Outcome Predictive Values of the Society of Fetal Urology (SFU) Grading System and Urinary Tract Dilation (UTD) Classification in Patients with High-Grade Ureteropelvic Junction Obstruction-like Prenatal Hydronephrosis

John Ivan S. Alonzo, MD and David T. Bolong, MD, FPUA

Section of Urology, Department of Surgery, University of Santo Tomas Hospital

Objectives: This study aimed to determine the surgical predictive value of both SFU and UTD classifications in a specific subset of patients presenting with high-grade (SFU 3 and 4; UTD P2 and P3) UPJO-like hydronephrosis on prenatal ultrasound. Furthermore, this study also aimed to determine the likelihood of spontaneous resolution of high grade UPJO-like hydronphrosis based on both grading systems.

Methods: Patients who presented with high-grade hydronephrosis on prenatal ultrasound based on the SFU grading system (Grades 3 and 4) were included in this study. The prenatal renal ultrasounds of these patients were reclassified by a single interpreter according to the UTD classification. Logistic regression was used to test the predictive value of SFU and UTD; ROC curves were plotted accordingly. Kaplan-Meier curves were used to model time to operation and mean time to operation was computed with a 95% confidence interval. Breslow Test was used to determine significant differences in survival curves across the different SFU grades and UTD classifications.

Results: Of the 163 patients in the database who presented with prenatal hydronephrosis, 25 patients presented with high-grade UPJO-like hydronephrosis (50 renal units). Logistic regression revealed that the SFU grading system was able to explain only 18.7% of the variance of the occurrence of pyeloplasty, thus, was a poor predictor of the occurrence of surgery. In contrast, logistic regression of the UTD classification was able to explain 47.3% of the occurrence of pyeloplasy with an accuracy of 86% making it a good predictor of surgical intervention. Both SFU and UTD classifications were poor predictors of spontaneous resolution. Mean time to pyeloplasty from the time of diagnosis was 2.98 years (95% CI: 2.45-3.53) Kaplan-Meier curve analysis for the time of pyeloplasty for the SFU grading system revealed no significant difference in the time to operation among the different SFU grades (p=0.110) while for the UTD classification, there was a significant difference in time to pyeloplasty across the different UTD classes with the higher classes correlating to a shorter time to pyeloplasty. (p<0.05)

Conclusion: The UTD classification system is a good predictor of surgery in patients presenting with high-grade UPJO-like hydronephrosis on prenatal ultrasound with a predictive accuracy of 86%. High-grade hydronephrosis based on the UTD classification equates to a shorter time to surgical intervention from the time of diagnosis necessitating closer follow-up of these patients.

Key words: urinary tract dilation, prenatal hydronphrosis

Introduction

Prenatal hydronephrosis is the most commonly diagnosed congenital anomaly comprising 1-5% of pregnancies.[1-5] Of the causes of prenatal hydronephrosis, "UPJO-like" hydronephrosis (those with dilation of the renal collecting system without ureteral dilation) is the most common comprising 50% of accounted cases.[6] Earlier schools of thought advocated that early repair of UPJO (at 2-4 weeks of age) was more effective for protection of renal function compared to later repair.[7-8] However, more recent studies reported that waiting for spontaneous resolution of the UPJO may not cause loss of renal function with a spontaneous resolution rate of 70-80%.[9-12] The mechanism of spontaneous resolution is still not certain but postulated reasons include maturation of the smooth muscle peristaltic mechanism of the proximal ureter and pelvis or that milder hydronephrosis may be brought about by high bladder pressures in immature boys.[12] The clinical challenge lies in the fact that the time to resolution or risk of progression of high grade hydronephrosis is difficult to ascertain necessitating continuous monitoring as progression of hydronephrosis, development of recurrent UTI, and development of signs of decreasing renal function would necessitate surgical intervention.

