INTER-OBSERVER AGREEMENT IN INTERPRETING UROFLOWMETRY MEASUREMENTS IN INDONESIA

[1]I Made Suyadnya, [1]Yacobda Sigumonrong.

[1]Department of Urology, Faculty of Medicine/University of Sumatera Utara, H. Adam Malik General Hospital, Medan.

ABSTRACT

Objective: *We evaluated the comparability and repeatability of analyses based on uroflowmetry reports among urologists in Indonesia.* ***Material & Methods:*** *We assessed the inter-observer agreement when interpreting urodynamic examinations. Four urologists analyzed 20 sets of uroflowmetry data and gave their interpretations of the normality of the curve, reason of abnormality, grade classification, and pattern of the uroflowmetry curves. The consensus among observers was analyzed using the kappa statistic.* ***Results:*** *The kappa values for the analyses of the normality of the curves indicated fair to the moderate agreement. Agreement on the reason of abnormality showed poor to fair consensus. The shape of the flow curve had kappa values ranging from 0.047 to 0.225, indicating poor to fair consensus. Based on grade also showed kappa value from 0.047 to 0.169, indicating a poor agreement.* ***Conclusion:*** *Interpretations of uroflowmetry tracings showed only the poor to a fair agreement despite the normality of the uroflow curve. Variability in interpretation can strongly impact patient treatment. Therefore, further work is needed to standardize the reporting and interpretation of uroflowmetry studies to optimize patient care.*

Keywords: *Uroflowmetry, observer variation, inter-observer.*

ABSTRAK

Tujuan: *Mengevaluasi perbandingan dan kemampuan pengulangan analisis berdasarkan laporan uroflowmetri di antara ahli urologi di Indonesia.* ***Bahan & Cara:*** *Kami menilai kesepakatan antar pengamat ketika menafsirkan pemeriksaan Urodinamik. Empat ahli urologi menganalisis 20 set data uroflowmetri dan memberikan interpretasi mereka tentang kurva normalitas, penyebab abnormalitas, tingkat klasifikasi dan pola kurva uroflowmetri. Konsensus di antara pengamat dianalisis dengan menggunakan statistik Kappa.* ***Hasil:*** *Nilai Kappa untuk analisis kurva normalitas menunjukkan hasil yang cukup untuk kesepakatan moderat. Kesepakatan tentang penyebab abnormalitas menunjukkan konsensus yang buruk hingga wajar. Bentuk aliran kurva memiliki nilai Kappa dari 0.047 sampai 0.225, menunjukkan konsensus yang buruk hingga wajar. Berdasarkan tingkat juga menunjukkan nilai Kappa dari 0.047 ke 0.169, dan menunjukkan kesepakatan yang buruk.* ***Simpulan:*** *Interpretasi uroflowmetri hanya menunjukkan kesepakatan yang buruk hingga cukup meskipun pada kurva uroflowmetri normal. Variabilitas dalam interpretasi sangat mempengaruhi perawatan pasien. Oleh karena itu, pekerjaan lebih lanjut diperlukan untuk menstandarisasi pelaporan dan interpretasi studi uroflowmetri untuk mengoptimalkan perawatan pasien.*

Kata Kunci: *Uroflowmetri, variasi pengamat, antar pengamat.*

Correspondence: I Made Suyadnya; c/o: Department of Urology, Faculty of Medicine/University of Sumatera Utara, H. Adam Malik General Hospital, Jl. Bunga Lau No.17, Kemenangan Tani, Kec. Medan Tuntungan, Kota Medan, Sumatera Utara 20136. Indonesia. Phone: (061)8360143. Email: suyadnya.im@gmail.com.

