

Increasing the Reliability of the Grading System for Voiding Cystourethrograms Using Ultrasonography: An Inter-Rater Comparison

Seyithan Ozaydin,^{1,*} Suleyman Celebi,¹ Ismail Caymaz,³ Cemile Besik,¹ Birgul Karaaslan,¹ Ozgur Kuzdan,¹ and Serdar Sander²

¹Department of Pediatric Surgery, Kanuni Sultan Suleyman Education and Research Hospital, Istanbul, Turkey

²Department of Pediatric Urology, Kanuni Sultan Suleyman Education and Research Hospital, Istanbul, Turkey

³Department of Pediatric Radiology, Faculty of Medicine, Baskent University, Istanbul, Turkey

* Corresponding author: Seyithan Ozaydin, Department of Pediatric Surgery, Kanuni Sultan Suleyman Education and Research Hospital, Istanbul, Turkey. Tel: +05-057456577, Fax: +090-2162453693, E-mail: celebisuleyman@hotmail.com

Received 2016 April 23; Revised 2016 June 11; Accepted 2016 June 12.

Abstract

Background: To assess the effectiveness of the current vesicoureteral reflux (VUR) grading system according to the international classification of VUR (ICVUR) and to evaluate whether VUR grading accuracy could be improved by renal ultrasonography (RU) according to the Society for Fetal Urology (SFU) grading system.

Objectives: Therefore, this study assessed the accuracy of the current VCUG staging system by assessing inter-rater reliability among pediatric radiologists and urologists; it also evaluated whether accuracy is increased by RU without consensus (with respect to VCUG grading).

Methods: Four pediatric urologists and four pediatric radiologists independently graded 120 voiding cystourethrograms (VCUGs). Middle VUR grades were divided into the following three groups: VUR consensus grade III (group 1), VUR consensus grade IV (group 3), and VUR non-consensus grades III and IV (group 2). All groups were compared with respect to hydronephrosis grade using RU.

Results: The intraclass correlation coefficient (ICC) values ranging from 0.86 to 0.89 reflected good reliability. The lowest agreement was associated with middle grades (III and IV). A marked difference in sensitivity was observed between groups 1 and 3 (35% and 95%, respectively, $P < 0.05$), indexed by SFU hydronephrosis grade, suggesting that VCUG cases in group 2 ($n = 16$ at SFU 0 or 1) could be accepted as grade III, and SFU scores of 2, 3, or 4 could be considered grade IV.

Conclusions: Inter-rater accuracy could be improved at middle grades using renal ultrasonography (USG), which could promote communication between different specialists.

Keywords: Inter-Rater Reliability, Vesicoureteral Reflux, Ultrasonography, Society for Fetal Urology, International Classification of Vesicoureteral Reflux

1. Background

Vesicoureteral reflux (VUR), defined as the retrograde flow of urine from the bladder back up the ureter into the kidney, is diagnosed in 30% to 40% of children who present with urinary tract infections (UTIs) (1). It is a congenital condition that may resolve or improve over time (1). Because of the relatively high prevalence of renal scarring, it is prudent to understand how to identify VUR, potential problems associated with chronic VUR, and the most effective therapeutic strategies.

Voiding cystourethrography (VCUG) or radionuclide cystourethrography is the gold standard for diagnosis of VUR (2, 3). During recent decades, new diagnos-

tic modalities have been introduced, including contrast-enhanced voiding USG and magnetic resonance voiding cystourethrography (MRVCUG) to detect VUR (3). However, there is no recommendation for using these newly introduced modalities as the first line diagnostic method. Currently, the standard test for diagnosing VUR is the VCUG, which is also used to classify the severity of reflux. One of the latest studies from Greenfield et al. (1) reported divergent grade interpretation in 9 out of 61 ureters initially assessed as middle grade (15%). Of these 9 discrepancies, 7 (78%) were adjudicated to the higher grade. Greenfield et al. concluded that discrepancies in the assessment of intermediate grade VUR were noteworthy and added in particular that there was considerable disagreement in the evalu-

ation of intermediate grades of reflux. It is very important that physicians reach a universal consensus regarding the stages of VUR by adhering to common principles, because staging determines whether each child should simply be closely observed, receive prophylactic antibiotics, or undergo endoscopic treatment or surgery. However, there is a discrepancy in staging at the middle grades (III and IV) (2), and the reliability of inter-rater reflux grading has largely been ignored in the literature.

