the sonographic assessment of Focal Bladder Wall Abnormalities (FBWA) between two physicians with different levels of experience was evaluated prospectively.

Materials and methods: In the same session the two operators examined independently the urinary bladder of 87 consecutive patients (Mean Age 68 yrs; range 33–80; 75 Males; 15 Females) who underwent cystoscopy within 1–2 days. Before beginning the study, the two observers standardized the process for US exam. Weighted kappa statistics were used to determine the degree of agreement as to the presence, size, location and number of FBWA detected at cystoscopy.

Results: the highest level of agreement ($k = 1$) was reached on the presence of FBWA, whereas a substantial agreement was reached on size ($k = 0.78$), number ($k = 0.72$) and location ($k = 0.62$) of FBWA.

Conclusions: In this study the gap in US experience between two operators did not emerge in assessing FBWA, provided a standardized technical approach is pursued.

Sommario *Introduzione:* È stata valutata prospetticamente la variabilità interosservatore relativamente alla diagnosi di alterazioni focali della parete vescicale tra due operatori con diverso livello di esperienza.

Materiali e metodi: Due operatori con diverso livello di esperienza in ecografia hanno esaminato indipendentemente nella stessa sessione 87 pazienti consecutivi (età media 68 anni; range 33–80; 75 maschi; 15 donne) sottoposti successivamente a cistoscopia entro 1–2 giorni. Prima dell'inizio dello studio la metodologia di studio è stata standardizzata. Il grado di accordo sulla presenza, sede, dimensioni e numero delle alterazioni parietali della vescica è stato calcolato con il metodo della Kw.

Risultati: Il grado più elevato di accordo ($k = 1$) è stato raggiunto relativamente alla presenza delle alterazioni parietali, mentre un accordo sostanziale è stato ottenuto per le dimensioni ($k = 0.78$), numero ($k = 0.72$) e sede ($k = 0.62$).

* Corresponding author. Servizio di Ecografia Diagnostica ed Interventistica, Presidio Ospedaliero Camilliani "S. Maria della Pietà", via S. Rocco 9, 80026 Casoria (NA), Italy.

E-mail address: giampierofrancica@tin.it (G. Francica).

Conclusioni: In questo studio la differenza di esperienza tra due operatori non è emersa nella diagnostica delle alterazioni parietali della vescica anche in ragione di una metodologia standardizzata di studio.

© 2010 Elsevier Srl. All rights reserved.

Introduction

Urinary bladder Ultrasound (US) has proved to be useful in a variety of clinical settings, as US has been used to establish postvoid bladder volume [1], as an aid for avoiding unnecessary catheterization and for diagnosing outlet obstruction by measuring bladder wall thickness and bladder mass [2]. However US diagnostic performance for focal bladder wall abnormalities (FBWA), particularly urothelial cancer, has been challenged because several factors (e.g. size, location and macroscopic aspect of FBWA) may hamper US diagnostic capabilities [3–12]. In addition, intrinsic characteristics of the sonographic technique, i.e. dependence on both patients' body habitus and the operator's skills, may be considered as adjunctive elements affecting reproducibility of US. Generally, clinicians and radiologists do not regard US as a valuable diagnostic tool for FBWA as other imaging techniques such as both excretory CT or MR [13,14] and virtual CT or MR cystoscopy [15,16], especially for evaluation of patients with hematuria. We have recently demonstrated that US performed with newer machines and under optimal technical conditions provides a valuable imaging tool for demonstrating focal alterations of the urinary bladder wall, considering cystoscopy and histology as reference standards [17]. However, in this study only one operator carried out US exams, a condition that cannot be reproduced in everyday practice where exams are performed by different radiologists or sonographers. The purpose of the present study was to evaluate the interobserver variability of sonographic assessment of FBWA according to the levels of experience of two operators in a population of patients undergoing cystoscopy.

Materials and methods

Between November 2007 and January 2008, two physicians with different levels of experience in abdominopelvic sonography (operator A: 20 years; operator B: 5 years) performed a US exam of the urinary bladder in all patients consecutively referred to our Urology Department for cystoscopic examination.

87 consecutive patients entered this study (Mean Age 68 yrs; range 33–80; 72 Males; 15 Females) who had been scheduled for cystoscopy on account of hematuria (43 cases [49.5%], including 32 with gross hematuria and 11 with microscopic hematuria associated with lower urinary tract symptoms, surveillance after TUR (33 cases [38%], including 18 with gross hematuria), and incidental findings during sonographic examinations of the abdomen and pelvis performed at other Institutions (11 cases [12.5%], including 7 with lower urinary tract symptoms).

