Med Ultrason 2020, Vol. 22, no. 3, 287-292
DOI: 10.11152/mu-2311

A quantitative grading system of vesicoureteral reflux by contrast-enhanced voiding urosonography

Hui-Ming Yi[1,2,3], Xin-Wu Cui[1], Bao-Huan Cai[4], Li-Ru Qiu[4], Peng-Fei Song[2,3], Wei Zhang[1,2,3]

[1]Department of Medical Ultrasound, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [2]Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA, [3]Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA, [4]Department of Paediatrics, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Abstract

Aims: Contrast-enhanced voiding urosonography (ceVUS) is a well-established imaging modality for the diagnosis of vesicoureteral reflux (VUR). However, discrepancies of grading diagnosis of VUR exist due to the qualitative grading criteria currently used in clinics. This study aimed to evaluate numerical markers for a quantitative VUR grading system. **Material and methods:** CeVUS images of grade II-VVUR were analysed. A quantitative indicator, i.e. sectional area ratio (SAR), on the imaging section with maximum cross-section area and the presence of kidney hilum was calculated to distinguish different grades of VUR. The diagnostic performance of SAR was evaluated using receiver operating characteristic curve (ROC) analysis, and the maximum Youden Index was used to determine the optimal cut-off values. **Results:** A total of 63 patients with 126 PelviUreteral Units were enrolled. The SAR value increased significantly along with the increase of VUR grade. SAR had an excellent diagnostic performance in grading VUR. For differentiating VUR of grade II vs III, III vs IV and IV vs V, the area under the ROC curve values of SAR were 0.967, 0.943 and 0.865, respectively, while the optimal SAR cut-off values were 14.3%, 34.9% and 51.0%, respectively. The quantitative grading system based on the optimal SAR cut-off values showed excellent consistency with the qualitative grading system of VUR currently used in clinic. **Conclusions:** The numerical indicator SAR calculated from ceVUS may be used to establish a quantitative VUR grading system with excellent diagnostic performance and can potentially serve as a reliable tool for the evaluation and follow-up of VUR.

Keywords: contrast-enhanced voiding urosonography; vesicoureteral reflux; quantitative grading system; cut-off value

Introduction

Vesicoureteral reflux (VUR) is the most common anomaly in paediatric urinary infection (UTI) cases, with an incidence rate of 30-50% [1]. The primary concern of VUR is the risk of recurrent pyelonephritis and secondary long-term clinical consequences of renal scarring, and end stage renal disease [2]. Nowadays, VUR management is based on two different approaches: surgical correction techniques including endoscopic injection of bulking agents and ureteric reimplantation (open or laparoscopic), or conservative treatments consisting of hygienic habit recommendations and continuous antibiotic administration of prophylaxis (CAP) [3-5]. Despite the high incidence rate of VUR and the increasing number of publications regarding its clinical management, there is still a lack of consensus regarding the optimal treatment for VUR aiming to prevent renal injury progression [3,4]. In addition, the grade of VUR is closely related to the prognosis of VUR, manifested as the fact that the higher the VUR grade, the more renal scarring is present [6]. Therefore, the VUR management should be individual-

Received 15.11.2019 Accepted 08.03.2020
Med Ultrason
2020, Vol. 22, No 3, 287-292
Corresponding author: Wei Zhang
Department of Medical Ultrasound,
Tongji Hospital, Tongji Medical College,
Huazhong University of Science and Technology
1095 Jiefang Avenue, Qiaokou District,
Wuhan City, Hubei Province, China, 430030
E-mail: zhangweitjh@qq.com
Phone: +86 15926289141

ized based on the grade and manifestations of the VUR [3-5] and the diagnosis and evaluation of VUR should be as objective and precise as possible.