The Society for Fetal Urology (SFU) grading system is the most widely used grading system to assess antenatal hydronephrosis.[6] However, the SFU grading system has several limitations, in particular, 1) the inability to distinguish between segmental and diffuse parenchymal thinning, and 2) the inability to clearly differentiate Grade 3 from Grade 4 hydronephrosis.[13] In addition, the SFU grading system was found to have only moderate interrater reliability.[14] In lieu of this, the Urinary Tract Dilation (UTD) classification system was devised in 2014 to address these limitations. The UTD classification system includes the following parameters: 1) anterior-posterior renal pelvic diameter (APRPD); 2) calyceal dilation; 3) renal parenchymal thickness; 4) renal parenchymal appearance; 5) bladder abnormalities; and 6) ureteral abnormalities.[15] A study evaluating the UTD classification system concluded that it is reliable for evaluating postnatal hydronephrosis and is valid in predicting surgical intervention while the SFU grading system was predictive of likelihood of resolution.[13] However, a recent study revealed that both grading systems equally allowed for proper risk-stratification and prediction of clinical outcomes based on initial ultrasound, correctly separating almost all infants who underwent surgery or developed febrile UTI from those with mild hydronephrosis who were managed non-surgically.[16]

This study aimed to determine the surgical predictive value of both SFU and UTD classifications in a specific subset of patients presenting with high-grade (SFU 3 and 4; UTD P2 and P3) UPJO-like hydronephrosis on prenatal ultrasound. Furthermore, this study also aims determine the likelihood of spontaneous resolution of high grade UPJO-like hydronphrosis based on both grading systems.

Materials and Methods

Patient database search was done from January 2014 to August 2016 for patients who presented with high grade hydronephrosis (SFU Grade 3 and 4) on prenatal ultrasound. The authors then selected those whose prenatal renal ultrasounds exhibited "UPJO-like" features (dilation of the renal collecting system without ureteral dilation) to be included in the study. They limited their population to those who are suspected to have ureteropelvic junction obstruction on prenatal ultrasound, excluding those who were suspected to have other congenital renal anomalies (i.e. vesicoureteral reflux, double collecting system, multicystic dysplastic kidneys) or those with only low grade prenatal hydronephrosis based on the SFU grading.

The prenatal renal ultrasound of these patients were reclassified according to the UTD classification by a single, blinded-interpreter, taking into account the APRPD and parenchymal thickness of each affected renal unit. Repeat renal ultrasounds were done after the 7th day of life to discount the immaturity of the urinary tract and the risk of underestimating the pathologic findings. Baseline diuretic DTPA renal scan were

done to all patients to determine baseline differential glomerular filtration rates on the 8th week of life ensuring maturity of the urinary tract and reliability of results obtained from the study. Voiding cystourethrography was done based on the discretion of the attending physician to rule out the suspicion of vesicoureteral reflux postnatally. All patients received Cefalexin single-dose per orem computed by weight as antibiotic prophylaxis. Repeat renal ultrasound using the SFU grading system was done during the immediate postnatal period (within 4 weeks postpartum), and on 2, 4, and 6 month follow-up. Repeat renal ultrasound beyond 6 months and repeat DTPA renal scans were done upon the discretion of the attending physician.

Pyeloplasty (either open or laparoscopic) was done at any time during the follow-up period if patients developed one of the following: 1) worsening of hydronephrosis on follow-up ultrasonography, 2) recurrent UTI, 3) differential GFR <35% or worsening of differential GFR by >5% of the affected kidney on follow-up renal scan. Patients who did not undergo surgery were noted for either spontaneous resolution or continuous observation.

Logistic regression was used to test the predictive value of SFU and UTD; ROC curves were plotted accordingly. Kaplan-Meier curves were used to model time to operation and mean time to operation was computed with a 95% confidence interval. Breslow Test was used to determine significant differences in survival curves across the different SFU grades and UTD classifications.