Although USG may not be an ideal diagnostic tool for the prediction of VUR when various reflux grades are considered simultaneously, it is useful for the prediction of high grades (i.e., grades IV and V) (3). Normal renal ultrasonography (RU) is infrequently applied to grade IV reflux, whereas moderate-to-severe hydronephrosis is rare in the context of reflux grades < IV (4).

2. Objectives

Therefore, this study assessed the accuracy of the current VCUG staging system by assessing inter-rater reliability among pediatric radiologists and urologists; it also evaluated whether accuracy is increased by RU without consensus (with respect to VCUG grading).

3. Methods

After approval by the local institutional review board, we prospectively recruited 120 children with primary unilateral VUR after their first occurrence of UTI between January, 2013 and January, 2015. Patients who underwent VCUG and had exstrophy vesicae or other abnormalities including ectopic ureterocele, ureteral duplication, and renal pelvic anomalies were excluded from the study. The VCUGs of these patients were shown to four experienced pediatric urologists and four experienced pediatric radiologists; raters were provided with de-identified images from our hospital image repository so that they were blind to the patients' clinical history. These physicians were asked to grade VUR (from I - V) according to the International Classification of Vesicoureteral Reflux (ICVUR) system (3). Each image was assessed 3 times by each doctor: at baseline, after 6 weeks, and after 3 months. The responses from the urologists and radiologists were compared, and discrepancies were adjudicated to a final assessment. Middle-grade VUR groups were further subdivided into groups 1 (consensus grade III), 2 (non-consensus grade III or IV), or 3 (consensus grade IV).

Finally, a single, experienced pediatric radiologist graded hydronephrosis on RU while blinded to patients' clinical history, using the method of the Society for Fetal Urology (SFU) (4). VUR grade according to the ICVUR

system was compared to hydronephrosis findings on RU using SFU. Fullness corresponds to SFU grade I, mild hydronephrosis, and visualization of the renal pelvis only. SFU grade II is characterized by moderate hydronephrosis, visualization of the renal pelvis, and some but not all calyces. SFU grade III is characterized by mild hydronephrosis, visualization of the renal pelvis, and virtually all calyces. Finally, SFU grade IV is characterized by severe hydronephrosis, and cortical thinning is also observed (4).

For statistical analysis, to examine the relationship between hydronephrosis and reflux grade, comparison of categorical variables was performed using chi-square tests and Fisher's exact test when the sample size was sufficiently small. The intraclass correlation coefficient (ICC) was used to calculate inter-rater reliability. A value of $P < 0.05$ was considered statistically significant, and the 95% confidence interval was evaluated. $ICC > 0.72$ indicated adequate reliability, whereas an $ICC > 0.8$ indicated good reliability.

4. Results

The study group included 120 children (86 girls and 34 boys) with unilateral VUR and a median age of 3.2 years (range: 6 months to 10 years), and most children (91%) were enrolled after recurrent UTI. The UTI prior to enrollment was both febrile and symptomatic in 69 children, only febrile in 28, and only symptomatic in 23 children. Symptoms were mainly voiding symptoms, such as dysuria and frequent urination. The other reflux nephropathies manifesting were abdominal pain or classic flank pain and tenderness. Less frequent symptoms included vomiting, diarrhea, anorexia, and lethargy. All cases of VUR were diagnosed using VCUG. In total, 120 VCUGs were reviewed by four pediatric urologists and four pediatric radiologists, yielding a total of 960 observations.