INTRODUCTION

Uroflowmetry is an indispensable, first-line non-invasive examination for the most patients with suspected LUT dysfunction. Objective, quantitative information, which helps to understand both storage and voiding symptoms, is obtainable.[1-2] A private bathroom is essential to perform this examination. The child is instructed to void when he/she feels a "normal" desire to urinate. The patient who sits to void should have a footrest, supporting their feet, to eliminate the possibility of a non-relaxed pelvic floor. Boys are instructed to aim their flow at a specific point in the funneled receptacle to minimize potential misrepresentations. Afterward, parents are asked if their child's flowmetry pattern was representative of their voiding.[3-4]

The maximum flow rate (Qmax) should be

sustained for >2 sec to eliminate artifacts (straining). If the square of Qmax equals or exceeds the voided volume, that value is considered real.[3] Adequate voided volumes should ≥50% of EBC for age, based on the Koff–Hjalmas equation or that of the MVV measured on the FVC. Voided volumes 20 ml or >10% bladder capacity is considered elevated.[4-6] In 7 years old patient, repetitive PVR >10ml or 6% bladder capacity is regarded as elevated. Ideally, 3 uroflows are representative but 2 will suffice as this maintains accuracy and consistency. First-morning uroflows should be avoided as they may exceed normal voided volumes leading to aberrant flow patterns.[5-6]

The Normal Uroflow: Normal voiding occurs when the bladder outlet relaxes and the detrusor contracts. During a normal detrusor contraction with minimal intraurethral resistance, the normal flow curve is bell-shaped with a high maximum flow rate. (Fig. 1 A). Abnormal shapes exist that are flat [plateau], asymmetric, or have multiple peaks (fluctuating [staccato] and/or intermittent with >1 complete stoppages of flow [interrupted]). (Fig. 1B, C) Although suggestive, these patterns do not predict a specific etiology. A normal flow does not always exclude dysfunction, nor does an abnormal pattern automatically mean LUT dysfunction, as abnormal patterns were found in a small but definite number of an asymptomatic normal patient.[3,7] A minimal number of the normal patient void with flattened or intermittent flow curves; most have a bell-shaped curve.[8]

