© Pol J Radiol 2021; 86: e449-e454
DOI: https://doi.org/10.5114/pjr.2021.107609

Received: 1.09.2020
Accepted: 04.02.2021
Published: 23.07.2021

Polish Journal of Radiology

http://www.polradiol.com

Original paper

Comparison of Urinary Tract Dilatation and Society of Fetal Urology systems in the detection of vesicourethral reflux and renal scar

Gulec Mert Dogan[1A,B,C,D,E,F,G], Ahmet Sigirci[1A,E], Aslinur Cengiz[1B,C,D], Sevgi Demiroz Tasolar[1E,F,G], Turan Yildiz[2A,B,D], Yilmaz Tabel[3C,D,E,F], Ahmet Taner Elmas[3C,D,E], Muge Otlu[4B,C], Sait Murat Dogan[5C,E,F]

[1]Department of Pediatric Radiology, Inonu University, Malatya, Turkey
[2]Department of Pediatric Surgery, Inonu University, Malatya, Turkey
[3]Department of Pediatric Nephrology, Inonu University, Malatya, Turkey
[4]Department of Nuclear Medicine, Inonu University, Malatya, Turkey
[5]Department of General Surgery, Inonu University, Malatya, Turkey

Abstract

Purpose: The presence and degree of hydronephrosis is very important in the management of many diseases of the urinary tract. In this study, we aim to compare the sensitivity and specificity of 2 classification systems that are used for hydronephrosis grading in ultrasound, for reflux and scar detection. The classification systems were the Society of Fetal Urology (SFU) and Urinary Tract Dilatation (UTD).

Material and methods: Ultrasounds and dimercaptosuccinic acid scintigraphies (DMSA) of all patients who underwent voiding cystourethrogram (VCUG) due to urinary tract infection were examined retrospectively. DMSA was accepted for scar detection and VCUG for reflux detection as reference methods. SFU classification was used for hydronephrosis in ultrasound reports, and UTD classification was made over the reports. Sensitivity, specificity, and positive and negative predictive values of UTD and SFU classification systems for reflux and scar detection were calculated, and these 2 systems were compared.

Results: 103 (39%) of the patients were male and 162 (61%) were female. Pathologies were detected in 192 (35%) of 530 kidneys in ultrasound. In 110 (42%) of the children, reflux was detected in VCUG. Scars in DMSA were detected in only 16% (44) of 266 kidneys. Sensitivity, positive and negative predictive values of the UTD classification system were statistically significantly higher than the SFU system for scar and reflux detection ($p < 0.01$).

Conclusions: If we use the UTD system in ultrasounds of patients with urinary tract infections, children reported as UTD 0 may not need VCUG, which reduces radiation exposure to children and the cost of the diagnostic interventions.

Key words: ultrasound, reflux, DMSA, voiding cystourethrogram.

Introduction

Urinary tract infections (UTI) are among the most common infections under the age of 5 years, which can cause very serious complications such as permanent kidney damage. Occurrence of these complications increase significantly when UTIs are accompanied by vesicourethral reflux (VUR) [1]. Ultrasound (US), dimercaptosuccinic acid scintigraphy (DMSA), and voiding cystourethrogram (VCUG) are used for the diagnosis of negative effects and complications of UTI [2]. Many guidelines have been published on which order and in which situations these im-

Correspondence address:
Prof. Sait Murat Dogan, Department of General Surgery, Inonu University, Malatya, Turkey, e-mail: saitmuratdogan@hotmail.com

Authors' contribution:
A Study design · **B** Data collection · **C** Statistical analysis · **D** Data interpretation · **E** Manuscript preparation · **F** Literature search · **G** Funds collection

aging methods should be used. Especially in recent years, it is believed that US findings are important in detecting high-grade VUR with the contribution of technological developments in US devices, and VCUG is not advised in the presence of a normal US [3,4].

The presence and degree of hydronephrosis are very important in the management of many congenital and acquired diseases of the urinary tract. For the standardization of hydronephrosis evaluation, in 1993, the Fetal Urology Association (SFU) proposed a classification system [5]. In addition, while many more grading systems are being used, the Urinary Tract Dilation (UTD) classification system was proposed in 2014 [6].

In this study, we aimed to compare the sensitivity and specificity of UTD and SFU classification systems in US examination for VUR detection by using VCUG as a reference method. Also, we aimed to compare the sensitivity and specificity of the UTD and SFU classification systems for scar detection by using DMSA as a reference method and to contribute to the guideline recommendations.