Among the pediatric urologists, ICC scores were consistently > 0.8 , between 0.81 and 0.87, indicating good reliability (Table 1). Among the radiologists' evaluations, the ICC scores were consistent, ranging from 0.83 to 0.91 (Table 1), indicating reliability of VCUG grading in this group of physicians. The inter-rater agreement was 0.84. Values were also compatible and reliable (> 0.8 ; Table 2). However, a closer inspection of scoring among pediatric urologists and radiologists revealed significant discrepancies between grades III and IV. For pediatric urologists, 1 rater graded 24 cases (20%) as III, whereas the other raters graded 33 cases (27%) as III. Similar ratings were given by the radiologists; 1 radiologist graded 23 cases (19%) as III, whereas the others graded 33 cases (27%) as III (Table 3). The same type of discrepancy also occurred for cases given a grade of IV. For pediatric urologists, 26 cases (21%) were graded as

IV by 1 rater, while another rater graded 35 cases (29%) as IV. The radiologists used similar grades; 1 radiologist graded 27 cases (22.5%) as IV, whereas the other radiologist graded 37 cases (30%) as IV (Table 3).

Table 1. Assessment VUR Grading Among Pediatric Urologists and Radiologists

	Intraclass Correlation Coefficient	95% Confidence Interval	
		Lower Bound	Upper Bound
Pediatric urologist 1	0.88	0.88	0.92
Pediatric urologist 2	0.86	0.84	0.87
Pediatric urologist 3	0.85	0.83	0.86
Pediatric urologist 4	0.81	0.80	0.83
Pediatric radiologist 1	0.87	0.86	0.88
Pediatric radiologist 2	0.85	0.83	0.86
Pediatric radiologist 3	0.84	0.83	0.86
Pediatric radiologist 4	0.86	0.84	0.91

Table 2. Agreement Between Radiologists and Pediatric Urologists on VCUG Assessment

	Intraclass Correlation Coefficient	95% Confidence Interval	
		Lower Bound	Upper Bound
Pediatric urologists	0.887	0.879	0.937
Radiologists	0.854	0.839	0.864

Middle and high reflux grades were associated with the degree of hydronephrosis revealed by USG, according to the SFU system (Table 4). USG did not detect any hydronephrosis at grades I or II according to SFU, but there was a marked difference in sensitivity between grades III and IV for VUR (35% versus 95%, respectively) and hydronephrosis on USG, suggesting that USG represents a more useful diagnostic tool for predicting grade IV than grade III (Table 4). Significantly fewer grade IV patients were classified without hydronephrosis compared to the other reflux grades (i.e., < grade IV; $P < 0.01$); 91% of group 1 patients were SFU 0 or SFU 1, compared to 55% in group 2 and 15% in group 3. Total SFU (2, 3, and 4) was greater in group 3 ($n = 13$; 85%) than in group 1 ($n = 2$; 9%). This sug-

gests that VCUG cases in group 2 ($n = 9$ at SFU 0 and 1) could also be accepted as VUR grade III, and that cases in SFU 2, 3, and 4 ($n = 7$; 43%) could be considered VUR grade IV (Table 4).

Significantly fewer grade IV patients were classified without hydronephrosis compared to the other reflux grades (i.e., < grade IV; $P < 0.05$). Total SFU (0 and 1) was significantly higher in group 1 than in group 3. Total SFU (2, 3, and 4) was greater in group 3 than in group 1. This suggests that VCUG cases in group 2 ($n = 9$ at SFU 0 and 1) could also be accepted as VUR grade III, and that cases in SFU 2, 3, and 4 ($n = 7$) could be considered VUR grade IV.

5. Discussion

VUR is caused by an abnormal vesicoureteral junction that is too short or is unattached between the ureter and the detrusor muscle; this results in retrograde passage of urine back up into the ureter (5). Reports from the 1960s and 1970s, when VUR was less frequently recognized, reveal that renal scarring due to VUR was the etiology of 50% of hypertension cases and 30% of end stage renal disease (ESRD) cases in children (6); therefore, the current standard of care includes imaging to assess the presence and extent of VUR (7). This has resulted in more effective recognition and treatment of VUR, which has considerably reduced ESRD rates, with scarring now accounting for only 5% of pediatric cases of significant renal impairment (8).