Results

Of the 163 patients who presented with prenatal hydronephrosis, 25 patients presented with high-grade UPJO-like hydronephrosis (50 renal units). The study population primarily consisted of males (22 males vs. 3 females.) Fifteen patients presented with high grade hydronephrosis on the left kidney while 10 had high-grade hydronephrosis on the right. Of the cases of high-grade hydronephrosis, 12 patients presented with Grade 4 hydronephrosis while the

rest presented with Grade 3 hydronephrosis on prenatal ultrasound based on the SFU grading system. In contrast, reclassification based on UTD classification revealed that 11 patients had UTD P3, 9 patients had UTD P2 , and 5 had UTD P1. Mean APRPD on the right and and left kidneys were 10.92 mm and 23.13 mm, respectively. On the right kidneys, 20 had normal parenchymal thickness while 5 were thinned out. With regards to the left kidney, 15 had normal parenchymal thickness while 10 were thinned out. (Table-1)

Table 1. Study population demographic data.

Demographic Data (n=25)		
Gender		
Male		22
Female		3
Laterality of High-grade Hydronephrosis		
Right		15
Left		10
SFU Grade		
Grade 3		13
Grade 4		12
Mean A-P Renal Pelvis Diameter (mm)		
Right		10.92
Left		23.13
Parenchymal Thickness	Right	Left
Normal	20	15
Thinned	5	10
UTD Classification		
P1		5
P2		9
P3		11
Outcomes		
Surgery(Pyeloplasty)		11
Resolved		10
Still on observation		4
Indication for Surgery		
Worsening hydronephrosis		6
Abnormal differential GFR		3
Recurrent UTI		2

From the study population, 11 patients eventually underwent pyeloplasty, 10 had

spontaneous resolution of high-grade hydronephrosis, while 4 were still undergoing observation. Of the patients who underwent pyeloplasty, 2 were operated on due to recurrent UTI, 3 had a decrease in differential GFR >5% or a differential GFR <35%, and 6 had worsening hydronephrosis on follow-up ultrasonography. (Table-1)

On logistic regression, the SFU grading system was able to explain only 18.7% of the variance of the occurrence of pyeloplasty. Table-2 shows that the SFU grading system was not able to predict the occurrence of pyeloplasty in this specific set of patients, hence, was not a good predictor of pyeloplasty. (Figure-1) In comparison, the UTD classification was able to explain 47.3% of the occurrence of pyeloplasy with an accuracy of 86% making it a good predictor of surgical intervention. (Table-3 & Figure-2)

Figure 1. ROC curve of SFU grading system-pyeloplasty.

Table 2. Classification table for SFU grading system-pyeloplasty.

Observed		Predicted		
		Surgery		Percentage
		No	Yes	Correct
Surgery	No	38	0	100.0
	Yes	12	0	.0
Overall Percentage				76.0

Table 3. Classification table for UTD classification-pyeloplasty.

Observed		Predicted		
		Surgery		Percentage
		No	Yes	Correct
Surgery	No	34	4	89.5
	Yes	3	9	75.0
Overall Percentage				86.0
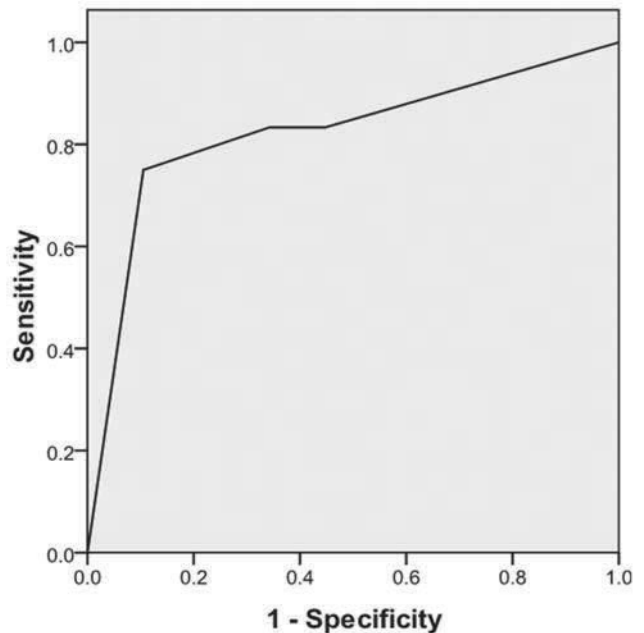
Figure 2. ROC curve of UTD grading system-pyeloplasty.