Study design

For each patient, the two observers, blind to clinical and previous US findings, carried out US examinations independently in the same session using the same US

machine available at the US unit (LOGIQ 9 GE Healthcare, Milwaukee, WI) equipped with tissue harmonic (THI) and spatial compound imaging (SCI) and a wide-band electronic convex probe (2–4 MHz). For each session, the second observer was outside the room while US was performed by the first observer and was therefore unaware of the first observer's results. The order of operators was randomly assigned. All the patients were scanned 1–2 days before cystoscopy. Sonography of the upper urinary tract was performed as well to exclude other causes of hematuria or transitional tumor extension. To prevent a possible bias caused by a different technical modality of performing US examination, the process for US exam was standardized as follows:

- a) *Technical modality of US exam:* an optimal filling of the urinary bladder was considered between 200 and 250 mL as measured automatically on 2 orthogonal images directly on the screen. B-mode parameters were adjusted to obtain a quality image as high as possible according to the patient's body habitus (Fig. 1). All examinations were started by coupling THI and SCI, but each operator could vary the protocol by using the 2 sonographic algorithms separately and in different sequences during the examination for each individual patient to obtain the most informative images. Similarly, the dynamic range, total gain, and time-gain compensation were set at low values to obtain the best image in each patient. Transverse, longitudinal and oblique scans were obtained for each patient, paying special care to areas of the bladder that are usually difficult to explore (the anterior wall and the dome); change in the patient's decubitus was obtained to assess mobility of parietal lesions especially if partially or entirely calcified.
- b) *Definition of Focal Bladder Wall Abnormalities (FBWA):* this parameter was defined as an echogenic (solid) alteration of the normal sonographic bladder wall anatomy, not mobile with decubitus changes, including one of the following morphologic aspects: (1) polypoid (height greater than width) (Fig. 2); (2) sessile (width greater than height) (Fig. 3), (3) infiltrating (obvious interruption of the bladder wall) (Fig. 4), and (4) mixed (a combination of either 1 and 3 or 2 and 3). For each patient, number (1, 2, 3, or >3), size (widest diameter, <5, 5–15, >15–25, and >25 mm; in cases of multiple lesions, the size of the largest FBWA was reported) and location (base, dome, right or left lateral wall, or anterior or posterior wall) of FBWA were recorded. Although Color Doppler interrogation was carried out whenever a FBWA was detected, we decided to discard this parameter in the final evaluation due to the disappointing results in terms of diagnostic effectiveness reported by other Authors [18,19] and by ourselves [17]. In addition, diagnostic pitfalls in bladder US were discussed: a wide spectrum of differential diagnosis

Fig. 1 High quality US transverse (a) and longitudinal (b) images of a well-distended, normal urinary bladder.

Fig. 2 Two small "polypoid" vegetations (white arrows) are demonstrated at the bladder base in a transverse US scan.

Fig. 3 A sessile, non infiltrating urothelial lesion (between white arrows) is appreciated on the right lateral wall.

was considered, including blood clots, coexisting columnar hypertrophy of the bladder wall in the setting of benign prostatic hypertrophy (Figs. 5 and 6), calculi (Fig. 7).

For each patient, sonographic parameters were reported to a personal computer database by each operator at the end of the sonographic examination.

Cystoscopy was carried out using a rigid cystoscope with a field of view of 30°–70° to 120° (Karl Storz, GmbH, Tuttlingen, Germany) under general or regional anesthesia. The number, location, and size of the tumors were individually determined and then entered into a database after cystoscopy for later comparison with the sonographic findings. All cystoscopic abnormalities were removed with a resectoscope. Multiple biopsy specimens were obtained with cold-cup biopsy forceps from areas suggestive of

Fig. 4 A deep neoplastic ulcer (*) of the posterior bladder wall is demonstrated in a longitudinal US scan; the echogenic parietal line is clearly disrupted and a neoplastic digitation invades peri-vesical tissue (black arrowhead).

Fig. 5 Multiple neoplastic polypoid lesions (white arrowheads) appear from posterior wall and base (longitudinal scan).

carcinoma in situ and from random sites throughout the bladder in high risk cases (e.g. previous TUR or previous high-grade carcinoma). Pathologic diagnosis was available for all lesions according to the 2004 World Health Organization/International Society of Urological Pathology classification for urothelial Diseases [20].