At present, there are three methods to detect VUR: 1) voiding cystoureterography (VCUG); 2) radionuclide cystography (RNC); and 3) contrast-enhanced voiding urosonography (ceVUS) [7]. VCUG was the most common approach and is recognized as the gold standard for the diagnosis of VUR; however, ceVUS has been approved by the FDA and is becoming a more widespread available technique [8]. As a dynamic imaging technique with high sensitivity and non-radioactivity, ceVUS can be repeatedly applied for the detection and follow-up of VUR in children [9-12]. With the increasing implementation and without severe adverse effects reported, ceVUS has become a reliable and preferred method for VUR diagnosis in children [13-16].

According to the morphology of pelvicalyceal system and ureteral imaging, VUR can be divided into five grades, which is recognized as the standard for evaluating the degree of VUR (Table I) [10,17,18]. In practice, however, the grading diagnosis subjectively depends on the experience of the observers who identify characteristics from images. Discrepancies involving VUR of grade II-V exist due to the lack of quantitative measurements, which leads to different interpretations of the same image and inconsistency in terms of the grading diagnosis of the same patient [19,20]. Hence, current ceVUS grading criteria for VUR diagnosis is still a qualitative system, which is not able to meet the requirements of quantitative and precise assessment of VUR.

To the best of our knowledge, there is no established quantitative parameter for VUR grading diagnosis at present. This study aimed to analyse ceVUS images and evaluate a specific quantitative indicator for the VUR grading system.

Material and methods

Patients

Retrospective analysis of ceVUS images of VUR of grade II-V from a total of 63 children between June 2017 and September 2018 was performed in this study. All the children included in this study were recruited from the Paediatric Nephrology section of the Department of Pediatrics at Tongji hospital. Recurrent urinary tract infection (UTIs) patients aged between 0-14 years old diagnosed with VUR by ceVUS were included in this study. Patients with a history of renal surgery, kidney duplication and atrophic kidney were excluded. The present study was approved by the Institutional Review Board of Tongji Hospital and was performed in accordance with the principles of the World Medical Association Declaration of Helsinki, revised in 2000, Edinburgh. Written informed consent was obtained from a parent of each child prior to ceVUS imaging.

Contrast-enhanced voiding urosonography

CeVUS was performed using the LOGIQ E9 ultrasound system (GE Healthcare, Wauwatosa, WI, USA) with a 4 MHz convex array harmonic imaging probe as described in our preliminary study [17]. Imaging sections with maximum cross-section area and the presence of kidney hilum were selected during ceVUS. The ceVUS procedures were performed by the same radiologist with 9 years of experience in ultrasound and 3 years of experience in evaluating VUR by ultrasound. Two experienced radiologists completed the VUR grading diagnosis of the ceVUS imaging independently. Both radiologists had three years of clinical experience in VUR diagnosis by ultrasound. The imaging of the highest grade of VUR diagnosis which mostly occurred during the voiding period, was selected for analysis in each child. The grading diagnoses were conducted individually by the two radiologists and made on the same day based on the aforementioned diagnosis standards [17]. All VUR grade diagnoses were made by the consensus of the two radiologists when discrepancies occurred.

The images during a period of maximum pelvicalyceal flux were selected for analyses in each child. Based on the imaging section with a maximum cross-section area and the presence of kidney hilum, the sectional area of kidney (SAK) and the sectional area of pelvicalyceal sys-

Table I. The grading system of vesicoureteral reflux [10,17,18]

Grade	Description
Grade I	Contrast agents present only in the ureter
Grade II	Contrast agents present in the ureter and renal pelvis without dilation of the renal pelvis
Grade III	Contrast agents present in the ureter and renal pelvis with mild dilation of the renal pelvis
Grade IV	Contrast agents present in the ureter and renal pelvis with dilations of the renal pelvis and calyces as well as clear calyceal fornices
Grade V	Contrast agents present in the ureter and pelvis with gross dilation of the renal pelvis and calyces as well as disappearance of calyceal fornices and tortuous expansion of the ureter