In terms of spontaneous resolution, the SFU grading system was able to explain only 6.4% of the variance of occurrences with an accuracy of 60%, hence, was a poor predictor of spontaneous resolution. (Table-4 & Figure-3) Likewise, the UTD classi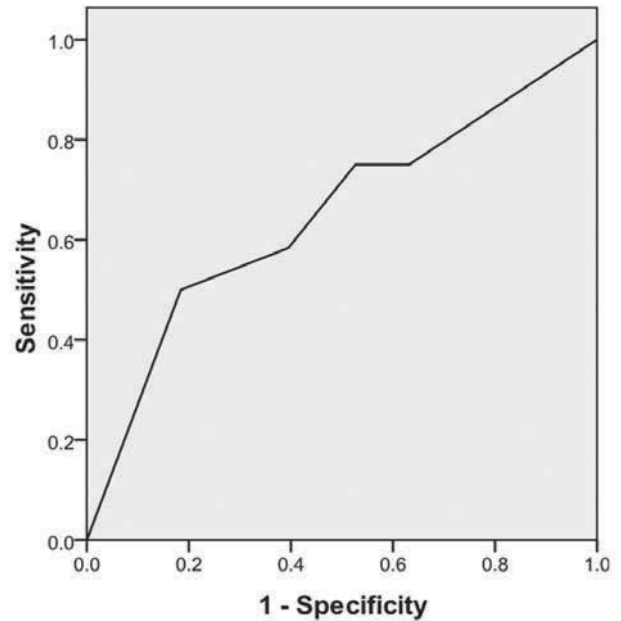fication was able to explain only 12.7% of the variance of the occurrence of spontaneous resolution with an accuracy of 62% which also makes it a poor predictor of spontaneous resolution. (Table-5 & Fugure-4).

Table 4. Classification table for SFU grading system-resolution.

Observed		Predicted		
		Resolved		Percentage
		0	1	Correct
Resolve	0	10	10	61.5
	1	10	14	58.3
Overall Percentage				60.0

Table 5. Classification table for UTD classification-resolution.