Statistics

Weighted kappa statistics were used to determine the degree of agreement after correction for the agreement expected by chance. The kappa statistic was interpreted as follows: less than 0.00, poor agreement; 0.00–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and 0.81–1.00, almost perfect agreement [21]. The level of statistically significant difference was $p < 0.05$.

Fig. 6 Multiple pseudo-polyps (white arrows) in the setting of columnar hypertrophy of the bladder wall due to benign prostatic hypertrophy.

Fig. 7 Similar images (a hyperechoic focus with faint posterior shadowing and twinkling artifact at the left uretral meatus) are obtained in case of a stone (a) and an urothelial cancer (b). Calculus was recognized on the basis of free movements seen in real-time during decubitus' changes.

This study was approved by the institutional review board of our institution, and informed consent was obtained from all patients.

Results

Cystoscopy displayed one or more FBWA in 67 patients in whom malignancy was diagnosed in 62 cases (transitional tumors in 60 patients, metastatic tumors in 2 cases) the remaining 5 cases being chronic inflammatory changes. In five out of 20 patients with negative cystoscopy transitional tumors were diagnosed on biopsy taken randomly (four of whom were on follow up for previously resected urothelial cancer).

An optimal bladder filling was obtained in 83 patients (95.4%). Of the detected FBWA 86% were within 25 mm in diameter and 89.5% presented as polypoid lesions.

Sensitivity, specificity and overall accuracy values for cystoscopic FBWA of operator A were 91%, 85% and 89.6%, which resulted not statistically different ($>p 0.05$) from those of operator B (91%, 80%, 88.5%). When only bladder cancer was considered, both operators detected 93.5% of the biopsy-proved cancers. As far as missed cancers are concerned, both observers failed to diagnose 4 cancers but they missed the same lesions only in two cases. False positive cases were due to impacted calculi in an edematous uretral meatus in 2 patients (both operators missed the correct diagnosis in the same patient whereas only the less experienced observer failed to make one more correct diagnosis),

Table 1 Interobserver agreement for diagnosis of Focal Bladder Wall Abnormalities (FBWA) parameters.

FBWA	K_w	95% CI
Presence	1	0.88–1
Size	0.78	0.78–0.90
Number	0.70	0.65–0.80
Location	0.62	0.59–0.78

pseudo-polypoid columnar hypertrophy in 2 cases (both operators overdiagnosed a FBWA).

In Table 1 levels of agreement for selected items (see methods) are reported. The highest level of agreement ($k = 1$) was reached on the presence of FBWA whereas a substantial agreement was reached on size ($k = 0.78$), number ($k = 0.72$) and location ($k = 0.62$) of FBWA.

Conclusions

Diagnostic capabilities of US for FBWA are generally considered poor especially in terms of sensitivity which is reported to range between 50% and 94% [3–12]. It is conceivable that both tumor-related factors (e.g. size, location, macroscopic aspect) and operator-dependent factors (i.e. low reproducibility due to different levels of experience and the technical modality of performing US exam) may explain such a wide variation in results.

In addition, another variable affecting diagnostic performance of the sonographic technique to be taken into account is the local policy of performing US exam. In fact, by comparing the diagnostic accuracy of multidetector computed tomography urography (CTU) and US for diagnosing bladder cancer Knox et al. [22] showed that sensitivity was substantially higher for CTU (89.7%) than US (69%). However, these Authors acknowledged that the US sensitivity of their study might have been increased if US examinations had been performed and interpreted by a radiologist with subspecialty experience in US instead of being carried out by sonographers with subsequent reporting by the radiologist.

We have recently reported good US diagnostic performance for FBWA detected at cystoscopy [17]: sensitivity, specificity, and overall accuracy of US for cystoscopic lesions were 91.4%, 79.3%, 88.2%. However, in that study only a single operator performed US exams, a condition difficult to reproduce in routine practice where patients may be evaluated by different physicians or sonographers.

In the present prospective study, we have investigated the degree of interobserver agreement for the US diagnosis of FBWA having as reference standard cystoscopy results. To the best of our knowledge, such an approach has not been reported in the literature. Our results show that both operators exhibited high sensitivity, specificity and overall diagnostic accuracy in detecting FBWA demonstrated at cystoscopy (with sensitivity as high as 93.5% when only bladder cancers were considered). A high level of agreement was reached on the presence of FBWA and a substantial agreement between observers was obtained for size, location and number of BWA. Indeed, the gap in US experience between two operators did not emerge in assessing the urinary bladder for FBWA.