Material and methods

In 2018-2019, 436 children in the 0-17 years age range, who were followed up with a diagnosis of UTI, and who had VCUG, were evaluated retrospectively. Children with neurogenic bladder, and children with congenital and acquired urogenital anomalies were excluded from the study. Children whose US examinations were not performed by a paediatric radiologist and those with US reports containing missing information were excluded from the study. Consequently, 265 children and 530 kidneys were included in this study. All 265 children in the study group also had US examinations, but only133 of 265 children had DMSA.

The patients were divided into 3 groups according to their ages. Group 1, 0-2 years old; Group 2, 2-5 years old; Group 3, > 5 years old. Urinary system US was performed in the first 15 days of the diagnosis of the index UTI. DMSA scintigraphy findings, applied at least 120 days after active urinary tract infection, were recorded. When the renal function was within the normal range, radioisotope intake was homogeneous without an evident scar and there was no cortical hypoactivity, DMSA was considered normal. Significant localized deformity, volume loss in the outer contours, and volume reduction in the cortex were considered as scar [7].

Sensitivity, specificity, and positive and negative predictive values of UTD and SFU classification systems in US examinations were calculated for reflux detection by using VCUG as a reference method. For this, VCUG and US results of 530 kidneys were compared. Sensitivity, specificity, and positive and negative predictive values of UTD and SFU classification systems in US examinations were calculated for scar detection by using DMSA as a reference method. For this, DMSA and US results of 266 kidneys were compared.

Categorical data were expressed as count and percentage. Pearson's χ^2 test was used for comparisons based on independent groups. Sensitivity and specificity comparisons were performed by McNemar test. The weighted generalized score statistic proposed by Kosinski was used for comparisons of 2 positive or negative predictive values. In all analyses the significance level was considered as 0.05. R software (4.0.0) was used for statistical analysis. All procedures followed were in accordance with the Helsinki Declaration, and all of the parents of the patients were informed and all were approved to participate in this study. Institutional ethics committee approval was provided, with number 2020/750 on 02-06-2020.

Voiding cystourethrogram findings

Reflux evaluation was done by 2 paediatric radiologists: A.S has 12 years' experience in paediatric radiology, and G.M. has 3 years' experience in paediatric radiology. VUR was classified as 0 to 5 according to the International Reflux Study [8]. According to this classification, Grade 1 and 2 reflux were accepted as low grade and Grade 3 to 5 reflux as high grade.

Ultrasound findings

US examinations of all patients were performed by the same paediatric radiologist, who has 3 years' experience in paediatric radiology, using a GE LOGIQ S8 (USA). The SFU grading system was used in the urinary system US examination in 2018-2019 in our clinic. The SFU classification system, which we use, is divided into 5 groups [5]. Reports were analysed retrospectively by the same paediatric radiologist, and UTD classifications were made additionally over all reports [6]. All the parameters required for the UTD classification system were already detailed in our existing reports.

Results

Demographic and clinical data of the patients are summarized in Table 1. Fifty-seven children (52%) with reflux in VCUG were > 5 years old, and most of them were girls (66%), but there was no statistically significant difference between age and gender in terms of reflux ($p > 0.05$).

The values of each kidney in the SFU and UTD systems are shown in in Table 2. 76% (28) of kidneys with SFU 1 were normal in UTD system and 19% of kidneys with SFU 0 were evaluated as UTD 2-3.

While the SFU classification system indicated pathological in 15% (57) of kidneys without reflux in VCUG, this rate was 14% (51) in the UTD classification system. Only 6 (7%) kidneys with high-grade reflux in VCUG were normal in the UTD system, but 56 (67%) kidneys with high-grade reflux were normal in the SFU system.

Table 1. Patient demographics

Parameters	*n* (%)
Gender	
Female	162 (61)
Male	103 (39)
Age groups	
Group 1	71 (27)
Group 2	59 (22)
Group 3	135 (51)
VUR	
Reflux (+)	110 (42)
Reflux (–)	155 (58)
VUR – laterality	
Unilateral	72 (65)
Bilateral	38 (35)
VUR – side	
Right	75 (49)
Left	79 (51)
VCUG (kidney)	
High grade	83 (54)
Low grade	71 (46)
SFU system (kidney)	
Grade 0	417 (79)
Grade 1	37
Grade 2	50
Grade 3	15
Grade 4	11
UTD system (kidney)	
Normal	366 (69)
P1	28
P2	45
P3	91
DMSA (kidney)	
Normal	222 (84)
Scar	44 (16)
No DMSA	264 (49)
Scar	
Right kidney	20 (47)
Left kidney	24 (53)

Fifty-four per cent (83) of the refluxes detected in our study were high grade (Figures 1 and 2). Because the VCUG is considered as the gold standard for the diagnosis of VUR, the sensitivity, specificity, positive (PPV) and negative (NPV) predictive values of the SFU and UTD systems for the diagnosis of any grade of reflux and high-grade reflux (grade 3-5) are shown in Tables 3 and 4, respectively. The sensitivity, PPV, and NPV of the UTD classification system were statistically significantly higher than those of the SFU system for reflux and high-grade reflux ($p < 0.05$).