Observed		Predicted		
		Resolved		Percentage
		0	1	Correct
Resolve	0	10	16	38.5
	1	3	21	87.5
Overall Percentage				62.0
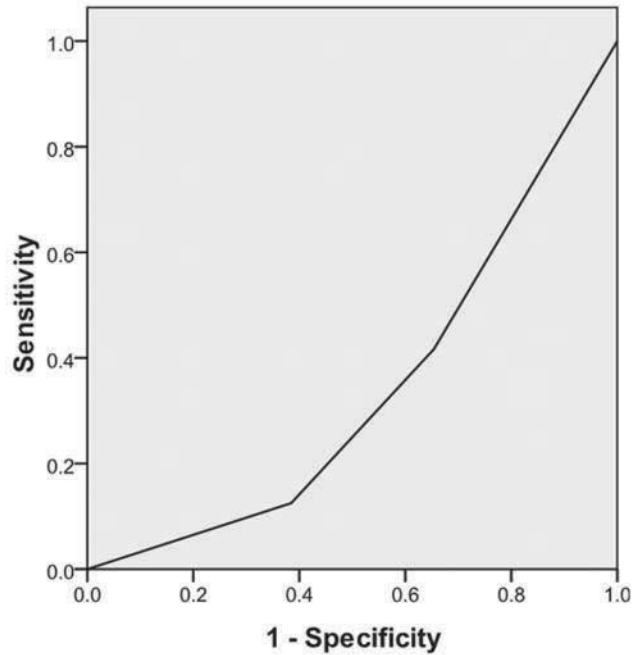
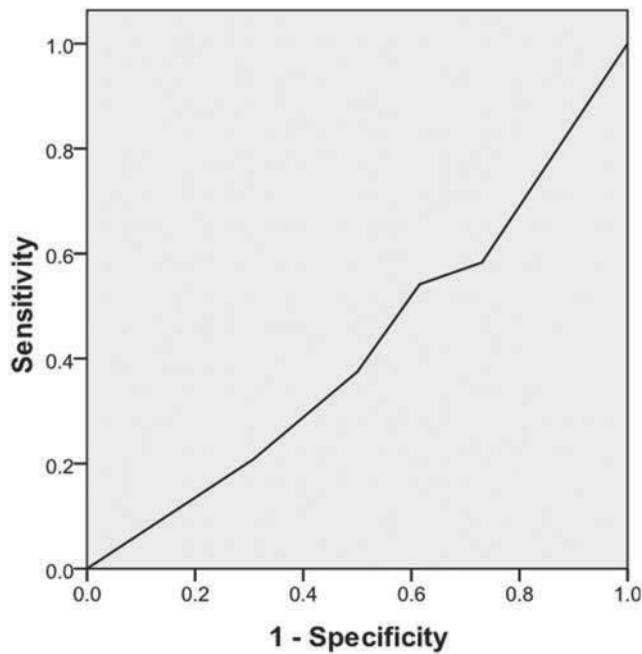**Figure 4**. ROC curve of UTD classification-resolution.

Figure 3. ROC curve of SFU grading system-resolution.

Mean time to pyeloplasty from the time of diagnosis was 2.98 years (95% CI: 2.45-3.53) (Figure-5) Analysis of the Kaplan-Meier curve for the time of pyeloplasty for the SFU grading system revealed no significant difference in the time to operation among the different SFU grades

(p=0.110.) (Figure-6) In contrast, the Kaplan-Meier curve for the UTD classification revealed that there is a significant difference in time to pyeloplasty across the different UTD classes with the higher classes correlating to a shorter time to pyeloplasty (p<0.05). (Figure-7)
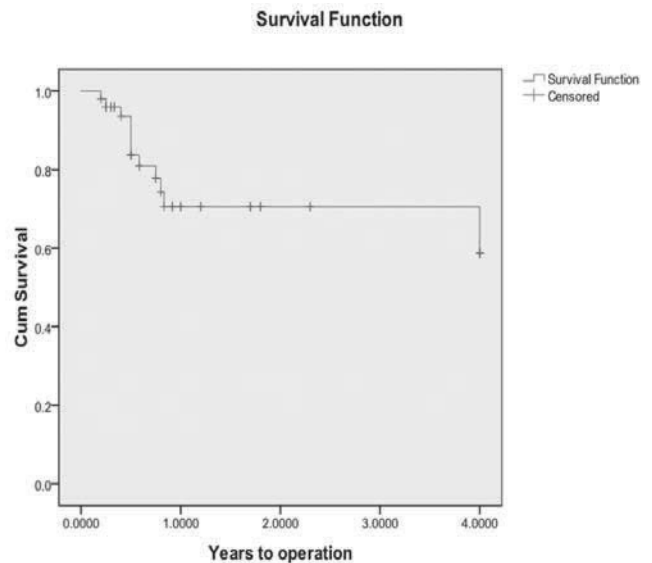