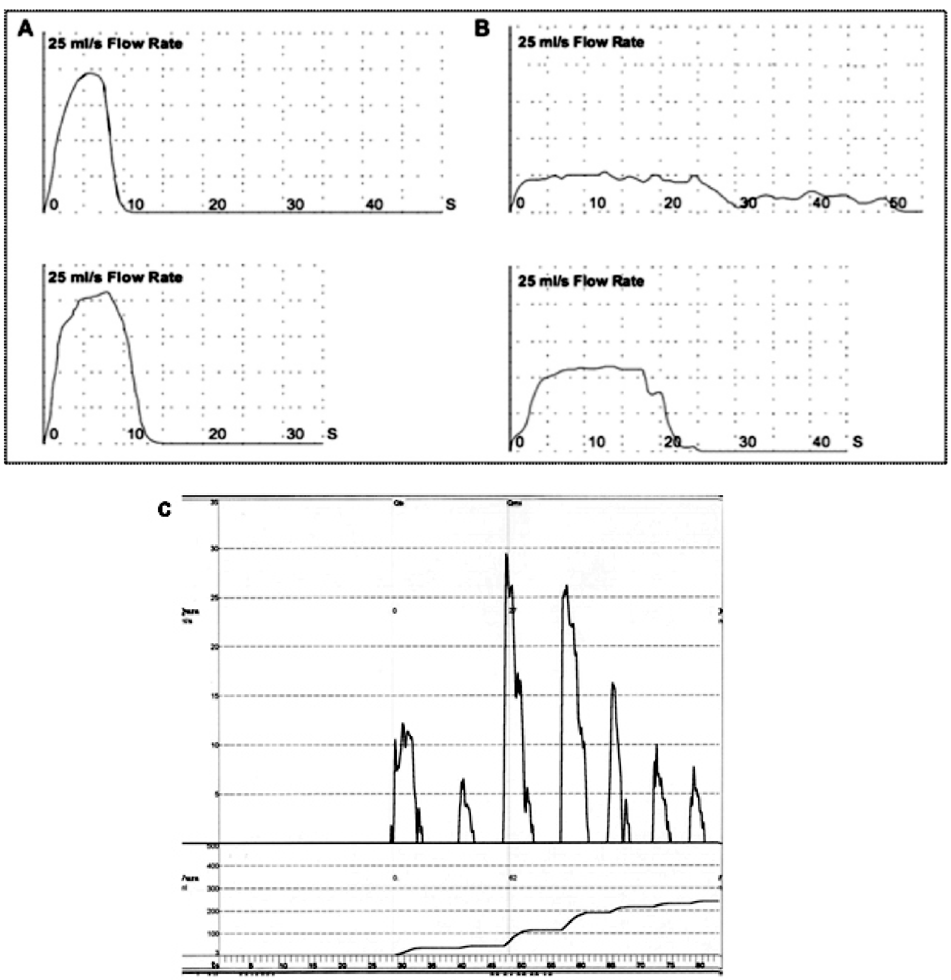

Figure 1. A: Normal (bell-shaped) urinary flow curves of 2 children. B: Flow curves of 2 children with a static, anatomic obstruction; the curve is continuous but the flow is lower than normal and extended in time. C: Interrupted flow curve in a child with either discoordination between bladder and contraction and sphincter relaxation (pelvic floor muscles) or underactive bladder with abdominal straining to empty.[3]

Complicated flow rate patterns may result from fluctuations in detrusor contractility, abdominal straining, or varying degrees of outlet resistance. External urethral sphincter or pelvic floor contraction and relaxation, mechanical compression of the urethra, or meatal stenosis can cause rapid changes in flow rate. Bladder volume may affect uroflowmetry.[3,9] As the volume increases and detrusor muscle fibers stretch, increases in potential detrusor power and work associated with a contraction are needed. This phenomenon is most evident from zero to 150 to 250 ml of filling. At higher volumes, the detrusor may be overstretched decreasing contractility again.[3,9,10]

Therefore, it is theorized Qmax is physiologically dependent on bladder volume. Some have questioned this theorem and are working to identify other factors that may be more important in what determines Qmax. Additionally, rapid changes in flow rate may be artifactual, when the flow rate signal is extracorporeally modified via the interference between the stream and the collecting funnel, the flowmeter, patient movements, or changes in aim of the stream; thus, proper positioning and instruction are necessary.[1,3]

Decreased detrusor power and/or consistently high urethral resistance will result in both a lower flow rate and a smooth flat flow curve. A constriction (e.g., urethral stricture), with reduced luminal size, produces a plateau-like flow curve. The same parameters used to characterize a continuous flow should be applied to the patient with interrupted, or staccato patterns.[3,11] When measuring flow time, the intervals between flow episodes are disregarded. Voiding time is the total duration of micturition, including interruptions.[10-11]

In Indonesia, the uroflowmetry study still limited at the center of urology education. Not every urologist has the same interpretation of uroflowmetry measurement. This study was aimed to evaluate the comparability and repeatability of analyses based on uroflowmetry reports among urologists in Indonesia.

OBJECTIVE

We evaluated the comparability and repeatability of analyses based on uroflowmetry reports among urologists in Indonesia.

MATERIAL & METHODS

Three urologists from three different university hospitals in Indonesia and urologists from Chinese Taipei Hospital assessed the same 20 sets of uroflowmetry data. The cases were randomly chosen by one nurse from H. Adam Malik General Hospital's medical record department. All patient identification data were removed from the uroflowmetry reports, leaving only age and sex. The observers were given a simple answer sheet on which normal or aberrant values were recorded. In the case of an abnormal finding, the observer was asked to state whether the value was pathologically high or low.

Specific suggested uroflowmetry diagnoses were requested regarding the reason of abnormality, the grade of uroflow, and pattern of the curves. We then compared the answer from the Indonesian urologist with an expert from Chinese Taipei. One year later, after the Indonesian urologist attending several workshops and symposiums related to uroflowmetry, we reassessed using another set of uroflowmetry data. Inter-observer agreement was analyzed using kappa statistics as outlined in Table 1. As your information, the authors weren't involved as a rater in this study.

Table 1. Interpretation of kappa statistics as indicators of agreement among observers.