Ultrasound is typically the first test performed following the diagnosis of a UTI; however, USG cannot identify VUR, particularly for low grades (9). Hoberman et al. (10) and Zamir et al. (11) despite finding 12% and 14% of abnormalities, respectively, by ultrasound in children with first-occurrence febrile UTI showed that management was not altered in any of them. The most common clinical practice is for those who present with a UTI to still be evaluated for VUR with a VCUG. The advantage of this method is the ability to grade reflux severity using the widely accepted 5-level International Scale (12). The extent of passage into the ureter is categorized hierarchically, with grades I and II including only the ureter. Grade III is reflux with mild to moderate dilation and minimal blunting of fornices, grade IV entails moderate ureteral tortuosity, and grade V reflux is distention of the renal pelvis and calyces, loss of papillary impressions, and ureteral tortuosity (13). The majority of children affected by this condition have low-grade VUR (grades I - II), for which certain guidelines exist concerning prophylaxis (14). The strategy depends on the hypothesis that reflux, especially VUR of grade III or greater, increases the risk of recurrent UTIs and renal scarring, which can lead to sequelae such as proteinuria, hypertension, and ESRD later in life (15). Thus, grading must

Table 3. Adjudication of Discrepancies in VUR grade Among Raters^a

Grade	Pediatric Urologists (PU)				Pediatric Radiologists (PR)			
	PU 1	PU 2	PU 3	PU 4	PR 1	PR 2	PR 3	PR 4
I	27 (22)	26 (21)	25 (20)	27 (22)	27 (22)	26 (21)	27 (22)	26 (21)
II	23 (19)	24 (20)	25 (20)	23 (19)	23 (19)	24 (20)	23 (19)	24 (20)
III	25 (20)	31 (26)	33 (27)	24 (20)	27 (22)	33 (27)	23 (19)	30 (25)
IV	34 (28)	29 (24)	26 (21)	35 (29)	32 (27)	27 (22)	37 (30)	29 (24)
V	11 (9)	10 (8)	11 (9)	11 (9)	11 (9)	10 (8)	10 (8)	11 (9)

^aValues are expressed as No. (%).**Table 4.** Severity of Hydronephrosis (SFU) Versus Reflux Grade (ICVUR)^a

	Reflux Grade					
	I (n = 27)	II (n: 23)	III (n = 23) Group 1	III or IV (n = 16) Group 2	IV (n = 20) Group 3	V (n = 11)
None (SFU 0)	22	18	15 (65)	7 (43)	1 (5)	0
Fullness (SFU grade 1)	5	4	6 (26)	2 (12)	2 (10)	1
Mild (SFU grade 2)	0	1	2 (9)	4 (25)	7 (35)	2
Moderate (SFU grade 3)	0	0	0	3 (18)	8 (40)	3
Severe (SFU grade 4)	0	0	0	0	2 (10)	5

^aValues are expressed as No. (%).**Table 5.** Severity Hydronephrosis (SFU) vs. Reflux Grade (ICVUR)^{a,b}

	Reflux Grade				
	I	II	III	IV	V
None (SFU 0)	22 (81.4)	18 (78.2)	15 (65.2)	1 (5)	0
Fullness (SFU grade 1)	5 (18.5)	4 (17.3)	6 (26)	2 (10)	1 (9)
Mild (SFU grade 2)	0	1 (4.3)	2 (8.6)	7 (35)	2 (18)
Moderate (SFU grade 3)	0	0	0	8 (40)	3 (27)
Severe (SFU grade 4)	0	0	0	2 (10)	5 (45)

^aValues are expressed as No. (%).^bSignificantly fewer grade IV patients were classified without hydronephrosis compared to reflux grade III (i.e., < grade IV; $P < 0.05$).

be accurate so that reflux of grade III or higher can reliably be distinguished from lower grade reflux to guide the decision of whether to initiate prophylactic antibiotics or to treat more aggressively with endoscopy or surgical treatment.