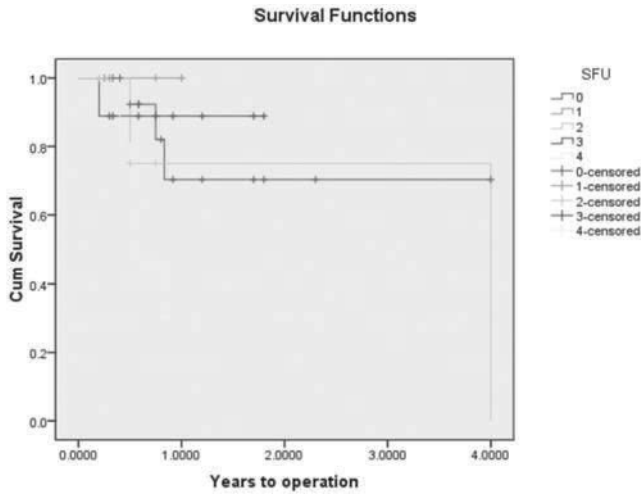**Figure 5**. Kaplan-Meier curve of occurrence of pyeloplasty.

Figure 6. Kaplan-Meier curve of time to pyeloplasty across the different SFU grades.
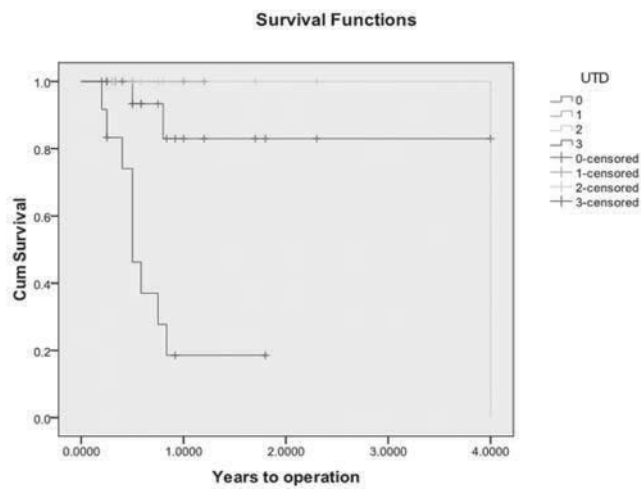

Figure 7. Kaplan-Meier curve of time to pyeloplasty across the different UTD classes.

Discussion

In a study done by Hodhod, et al. involving 490 patients who either had prenatal hydronephrosis or incidentally diagnosed hydronephrosis during the first year of life, they found that the UTD classification system was reliable in the assessment of hydronephrosis with parallel forms of 0.92. Univariate analysis showed that both SFU and UTD systems were significantly associated with spontaneous resolution and surgical intervention. Multivariate

analysis further revealed that SFU grade was an independent predictor of resolution while UTD classification was an independent predictor of surgical intervention.[13] In this study, the UTD classification was likewise a good predictor of surgical intervention based on logistic regression. However, logistic regression of SFU grading revealed that it is a poor predictor of both surgical intervention and resolution. This may be due to the fact that a good number of patients in this study were still under observation at the time of analysis hence contributing to the poor predictability.

In the study by Braga, et al. involving 322 patients with either low or high grade UPJO-like hydronephrosis, the cumulative predictive rates for surgery in patients with high-grade hydronephrosis based SFU and UTD classifications were 50% and 52%, respectively.[16] This is congruent with present findings that the UTD classification was able to predict the need for pyeloplasty in patients with high-grade hydronephrosis with an accuracy of 86%. The SFU grading system was a poor predictor of pyeloplasty in the present study probably owing to the limited sample size.

Of note in the present study is that, based on the UTD classification, as the grade of hydronephrosis becomes higher, the need for surgical intervention from the time of diagnosis becomes shorter. The clinical impact of this is that patients with low-grade hydronephrosis based on the UTD classification (UTD P1) can benefit more with longer follow-up periods compared to those with high-grade hydronephrosis (UTD P2/P3.)