Range	Interpretation
≤ 0.20	Poor agreement
0.21 – 0.40	Fair agreement
0.41 – 0.60	Moderate agreement
0.61 – 0.80	Good agreement
0.81 – 1.00	Very Good agreement

RESULTS

Agreement on the normality of the curves and reason for abnormality. On the first observation of uroflowmetry normality of curves and its reason of abnormality had kappa values ranging from 0.310 to 0.519, indicating a poor to moderate in consensus, and from 0.198 to 0.336, indicating poor to a fair agreement. From this observation, the interpretation of the uroflow curve and the reason of abnormality was seen not quite identical among the observers. After a year, the diagnosed based on that flow curve had increased, particularly on lower range from poor

Table 2. Kappa statistics among observers regarding uroflow normality curves and reason for abnormality.

Observer	Normality of the Curves			Reason of abnormality		
	2	3	4	2	3	4
1	0.310	0.519	0.375	0.241	0.336	0.198

Table 3. Kappa statistics among observers regarding uroflow normality curves and reason of abnormality after a year.

Observer	Normality of the Curves			Reason of abnormality		
	2	3	4	2	3	4
1	0.519	0.600	0.519	0.476	0.602	0.476

Table 4. Kappa statistics among observers regarding uroflow grade and pattern of the curves.

Observer	Grade of the Curve			Pattern of the Curve		
	2	3	4	2	3	4
1	0.169	0.047	0.110	0.225	0.047	0.153

Table 5. Kappa statistics among observers regarding uroflow grade and pattern of the curves after a year.

Observer	Grade of the Curve			Pattern of the Curve		
	2	3	4	2	3	4
1	0.387	0.224	0.381	0.510	0.548	0.582

to moderate in consensus. The kappa values range from 0.476 to 0.602 (Table 3).

Agreement on uroflow grade curves and pattern of the curve. The analyzes of uroflow grade curves varied markedly. The grade of the curve was interpreted with the poor agreement (kappa range 0.310 to 0.519, Table 4). The observers agreed only to a fair degree in the analyses of the pattern of the curve. The next year, the observers gave more similar answers in the grade of the curve and especially, on the pattern of the curve analysis. The kappa values for the grade of the curve indicated fair agreement (kappa range 0.224 to 0.602, Table 5) and moderate agreement for the pattern of the curve (kappa range 0.510-0.582).

DISCUSSION

The uroflowmetry test is an important screening tool in daily clinical urology practice.[4,12] An abnormal uroflowmetry usually requires further sophisticated study to clarify its nature and etiology.[4] However, interpretation of uroflowmetry is usually subjective and empirical.[13] This study aims to evaluate the inter-observer agreement in the screening uroflowmetry. In this study, the uroflowmetry tests were interpreted independently by four urologists, including one urodynamic expert from Chinese Taipei Hospital. to evaluate inter-observer and intra-observer agreement in the screening uroflowmetry, as a comparator three other urologists form our hospital.

ICS (International Continence Society) has introduced the standardization and definitions for each specific type of uroflowmetry curve in 1998 and 2006.[4,9] Typically, bell-shaped, plateau, staccato, or interrupted curves are easy to recognize. Wider fluctuations of the plateau curves made the diagnosis difficult and raised the disagreements.[3,9] Frequently, different interpretations were made for a uroflowmetry curve bearing characteristics of two specific uroflow patterns.

We found poor to the fair agreement among the observer in the interpretation of the uroflowmetry curve and this discrepancy was especially evident in the inconsistency between the uroflowmetry diagnoses of flow patterns obtained in the intra-observer analyses, where no investigator interpreted the same patient flow curve identically on the 2 occasions.[12]

Poor agreement in classifying specific types of 'abnormal' uroflowmetry curve was noted by Chang et al.[14] and Van de Beek et al.[15] who invited 58 urologists to evaluate 25 randomly selected uroflowmetry curves and four of the curves were evaluated twice. Because the invited urologists predicted correctly the actual diagnosis only in 36% of all cases, they urged the urological community to reconsider the diagnostic use of uroflowmetry in daily urological practice. Van de Beek might focus too much on the agreement on specific types of abnormal patterns. ICCS had stated that uroflow patterns only serve as a guide to the existence of specific conditions while not diagnostic itself.[4]