Recently, several authors have suggested that treatment be limited initially, with follow-up imaging applied to those diagnosed with low grades (16, 17). If this advice is followed, diagnostic differences at this cut-off are critical during the determination of treatment. One possible strategy to reduce the incidence of inappropriate treatment would be to employ a second rater to review any VCUG

grade III or IV cases, as well as for raters to reach a consensus regarding the most appropriate grade. When treatment is provided based on the recommendations of a single rater, undergrading and possible undertreatment occurs in 23% - 38% of cases (18). Craig et al. (19) reported near perfect agreement (kappa 90% to 91%) when three radiologists separately graded contrast VCUGs. However, Kronemer et al. (20) reported divergent grade interpretation in 20 of 39 patients with VUR when 2 radiologists separately read the studies. Keays et al. (20), also analyzed reflux grades and concluded that although the overall VUR grading of VCUGs was shown to be reliable, agreement was

highest at the extremes of the scale (grades I and V); scoring discrepancies were more common at the middle grades (II-IV). Our study found that among groups of pediatric urologists and radiologists, as well as in comparisons between the two groups, the ICC value was close to 0.9, indicating reliable grading. We found that most discrepancies concerned the VUR of grades III and IV, because both grades subjectively depend on the appearance of the renal calyx, without numerical values being taken into account. In addition, inter-observer assessment differences typically spanned only a single grade, which could nonetheless undermine reliability.

Treatment of a VUR case greatly depends upon the grade it is assigned (21); however, as our results demonstrated, VUR grade often varies at middle grades depending on the observer. Before a treatment algorithm can be discussed, the diagnostic grade must be expressed via objective numerical values. Confusion in this staging system arises from the fact that the current 5-level grading system cannot be easily applied to VCUGs, which only include characteristics of four stages: stage 1, in which the ureter is affected but reflux does not reach the renal pelvis; stage 2, in which the ureter is affected and reflux reaches the renal pelvis; stage 3, in which the renal calyceal system is affected; and stage 4, characterized by gross dilation and kinking of the ureter with papillary impressions no longer visible.

More objective and quantitative data are required to divide VCUG findings into five stages according to the ICVUR. The poor agreement on moderate grades may stem from differences in judging the degree of dilation of the calyceal system.

The SFU grading system emphasizes the importance of internal calyceal dilation rather than the size of the renal pelvis (4). One meta-analysis indicated that the SFU grading system is the most consistent and widely used (11/25 studies) (20), with good intra-rater reliability (21). The SFU grading system (22) comprises five grades and evaluates dilation of the renal pelvis, distinguishes between central (major) and peripheral (minor) calyceal dilation, and measures parenchymal thickness. Another meta-analysis revealed that the severity of UT dilation, based on SFU criteria, was correlated with urological pathology (23).

Abnormal USG finding was defined if the patient had hydronephrosis, dilatation of the ureter, elevated cortical echogenicity, decreased cortical thickness, and increased kidney size. USG cannot predict low grade VUR, whereas VCUG is more reliable and the standard method for detection of VUR. However, in a study by Lee et al. (24), the successful prediction rates of VUR by USG were 41.7% and 86% in low and high grade VUR by USG, respectively. RU may not be optimal for the prediction of VUR when vari-

ous grades of reflux are considered simultaneously. Normal RU is rarely applied to grade IV reflux, and moderate to severe hydronephrosis is rarely observed at reflux grade III. This suggests that the proportion of patients correctly diagnosed increases with the use of USG, particularly for grades III and IV. We herein demonstrate that grade IV specificity remains high, with a significant increase in sensitivity compared to grade III, during USG use. Kovanlıkaya et al. (25) indicated that, for reflux grades IV and V, RU can accurately predict the presence of reflux, with only 4.5% of grade IV patients and 0.7% of grade V patients misdiagnosed using renal bladder USG alone. In another recent study, RU was highly accurate for the prediction of VUR; the authors concluded that a normal RU largely excluded high grades (i.e., grades IV and V) in pediatric patients with UTI and mild renal scarring (21). These data indicate that USG can reliably predict high-grade reflux by distinguishing such cases from patients with lower grade or absent reflux. In our study, the marked difference in sensitivity between grades III and IV for hydronephrosis suggests that USG represents a useful diagnostic tool with which to differentiate these grades (21).