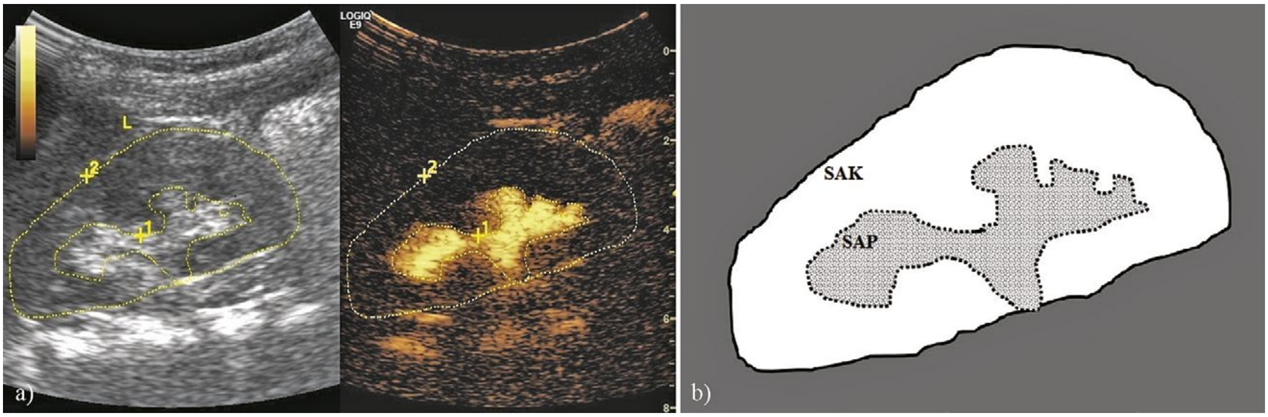

Fig 1. CeVUS images of left kidney based on the imaging section with maximum cross-section area and the presence of kidney hilum: contrast agent presented in the renal pelvis and calyces. a) the sectional area of renal pelvicalyceal system (SAP, measurement 1) and the cross-section area of kidney (SAK, measurement 2) were measured. b) the sketch of ceVUS imaging: the areas enclosed by the solid line & the dotted line were SAK & SAP, separately.

tem (SAP) were measured from ceVUS images of VUR of grade II-V. All the measurements were performed by a trained pediatrician, who was blind to the grading diagnosis of ceVUS. The sectional area ratio (SAR) of SAP to SAK were calculated subsequently (fig 1).

Statistical analysis

SPSS-19 (SPSS Inc., Chicago, Illinois, USA) was used for all statistical analyses. Continuous variables were presented as the mean ± standard deviation (SD) and were analysed using one-way analysis of variance (ANOVA). The diagnostic performances of indicator SAR to identify VUR grades were evaluated using receiver operating characteristic curve (ROC) analysis. The maximum Youden Index defined as sensitivity+specificity−1 was used to determine the optimal cut-off values [21]. Diagnostic performance was classified as low (the area under the ROC curve (AUC) = 0.50–0.70), moderate (AUC = 0.70–0.90), or high (AUC = 0.90–1.0) [22]. The threshold level for statistical significance was $p<0.05$. The consistency between qualitative grading diagnoses and quantitative grading diagnoses of VUR was calculated using the weighted kappa coefficient and intraclass correlation coefficient (ICC) [23]. Kappa and ICC values were interpreted using the following criteria: ≤0.20, poor; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, good; 0.81–1.00, very good [24].

Results

Clinical features

A total of 63 patients with 126 PelviUreteral Units (PUUs) (24 patients with left unilateral VUR, 14 patients with right unilateral VUR and 28 patients with bilateral VUR) were enrolled in this study. Of these patients, 34 were female and 29 were male with a mean age of 32.5±32.7 months (0.5-155 months). According to the VUR grading criteria, VUR of grade I does not involve the renal pelvis, thus 35 PUUs without VUR and 2 PUUs with VUR of grade 1 were excluded. Finally, 89 PUUs with VUR of grade II-V were included for analyses. The results of SAR of PUUs with VUR of grade II-V were presented in Table II.