In our opinion, the low rate of inconsistencies at the interpretation of bladder scans for FBWA which we encountered in our study is likely to depend on several factors, mainly related to the study design itself: use of a state-of-art US machine equipped with new algorithms of image acquisition (e.g. THI and SCI) which substantially improve image quality and diagnostic yield in comparison to conventional US [23,24]; a consensus obtained between the two operators as to technical modalities of carrying out US examination of the urinary bladder; preliminary discussion of most frequently encountered pitfalls in bladder US diagnostics. On clinical grounds, this implies that well-trained and motivated operators independently of their experience may achieve high diagnostic capabilities in identifying FBWA. However, the main limit of our study should be mentioned: the high pre-test prevalence rate of FBWA in our selected population implies that test sensitivity is overestimated and not entirely applicable to an open population.

Although CT urography is nowadays considered as the most appropriate [25] and the most clinically effective imaging method in the evaluation of patients with hematuria [13,14,26], yield and clinical effectiveness of US should not be disregarded in the evaluation of the urinary bladder. Our experience indicated that US has a good sensitivity and overall accuracy for cystoscopy-proved FBWA even when different but dedicated operators perform sonographic exploration of the bladder. A careful US examination using state-of-art machines as initial imaging test could avoid cystoscopy without resorting to more expensive, cumbersome, and potentially hazardous techniques (both for exposure to ionizing radiations and use of contrast media) at least in some patients, particularly those with microscopic hematuria thought not to be at high risk for bladder cancer [27] and those suffering from small, recurrent, papillary bladder tumors (Ta) in whom a watchful waiting policy has been recently proposed because they pose minimal risk for the patient [28].

In addition, further improvements in diagnostic US capabilities for bladder cancer might be expected from introduction of new technologies such as 3D reconstruction [29] and contrast-enhanced US [30].

Larger studies addressing these specific issues might confirm US utility in the diagnostic algorithm for FBWA.

Conflict of interest statement

The authors have no conflict of interest.

Supplementary material

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.jus.2010.11.002](https://doi.org/10.1016/j.jus.2010.11.002).

References

- [1] Simforoosh N, Dadkhah F, Hosseini SY, Asgari MA, Nasser A, Safarinejad MR. Accuracy of residual urine measurement in