Table 2. Number of renal units in each grade according to the Society for Fetal Urology (SFU) grading system and the Urinary Tract Dilatation (UTD) classification system

	UTD 0	UTD 1	UTD 2	UTD 3	SFU system
SFU 0	337	1	20	59	417
SFU 1	28	6	2	1	37
SFU 2	1	21	16	12	50
SFU 3	0	0	7	8	15
SFU 4	0	0	0	11	11
UTD system	366	28	45	91	530

DMSA was performed on 133 of 265 children. Scars in DMSA were detected in only 16% (44) of 266 kidneys. Seventy-three per cent (32) of these kidneys had pathology in US. Sixty-two per cent (28) of these patients were girls, and 59% (26) were > 5 years old. Because scintigraphy is considered as the gold standard for the detection of renal scars, the sensitivity, specificity, PPV, and NPV of SFU and UTD systems for scar detection are shown in Table 5. The sensitivity, PPV, and NPV of the UTD classification system were statistically significantly higher than those of the SFU system for scar detection ($p < 0.05$).

Discussion

In our study, reflux was observed in 110 of 265 patients in VCUG. Sixty-six per cent (73) of them were girls, and 58% (64) of them were > 5 years old. This may be related to the higher occurrence of UTI in girls. Although UTI is seen most often in patients between 2 and 24 months of age [9], our patients were not within this range. The development of VUR in children with UTI varies depending on age, gender, and clinical picture [9].

In this study, the sensitivity of US was 96% and the NPV was 98% for high-grade VUR. Our results showed that, if an US examination of a kidney was normal but had VUR in VCUG, the grade of this reflux was highly likely to be low. In this study, only 6 kidneys had high-grade reflux, and no abnormalities were found in their US examinations. This supports that US can be a guide for need of VCUG examination. In the literature, the sensitivity and NPV of US is reported to be between 16 and 40%, and between 25 and 86% for reflux [10,11]. However, these rates increase significantly for high-grade VUR and are reported as 63-86% for sensitivity and 70-94% for NPV [12,13]. Our study is similar to the literature.

Sixteen per cent of kidneys had scar in DMSA in our study. Seventy-three per cent of these kidneys had pathology in US, and most of these pathologies were related

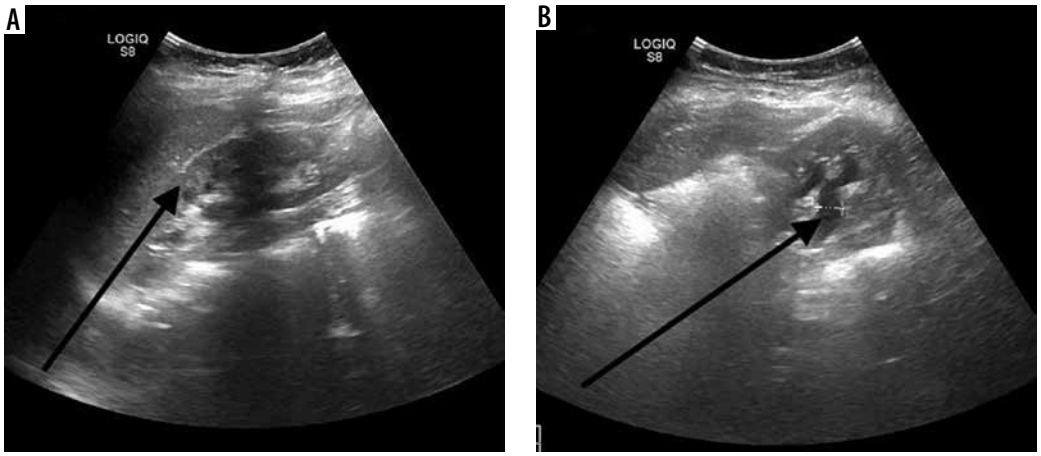

Figure 1. **A**) Ultrasound imaging in the sagittal plane. The parenchyma of the left kidney was abnormal and Urinary Tract Dilatation (UTD) classification system in this case was UTD 3. **B**) Imaging in the transverse plane US. Anterior-posterior diameter of the renal pelvis (APRPD) is 8 mm. Society for Fetal Urology (SFU) classification system in the same case was grade 2