Relationship between different grades of VUR assessed by SAR

There were statistically significant differences in SAR between different grades of VUR (II *vs* III $p<0.01$, III *vs* IV $p<0.01$, IV *vs* V $p<0.01$). The SAR value increased significantly along with the increase of VUR grade (fig 2).

Diagnostic performance of SAR to identify VUR grades

The AUC values of SAR for differentiating VUR of grade II *vs* III and III *vs* IV were 0.967 and 0.943, respectively, corresponding to a high diagnostic ability. The AUC value of SAR for differentiating VUR of grade IV from V was 0.865, corresponding to a moderate diagnostic ability.

Table II. Results of sectional area ratio (SAR) of PUUs with VUR of grade II-V

VUR grade of PUUs	SAR
II (n=12)	9.92±3.73 (5.05-17.29)
III (n=25)	24.69±7.00 (12.98-36.54)
IV (n=30)	44.94±10.42 (22.20-64.92)
V (n=22)	67.17±17.81 (39.39-100.00)

Values are expressed as average±standard deviation (minimum-maximum). VUR, vesicoureteral reflux; PUUs, PelviUreteral units

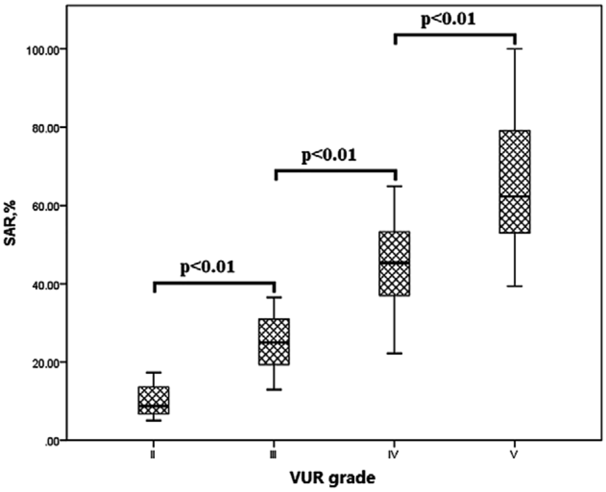

Fig 2. Relationship between different grades of VUR assessed by sectional area ratio (SAR).

For diagnosing VUR of grade II *vs* III, III *vs* IV and IV *vs* V, the optimal cut-off values of SAR were 14.3%, 34.9% and 51.0%, respectively. The sensitivity, specificity and Youden Index of each optimal cut-off value are presented in Table III.

Consistency between qualitative grading system and quantitative grading system based on the optimal cut-off values of SAR

Cases of VUR evaluated by quantitative and qualitative grading system were presented in Table IV. Weighted kappa coefficient was 0.695 ($p<0.001$) and intraclass correlation coefficient (ICC) was 0.82 ($p<0.001$), demonstrating excellent grading consistency of VUR between qualitative and quantitative grading systems.

Discussion

To the best of our knowledge, this is the first study aiming at establishing a quantitative VUR grading system by ceVUS. In order to improve the qualitative VUR grading system currently used in clinical practice, this study explored the feasibility and accuracy of a quantitative indicator SAR, generated from the ceVUS imaging of VUR involved kidney. With AUC of SAR for distinguishing VUR of grade II *vs* III, III *vs* IV and IV *vs* V being 0.967, 0.943, 0.865, respectively, we found that SAR had excellent diagnostic performance for grading VUR. Based on the optimal cut-off values of SAR obtained in the present study, a continuous numerical grading system was available for VUR evaluation. With weighted kappa coefficient value of 0.695 ($p<0.001$) and ICC value of 0.82 ($p<0.001$), the quantitative grading system showed excellent consistency with currently qualitative grading system of VUR.In addition, the indicator SAR can make the assessment of VUR more objective and precise. Moreover, the acquisition of SAR is not complicated, since most commercial ultrasound instruments are equipped with the function of area ratio calculation. Therefore, the quantitative VUR grading system established in this study is worthy of further testing in clinical applications.