The major limitation of the present study lies in the fact that a specific follow-up period is not well-defined as 4 patients were still being observed. As a result, the predictive value for spontaneous resolution was not established by either classifications. A specific long-term follow-up period may obviate this ensuring that all of the patients would have well-defined outcomes. In addition, the small sample population decreases the impact of the results obtained. This can be addressed in future studies by involving several physicians and institutions who can contribute to the database which will further strengthen the surgical predictive value of both classifications. Lastly, the retrospective nature of this study also

limits the power of the results. A future prospective study on this topic will increase the power of the information derived from the data.

Conclusion

The UTD classification system is a good predictor of surgery in patients presenting with high-grade UPJO-like hydronephrosis on prenatal ultrasound with a predictive accuracy of 86%. High-grade hydronephrosis based on the UTD classification equates to a shorter time to surgical intervention from the time of diagnosis necessitating closer follow-up of these patients.

References

1. Woodward M and Frank D. Postnatal management of antenatal hydronephrosis. BJU Int 2002; 89: 149.

2. Gunn TR, Mora JD and Pease P: Antenatal diagnosis of urinary tract abnormalities by ultrasonography after 28 weeks' gestation: incidence and outcome. Am J Obstet Gynecol 1995; 172: 479.

3. Grandjean H, Larroque D and Levi S. The performance of routine ultrasonographic screening of pregnancies in the Eurofetus study. Am J Obstet Gynecol 1999; 181: 446.

4. Sairam S, Al-Habib A, Sasson S, et al. Natural history of fetal hydronephrosis diagnosed on mid-trimester ultrasound. Ultrasound Obstet Gynecol 2001; 17: 191.

5. Ismaili K, Hall M, Donner C et al: Results of systematic screening for minor degrees of fetal renal pelvis dilatation in an unselected population. Am J Obstet Gynecol 2003; 188: 242.

6. Nguyen HT, Herndon CD, Cooper C, et al. The society for fetal urology consensus statement on the evaluation and management of antenatal hydronephrosis. J Pediatr Urol 2010; 6: 212.

7. King LR et al. The case for immediate pyeloplasty in the neonate with ureteropelvic junction obstruction. J Urol 1984; 132: 636-40.

8. King LR, Hatcher PA. Natural history of fetal and neonatal hydronephrosis. Urology 1990; 35(5): 433-8.

9. Ransley PG et al.The postnatal management of hydronephrosis diagnosed by prenatal ultrasound. J Urol 1990; 144(2 Pt 2):584-587; discussion 593-4.

10. Koff SA. Neonatal management of unilateral hydronephrosis. Role for delayed intervention. Urol Clin North Am 1998; 25(2): 181-6.

11. Ulman I, Jayanthi VR, Koff SA The long-term follow-up of newborns with severe unilateral hydronephrosis initially treated non-operatively. J Urol 2000; 164(3 Pt 2): 1101-5.

12. Peters CA. Congenital ureteropelvic junction obstruction: A pragmatic approach. In. Lima M, Manzoni G. (Eds:) Pediatric Urology: Contemporary Strategies from Fetal Life to Adolescence. Italy: Springer; 2015; 89-101.

13. Hodhod A, Capolicchio JP, Jednak R,El-Sherif E, El-Doray AEA, El-Sherbiny M. Evaluation of urinary tract dilation classification system for grading hydronephrosis. J Urol 2016; 195: 725-30.

14. 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: 1680.

15. Nguyen HT, Benson CB, Bromley B, et al. Multidisciplinary concensus on the classification of prenatal and postnatal urinary tract dilation (UTD classification system). J Pediatr Urol 2014; 10: 982-99.

16. Braga LH, McGrath M, Farrokhyar F, Jegatheeswaran K, Lorenzo AJ. Association of initial Society for Fetal Urology grades and urinary tract dilation risk groups with clinical outcomes in patients with isolated prenatal hydronephrosis.J Urol 2016 (accepted for publication).