Although no study evaluating reliability can completely replicate daily practice, during VCUG evaluation in the present study, we aimed to replicate daily clinical practice to the greatest degree possible, such that our results are highly reliable except at grades III and IV with respect to overall grading. We suggest that RU can be used to distinguish between these grades.

In conclusion, VCUG has long been a mainstay of the diagnosis and grading of VUR, and our study confirmed the reliability of this method. However, discrepancies arise in grading abnormalities of the calyceal system seen on VCUGs at middle grades, which could greatly impact the treatment method used. Although USG may not be the ideal tool for the prediction of VUR when various reflux grades are considered simultaneously, it represents a useful method for differentiating grades III and IV and reducing grading discrepancies, which could facilitate communication and collaboration among different specialists.

References

1. Greenfield SP, Ng M, Wan J. Resolution rates of low grade vesicoureteral reflux stratified by patient age at presentation. *J Urol*. 1997;157(4):1410-3. [PubMed: 9120966].
2. Metcalfe CB, Macneily AE, Afshar K. Reliability assessment of international grading system for vesicoureteral reflux. *J Urol*. 2012;188(4):1490-2. doi: 10.1016/j.juro.2012.02.015. [PubMed: 22906646].
3. Lebowitz RL, Olbing H, Parkkulainen KV, Smellie JM, Tamminen-Mobius TE. International system of radiographic grading of vesicoureteric reflux. International Reflux Study in Children. *Pediatr Radiol*. 1985;15(2):105-9. [PubMed: 3975102].