CeVUS is a dynamic and continuous imaging method, while VUR may occur transiently and the imaging features of the pelvicalyceal system change with time during the ceVUS procedure. In this study, the images during the period of maximum pelvicalyceal flux mostly occurring during the voiding period were selected for analyses in each child. Therefore, there was no case in this study where the grade of VUR was underestimated. We also suggest that ceVUS scanning of unilateral kidney should be continuously performed and last till the end of voiding, then the dynamic and intermittent occurrence of the contrast agent reflux will not affect the results of ceVUS.

The results of ultrasound examination depend partially on the experience of the operator in terms of scanning and image reading. The accuracy of ultrasound diagnosis

Table III. Optimal sectional area ratio (SAR) cut-off values of VUR grading system

VUR grade of PUUs	optimal cut-off value	sensitivity	specificity	Youden Index
II *vs* III	14.30%	92.00%	91.70%	0.84
III *vs* IV	34.90%	83.30%	96.00%	0.79
IV *vs* V	51.00%	86.40%	73.30%	0.6

VUR, vesicoureteral Reflux; PUUs, PelviUreteral units

Table IV. Cases of vesicoureteral reflux evaluated by a quantitative and qualitative grading system

	Conventional grading system based on qualitative diagnosis criteria								
	Grade	**II**	**III**	**Grade**	**III**	**IV**	**Grade**	**IV**	**V**
Quantitative grading system	II	11	2	III	22	5	IV	17	3
	III	1	22	IV	1	17	V	8	19

can be affected by the subjectivity of the operator. Therefore, the establishment of standardized operating procedures is crucial to improve the repeatability and reliability of ceVUS diagnosis [25]. In this study, when ceVUS imaging was performed and measured, the imaging section with maximum cross-section area and the presence of kidney hilum was used. We recommend this imaging section to be applied to the two-dimensional scanning of all ceVUS. On the other hand, the use of real-time, three-dimensional imaging may reduce operator dependence and improve the robustness of the quantitative indicator. In addition, the establishment of the quantitative cut-off values could not only make the VUR grading diagnosis easier and open new doors more objectively, but also help to reduce the interrater inconsistency as well as to promote intelligent automatic measurements and analysis based on machine learning in the future. An objective and quantitative grading system of VUR can also be applied to decrease the dependence on expertise and promote more extensive application and research of ceVUS, especially in developing regions.

Our study has several limitations. Firstly, the results should be interpreted with caution because of the small number of patients, the single-center design and performance inherently limited by two-dimensional ultrasound. Further multi-center prospective studies are required to systematically investigate the reliability and accuracy of the VUR quantitative grading system. Secondly, the quantitative grading threshold was obtained only based on image analysis: however, a quantitative grading system related to renal function damage and prognosis of VUR is more instructive to clinical management of VUR. Therefore, the results of this pilot study are preliminary. Further clinical studies incorporating renal function damage and prognosis of VUR are required to refine the quantitative grading system.

Conclusions

The numerical indicator SAR calculated from ceVUS imaging may be used to establish a quantitative grading system of VUR with excellent diagnostic performance and can potentially serve as a reliable tool for the evaluation and follow-up of VUR.