- men: comparison between real-time ultrasonography and catheterization. *J Urol* 1997;158:59–61.
- [2] Kelly CE. Evaluation of voiding dysfunction and measurement of bladder volume. *Rev Urol* 2004;6:S32–7.
 - [3] Itzchak Y, Singer D, Fischelovitch Y. Ultrasonographic assessment of bladder-tumors. 1. Tumor detection. *J Urol* 1981;126:31–3.
 - [4] Abu-Yousef MM, Narayana AS, Franken Jr EA, Brown RC. Urinary bladder tumors studied by cystosonography. Part I: detection. *Radiology* 1984;153:223–6.
 - [5] Brun B, Gammelgaard J, Christoffersen J. Transabdominal dynamic ultrasonography in detection of bladder tumors. *J Urol* 1984;132:19–20.
 - [6] Juul N, Torppedersen S, Larsen S, Rasmussen F, Holm HH. Bladder tumor control by abdominal ultrasound and urine cytology. *Scand J Urol Nephrol* 1986;20:275–8.
 - [7] Vallancien G, Veillon B, Charton M, Brisset JM. Can transabdominal ultrasonography of the bladder replace cystoscopy in the follow up of superficial bladder tumors? *J Urol* 1986;136:32–4.
 - [8] Malone PR, Weston-Underwood J, Aron PM, Wilkinson KW, Joseph AE, Riddle PR. The use of transabdominal ultrasound in the detection of early bladder tumors. *Br J Urol* 1986;58:520–2.
 - [9] Davies AH, Cranston D, Meagher T, Fellows GJ. Detection of recurrent bladder tumours by transrectal and abdominal ultrasound compared with cystoscopy. *Br J Urol* 1989;64:409–11.
 - [10] Vera-Donoso CD, Llopis B, Oliver F, Server G, Alonso M, Jimenez-Cruz JF. Follow-up of superficial bladder cancer: how to spare cystoscopies? *Eur Urol* 1990;17:17–9.
 - [11] Cariou G, Maaraoui N, Cortesse A. Can the combination of bladder ultrasonography and urinary cytodiagnosis replace cystoscopy in the diagnosis and follow-up of tumors of the bladder? *Prog Urol* 1997;7:51–5.
 - [12] Datta SN, Allen GM, Evans R, Vaughton KC, Lucas MG. Urinary tract ultrasonography in the evaluation of haematuria—a report of over 1000 cases. *Ann R Coll Surg Engl* 2002;84:203–5.
 - [13] Townsend B, Silverman S, Morteale K, Tuncali K, Cohan R. Current practice of CT urography by urologists: a survey of the society of urology. *J Comput Assist Tomogr* 2009;33:96–100.
 - [14] Silverman SG, Leyendecker JR, Amis Jr ES. What is the current role of CT urography and MR urography in the evaluation of the urinary tract. *Radiology* 2009;250:309–23.
 - [15] Lammle M, Beer A, Settles M, Hannig C, Schwaibold H, Drews C. Reliability of MR imaging-based virtual cystoscopy in the diagnosis of cancer of the urinary bladder. *AJR Am J Roentgenol* 2002;178:1483–8.
 - [16] Kim JK, Ahn JH, Park T, Ahn HJ, Kim CS, Cho KS. Virtual cystoscopy of the contrast material-filled bladder in patients with gross hematuria. *AJR Am J Roentgenol* 2002;179:763–8.
 - [17] Francica G, Bellini SA, Scarano F, Miragliuolo A, De Marino FA, Maniscalco M. Correlation of transabdominal sonographic and cystoscopic findings in the diagnosis of focal abnormalities of the urinary bladder wall: a prospective study. *J Ultrasound Med* 2008;27:887–94.
 - [18] Horstman WG, McFarland RM, Gorman JD. Color Doppler sonographic findings in patients with transitional cell carcinoma of the bladder and renal pelvis. *J Ultrasound Med* 1995;14:129–33.
 - [19] Karahan OI, Yikilmaz A, Ekmekcioglu O, Ozturk F, Sevinc H. Color Doppler ultrasonography findings of bladder tumors: correlation with stage and histopathologic grade. *Acta Radiol* 2004;45:481–6.
 - [20] Oosterhuis JW, Schapers RF, Janssen-Heijnen ML, Pauwels RP, Newling DW, Kate F. Histological grading of papillary urothelial carcinoma of the bladder: prognostic value of the 1998 WHO/ISUP classification system and comparison with conventional grading systems. *J Clin Pathol* 2002;55:900–5.
 - [21] Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
 - [22] Knox MK, Cowan NC, Rivers-Bowerman MD, Turney BW. Evaluation of multidetector computed tomography urography and ultrasonography for diagnosing bladder cancer. *Clin Radiol* 2008;63:1317–25.
 - [23] Yücel C, Özdemir H, Asık E, Öner Y, Isık S. Benefits of tissue harmonic imaging in evaluation of abdominal and pelvic lesions. *Abdom Imaging* 2003;28:103–9.
 - [24] Oktar SO, Yucel C, Ozdemir H, Uluturk A, Isik S. Comparison of conventional sonography, real-time compound sonography, tissue harmonic sonography, and tissue harmonic compound sonography of abdominal and pelvic lesions. *AJR Am J Roentgenol* 2003;181:1341–7.
 - [25] Choyke PL. Radiologic evaluation of hematuria: guidelines from the American college of radiology's appropriateness criteria. *Am Fam Physician* 2008;78:347–52.
 - [26] Joffe SA, Servaes S, Okon S, Horowitz M. Multi-detector row CT urography in the evaluation of hematuria. *RadioGraphics* 2003;23:1441–55.
 - [27] Grossfeld GD, Litwin MS, Wolf Jr JS, Hricak H, Shuler CL, Agerter DC, et al. Evaluation of asymptomatic microscopic hematuria in adults: the American urological association best practice policy. II. Patient evaluation, cytology, voided markers, imaging, cystoscopy, nephrology evaluation, and follow-up. *Urology* 2001;57:604–10.
 - [28] Soloway MS, Bruck DS, Kim SS. Expectant management of small, recurrent, noninvasive papillary bladder tumors. *J Urol* 2003;170:438–41.
 - [29] Kocakoc E, Kiris A, Orhan I, Poyraz AK, Artas H, Firdolas F. Detection of bladder tumors with 3-dimensional sonography and virtual sonographic cystoscopy. *J Ultrasound Med* 2008 Jan;27:45–53.
 - [30] Caruso G, Salvaggio G, Campisi A, Melloni D, Midiri M, Bertolotto M, et al. Bladder tumor staging: comparison of contrast-enhanced and gray-scale ultrasound. *AJR Am J Roentgenol* 2010 Jan;194:151–6.