4. Fernbach SK, Maizels M, Conway JJ. Ultrasound grading of hydronephrosis: introduction to the system used by the Society for Fetal Urology. *Pediatr Radiol*. 1993;**23**(6):478-80. [PubMed: 8255658].
5. Schwab CW, Wu HY, Selman H, Smith GH, Snyder HM, Canning DA. Spontaneous resolution of vesicoureteral reflux: a 15-year perspective. *J Urol*. 2002;**168**(6):2594-9. doi: 10.1097/01.ju.0000037530.11361.8b. [PubMed: 12441993].
6. Ardissino G, Avolio L, Dacco V, Testa S, Marra G, Vigano S, et al. Long-term outcome of vesicoureteral reflux associated chronic renal failure in children. Data from the ItalKid Project. *J Urol*. 2004;**172**(1):305-10. [PubMed: 15201801].
7. Wan J, Skoog SJ, Hulbert WC, Casale AJ, Greenfield SP, Cheng EY, et al. Section on Urology response to new Guidelines for the diagnosis and management of UTI. *Pediatrics*. 2012;**129**(4):1051-3. doi: 10.1542/peds.2011-3615. [PubMed: 22412033].
8. Vallee JP, Vallee MP, Greenfield SP, Wan J, Springate J. Contemporary incidence of morbidity related to vesicoureteral reflux. *Urology*. 1999;**53**(4):812-5. [PubMed: 10197863].
9. Foresman WH, Hulbert WC, Rabinowitz R. Does urinary tract ultrasonography at hospitalization for acute pyelonephritis predict vesicoureteral reflux?. *J Urol*. 2001;**165**(6):2232-4. [PubMed: 11371951].
10. Hoberman A, Charron M, Hickey RW, Baskin M, Kearney DH, Wald ER. Imaging studies after a first febrile urinary tract infection in young children. *N Engl J Med*. 2003;**348**(3):195-202. doi: 10.1056/NEJMoa021698. [PubMed: 12529459].
11. Zamir G, Sakran W, Horowitz Y, Koren A, Miron D. Urinary tract infection: is there a need for routine renal ultrasonography?. *Arch Dis Child*. 2004;**89**(5):466-8. [PubMed: 15102643].
12. Weiss R, Tamminen-Mobius T, Koskimies O, Olbing H, Smellie JM, Hirche H, et al. Characteristics at entry of children with severe primary vesicoureteral reflux recruited for a multicenter, international therapeutic trial comparing medical and surgical management. The International Reflux Study in Children. *J Urol*. 1992;**148**(5):1644-9. [PubMed: 1433580].
13. Smellie JM, Jodal U, Lax H, Mobius TT, Hirche H, Olbing H, et al. Outcome at 10 years of severe vesicoureteric reflux managed medically: Report of the International Reflux Study in Children. *J Pediatr*. 2001;**139**(5):656-63. doi: 10.1067/mpd.2001.117583. [PubMed: 11713442].
14. Roussey-Kesler G, Gadjos V, Idres N, Horen B, Ichay L, Leclair MD, et al. Antibiotic prophylaxis for the prevention of recurrent urinary tract infection in children with low grade vesicoureteral reflux: results from a prospective randomized study. *J Urol*. 2008;**179**(2):674-9. doi: 10.1016/j.juro.2007.09.090. [PubMed: 18082208].
15. Montini G, Rigon L, Zucchetta P, Fregonese F, Toffolo A, Gobber D, et al. Prophylaxis after first febrile urinary tract infection in children? A multicenter, randomized, controlled, noninferiority trial. *Pediatrics*. 2008;**122**(5):1064-71. doi: 10.1542/peds.2007-3770. [PubMed: 18977988].
16. Hoberman A, Keren R. Antimicrobial prophylaxis for urinary tract infection in children. *N Engl J Med*. 2009;**361**(18):1804-6. doi: 10.1056/NEJMe0907623. [PubMed: 19864680].
17. Peters CA, Skoog SJ, Arant BS, Copp HL, Elder JS, Hudson RG, et al. Summary of the AUA Guideline on Management of Primary Vesicoureteral Reflux in Children. *J Urol*. 2010;**184**(3):1134-44. doi: 10.1016/j.juro.2010.05.065. [PubMed: 20650499].
18. Craig JC, Irwig LM, Christie J, Lam A, Onikul E, Knight JF, et al. Variation in the diagnosis of vesicoureteric reflux using micturating cystourethrography. *Pediatr Nephrol*. 1997;**11**(4):455-9. [PubMed: 9260245].
19. Kronemer KA, Don S, Luker GD, Hildebolt C. Soft-copy versus hard-copy interpretation of voiding cystourethrography in neonates, infants, and children. *AJR Am J Roentgenol*. 1999;**172**(3):791-3. doi: 10.2214/ajr.172.3.10063884. [PubMed: 10063884].
20. Keays MA, Guerra LA, Mihill J, Raju G, Al-Asheeri N, Geier P, et al. Reliability assessment of Society for Fetal Urology ultrasound grading system for hydronephrosis. *J Urol*. 2008;**180**(4):1680-2. doi: 10.1016/j.juro.2008.03.107. [PubMed: 18708207].
21. Bayram MT, Kavukcu S, Alaygut D, Soyul A, Cakmakci H. Place of ultrasonography in predicting vesicoureteral reflux in patients with mild renal scarring. *Urology*. 2014;**83**(4):904-8. doi: 10.1016/j.urology.2013.10.019. [PubMed: 24315308].
22. Sidhu G, Beyene J, Rosenblum ND. Outcome of isolated antenatal hydronephrosis: a systematic review and meta-analysis. *Pediatr Nephrol*. 2006;**21**(2):218-24. doi: 10.1007/s00467-005-2100-9. [PubMed: 16362721].
23. Costi GA, Merguerian PA, Cheerasarn SP, Shortliffe LM. Sonographic renal parenchymal and pelvicaliceal areas: new quantitative parameters for renal sonographic followup. *J Urol*. 1996;**156**(2):725-9. [PubMed: 8683769].
24. Lee HY, Soh BH, Hong CH, Kim MJ, Han SW. The efficacy of ultrasound and dimercaptosuccinic acid scan in predicting vesicoureteral reflux in children below the age of 2 years with their first febrile urinary tract infection. *Pediatr Nephrol*. 2009;**24**(10):2009-13. doi: 10.1007/s00467-009-1232-8. [PubMed: 19593590].
25. Kovanlikaya A, Kazam J, Dunning A, Poppas D, Johnson V, Medina C, et al. The role of ultrasonography in predicting vesicoureteral reflux. *Urology*. 2014;**84**(5):1205-10. doi: 10.1016/j.urology.2014.06.057. [PubMed: 25443935].