Figure 2. Voiding cystourethrogram. There was a grade 3 (high grade) reflux to the left kidney

Table 3. The comparison of Urinary Tract Dilatation (UTD) and Society for Fetal Urology (SFU) classification for reflux (grade 1-5)

	SFU	UTD
Sensitivity	32.5	73.4
Specificity	84.8	85.6
PPV	49.6	67.7
NPV	76.5	88.7

PV – positive predictive value, NPV – negative predictive value

Table 4. The comparison of Urinary Tract Dilatation (UTD) and Society for Fetal Urology (SFU) classification for high-grade reflux

	SFU	UTD
Sensitivity	36.4	92.8
Specificity	80.8	79.9
PPV	23.9	46.1
NPV	86.6	98.3

PPV – positive predictive value, NPV – negative predictive value

Table 5. The comparison of Urinary Tract Dilatation (UTD) and Society for Fetal Urology (SFU) classification for scar

	SFU	UTD
Sensitivity	14	72.1
Specificity	78.4	67.1
PPV	11.1	29.8
NPV	82.5	92.5

PPV – positive predictive value, NPV – negative predictive value

to renal parenchymal changes. The sensitivity of the UTD system was much higher than that of the SFU system in our study for scar detection. It was reported that US was not sensitive enough to detect kidney damage in the study of Bush *et al.* [14]. In another study, it was stated that the sensitivity and the specificity of US in detecting renal scar could be 37-100% and 65-99%, respectively. It was stated that this wide range could be related to the differences of personal experiences [15]. The fact that the US examinations were all performed by one radiologist in our study and there were no interobserver variations might explain the high specificity and sensitivity values of our study.

Most of the kidneys (75,6%) classified as SFU 1 were found to be normal according to the UTD classification system. This is one of the most important results of our study. Most of these children had no reflux in VCUG or had low-grade reflux. Therefore, the false positivity of SFU in detecting VUR was high. Because this pathology in SFU is interpreted as abnormal in US, these children may be subjected to over-examination and -treatment.

Our other important result was that the sensitivity, PPV, and NPV of the UTD system were statistically significantly higher than those of the SFU system in detecting reflux and scar. The UTD system was able to classify 90% of high-grade VUR as moderate or high-risk hydronephrosis (P2 or P3). However, this was just 15% in the SFU system. In this new UTD classification system, unused parameters such as ureteral dilatation, bladder status, and renal parenchymal appearance are used [16]. In our study, the most important reason for the superiority of UTD sensitivity, PPV, and NPV over SFU was that the ureter dilation and parenchymal echogenicity increased the UTD degree even with a normal pelvicalyceal system. In many studies, these parameters are shown to increase the sensitivity and specificity of US detection for VUR [17,18].

Han *et al*. evaluated the reliability of the UTD and SFU grading system for postpartum urinary dilatation in infants [19]. As a result, they said that the UTD system had a better agreement between the observers. In their study, similarly to ours, half of the kidneys with SFU grade 1-2 were normal according to the UTD classification. Although there are similar studies in the literature for antenatal and postnatal hydronephrosis monitoring and treatment related to the comparison of the UTD and SFU classification systems [20,21], our study is the first study to compare the UTD classification system with the SFU classification system in detecting of VUR.

Contrast-enhanced ultrasound (CEUS) is a technique whereby biocompatible microspheres of inert gas are administered intra venously, which reflect ultrasonography sound waves. This technique has the potential to replace imaging studies such as DMSA because it does not involve radiation [22]. In recent years, there have been studies on the sensitivity of this method, especially in renal scar detection, and it was concluded that CEUS is a very sensitive and cost-effective diagnostic imaging method for the detection and monitoring of renal scars in children with VUR [23].

There were some limitations to our study. Our study was retrospective, all patients did not have DMSA, and the number of scars detected in DMSA was low. The inclusion of both index and recurrent UTIs in our study may have affected our results.

Conclusions

If the UTD system is used in the US of patients with UTI, children reported as UTD 0 may not need VCUG, which reduces the radiation exposure and cost. When urinary system US for a paediatric patient is performed by an experienced radiologist, examining all parameters in detail, reflux and scar can be predicted.

Conflict of interest

The authors report no conflict of interest.