Conflict of interest: None

References

1. Cooper CS, Austin JC. Vesicoureteral reflux: who benefits from surgery? Urol Clin North Am 2004;31:535-541.
2. Roussey-Kesler G, Gadjos V, Idres N, 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;179:674-679.
3. Arlen AM, Cooper CS. Controversies in the Management of Vesicoureteral Reflux. Curr Urol Rep 2015;16:64.
4. Routh JC, Bogaert GA, Kaefer M, et al. Vesicoureteral reflux: current trends in diagnosis, screening, and treatment. Eur Urol 2012;61:773-782.
5. Wang HH, Gbadegesin RA, Foreman JW, et al. Efficacy of antibiotic prophylaxis in children with vesicoureteral reflux: systematic review and meta-analysis. J Urol 2015;193:963-969.
6. Yilmaz I, Peru H, Yilmaz FH, et al. Association of vesicoureteral reflux and renal scarring in urinary tract infections. Arch Argent Pediatr 2018;116:e542-e547.
7. Lim R. Vesicoureteral reflux and urinary tract infection: evolving practices and current controversies in pediatric imaging. AJR Am J Roentgenol 2009;192:1197-1208.
8. Correas JM, Anglicheau D, Joly D, Gennisson JL, Tanter M, Hélénon O. Ultrasound-based imaging methods of the kidney-recent developments. Kidney Int 2016;90:1199-1210.
9. Darge K, Zieger B, Rohrschneider W, Ghods S, Wunsch R, Troeger J. Contrast-enhanced harmonic imaging for the diagnosis of vesicoureteral reflux in pediatric patients. AJR Am J Roentgenol 2001;177:1411-1415.
10. Darge K. Diagnosis of vesicoureteral reflux with ultrasonography. Pediatr Nephrol 2002;17:52-60.
11. Duran C, del Riego J, Riera L, Martin C, Serrano C, Palaña P. Voiding urosonography including urethrosonography: high-quality examinations with an optimised procedure using a second-generation US contrast agent. Pediatr Radiol 2012;42:660-667.
12. Papadopoulou F, Ntoulia A, Siomou E, Darge K. Contrast-enhanced voiding urosonography with intravesical administration of a second-generation ultrasound contrast agent for diagnosis of vesicoureteral reflux: prospective evaluation of contrast safety in 1,010 children. Pediatr Radiol 2014;44:719-728.
13. Papadopoulou F, Anthopoulou A, Siomou E, et al. Harmonic voiding urosonography with a second-generation contrast agent for the diagnosis of vesicoureteral reflux. Pediatr Radiol 2009;39:239-244.
14. Darge K, Back SJ. Invited Commentary: Prime Time for Contrast-enhanced Voiding Urosonography after Approval of a US Contrast Agent for Children From. Radiographics 2017;37:1869-1871.
15. Duran C, Beltran VP, Gonzalez A, Gomez C, Riego JD. Contrast-enhanced Voiding Urosonography for Vesicoureteral Reflux Diagnosis in Children. Radiographics. 2017;37:1854-1869.
16. Darge K. Voiding urosonography with US contrast agent for the diagnosis of vesicoureteric reflux in children: an update. Pediatr Radiol 2010;40:956-962.
17. Zhang W, Cai B, Zhang X, Zhou J, Qiu L, Yi. Contrast-enhanced voiding urosonography with intravesical administration of ultrasound contrast agent for the diag-

nosis of pediatric vesicoureteral reflux. Exp Ther Med 2018;16:4546-4552.
18. 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;15:105-109.
19. Schaeffer AJ, Greenfield SP, Ivanova A, et al. Reliability of grading of vesicoureteral reflux and other findings on voiding cystourethrography. J Pediatr Urol 2017;13:192-198.
20. Celebi S, Ozaydin S, Bastas CB, et al. Reliability of the Grading System for Voiding Cystourethrograms in the Management of Vesicoureteral Reflux: An Interrater Comparison. Adv Urol 2016;2016:1684190.
21. Youden WJ. Index for rating diagnostic tests. Cancer 1950;3:32-35.
22. Swets JA. Measuring the accuracy of diagnostic systems. Science 1988;240:1285-1293.
23. Mehta S, Bastero-Caballero RF, Sun Y, et al. Performance of intraclass correlation coefficient (ICC) as a reliability index under various distributions in scale reliability studies. Stat Med 2018;37:2734-2752.
24. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ Psychol Meas 1973;33:613-619.
25. Wong LS, Tse KS, Fan TW, et al. Voiding urosonography with second-generation ultrasound contrast versus micturating cystourethrography in the diagnosis of vesicoureteric reflux. Eur J Pediatr 2014;173:1095-1101.