References

1. Roberts KB. Urinary tract infection: Clinical practice guideline for the diagnosis and management the initial UTI in febrile infants and children 2 to 24 months. Pediatrics 2011; 128: e595-610.
2. La Scola C, De Mutiis C, Hewitt IK, et al. Different guidelines for imaging after first UTI in febrile infants: yield, cost and radiation. Pediatrics 2013; 131: e665-671.
3. Whiting P, Westwood M, Bojke L, et al. Clinical effectiveness and cost-effectiveness of tests for the diagnosis and investigation of urinary tract infection in children: a systematic review and economic model. Health Technol Assess 2006; 10: iii-iv, xi-xiii, 1-154.
4. Levart TK, Kljucewsek D, Kenig A, et al. Sensitivity of ultrasonography in detecting renal parenchymal defects: 6 year's followup. Pediatr Nephrol 2009; 24: 1193-1197.
5. Fernbach S, Maizels M, Conway J. Ultrasound grading of hydronephrosis: introduction to the system used by the Society for Fetal Urology. Pediatr Radiol 1993; 23: 478-480.
6. Nguyen HT, Benson CB, Bromley B, et al. Multidisciplinary consensus on the classification of prenatal and postnatal urinary tract dilation (UTD classification system). J Pediatr Urol 2014; 10: 982-998.
7. Piepsz A, Colarinha P, Gordon I, et al. Guidelines for 99mTc-DMSA scintigraphy in children. Eur J Nucl Med 2001; 28: BP37-BP41.
8. Lebowitz RL, Olbing H, Parkkulainen KV, et al. International system of radiographic grading of vesicoureteric reflux. International Reflux Study in Children. Pediatr Radiol 1985; 15: 105-109.
9. Cleper R, Krause I, Eisenstein B, et al. Prevalence of vesicoureteral reflux in neonatal urinary tract infection. Clin Pediatr (Phila) 2004; 43: 619-625.
10. Zamir G, Sakran W, Horowitz Y, et al. Urinary tract infection: is there a need for routine renal ultrasonography? Arch Dis Child 2004; 89: 466e8.
11. Massanyi EZ, Preece J, Gupta A, et al. Utility of screening ultrasound after first febrile UTI among patients with clinically significant vesicoureteral reflux. Urology 2013; 82: 905e9.
12. Preda I, Jodal U, Sixt R, et al. Value of ultrasound in evaluation of infants with first urinary tract infection. J Urol 2010; 183: 1984e8.
13. Tsai JD, Huang CT, Lin PY, et al. Screening high-grade vesicoureteral reflux in young infants with a febrile urinary tract infection. Pediatr Nephrol 2012; 27: 955e63.
14. Bush NC, Keays M, Adams C, et al. Renal damage detected by DMSA, despite normal renal ultrasound, in children with febrile UTI. J Pediatr Urol 2015; 11: 126.e1-7.
15. Roebock DJ, Howard RG, Metreweli C. How sensitive is ultrasound in the detection of renal scars? Br J Radiol 1999; 72: 345.

16. Polito C, Rambaldi PF, Signoriello G, et al. Permanent renal parenchymal defects after febrile UTI are closely associated with vesicoureteric reflux. Pediatr Nephrol 2006; 21: 521-526.
17. Leroy S, Vantolon S, Larakeb A, et al. Vesicoureteral reflux in children with urinary tract infection: comparison of diagnostic accuracy of renal US criteria. Radiology 2010; 255: 890-898.
18. KiM J, Lim YJ, Yi J, et al. Diagnostic accuracy of renal ultrasonography for vesicoureteral reflux in infants and children aged under 24 months with urinary tract infections. J Korean Soc Radiol 2019; 80: 1179-1189.
19. Han M, Kim HG, Lee JD, et al. Conversion and reliability of two urological grading systems in infants: the Society for Fetal Urology and the urinary tract dilatation classifications system. Pediatr Radiol 2017; 47: 65-73.
20. Hodhod A, Capolicchio JP, Jednak R, et al. Evaluation of tract dilation classification system for grading postnatal hydronephrosis. J Urol 2016; 195: 725-730.
21. Nelson CP, Heller HT, Benson CB, et al. Interobserver reliability of the antenatal consensus classification system for urinary tract dilatation. J Ultrasound Med 2020; 39: 551-557.
22. Correas J.M, Anglicheau D, Joly D. Ultrasound-based imaging methods of the kidney-recent developments. Kidney Int 2016; 90: 1199-1210.
23. Hains DS, Cohen HL, McCarville MB. Elucidation of renal scars in children with vesicoureteral reflux using contrast-enhanced ultrasound: a pilot study. Kidney Int Rep 2017; 2: 420-424.