Gacci et al. invited 105 urologists to evaluate 10 selected uroflowmetry curves. They found that there was a substantial agreement for the 'no abnormality' diagnosis (kappa Z 0.72), and that flow curves from healthy men or patients with urethral stricture or benign prostatic obstruction were easily recognizable.[16] Since one of the main purposes of uroflowmetry is to define those who may need an invasive sophisticated study or not, the high agreement rate of 'no abnormality' in children and adults justify the use of uroflowmetry as a screening tool of voiding dysfunction.[12]

In order to see the improvement in agreement among observers and to facilitate comparisons between different studies, the observer was reevaluated in diagnosing the uroflowmetry curve after following several workshops in uroflowmetry. In this study, even slightly, we found improvement in agreement among the observers.

Since the better agreement in 'normality vs abnormality of a uroflowmetry curve' could be reached after a year of "self-evaluation", we hypothesize that better agreement among observers and comparisons between different series could be achieved.[17] It may be wise to classify a uroflowmetry curve as typically normal, typically abnormal, and undetermined, rather than to argue in specifying the abnormal patterns.

Compared to previous studies, our study reevaluated the agreement in the interpretation of uroflowmetry patterns of abnormal uroflowmetry after a year. However, there are several limitations to the study. First, the current study had a small number of uroflowmetry curves enrolled for analysis. Second, the etiology of the abnormal uroflowmetry curves was not confirmed by the urodynamic study. Third, our observer did not attend the same course or workshop, so the variation of interpretation may still exist. Future studies involving more urodynamics and more uroflowmetry curves are warranted to confirm the inter-observer and intra-observer agreement on normality for uroflowmetry curves. Agreement in the interpretation of uroflowmetry was good, particularly in identifying 'no abnormality'. Uroflowmetry is a good screening tool, while not recommended to define the specific type and etiology of abnormal uroflowmetry.

CONCLUSION

Interpretations of uroflowmetry tracings showed only the poor to a fair agreement despite the normality of the uroflow curve. Variability in interpretation can strongly impact patient treatment. Therefore, further work is needed to standardize the reporting and interpretation of uroflowmetry studies to optimize patient care.

REFERENCES

1. Lacovelli V, Farullo G, Turbanti A, Agro EF. Videourodynamic in Children. In: Giovanni M, Del Popolo G, editors. Clinical Urodynamics in Childhood and Adolescence. Springer; 2018. p. 96
2. National Clinical Guideline Centre (UK). The Management of Lower Urinary Tract Symptoms in Men. London: Royal College of Physicians (UK); 2010.
3. Bauer SB, Nijman RJ, Drzewiecki BA, et al. International Children's Continence Society Standardization Report on Urodynamic Studies of the Lower Urinary Tract in Children. Neurourology and Urodynamics. 2015; 34: 640–7.
4. Neveus T, Von Gontard A, Hoebeke P, Hjalmas K, Bauer S, Bower W, et al. The standardization of terminology of lower urinary tract function in children and adolescents: report from the Standardisation Committee of the International Children's Continence Society. J Urol. 2006; 176: 314e24.
5. Pridgeon S, Harding C, Newton D, Pickard R. Clinical evaluation of a simple uroflowmeter for categorization of maximum urinary flow rate. Indian J Urol. 2007; 23(2): 114-8.
6. Austin PF, Bauer SB, Bower W, et al. The standardization of terminology of lower urinary tract function in children and adolescents: update report

from the standardization committee of the international children's continence society. J Urol. 2014; 191: 1863–5.
7. Bower WF, Kwok B, Yeung CK. Variability in normative urine flow rates. J Urol. 2004; 171: 2657–9.
8. Hoebeke P, Bower W, Combs A, et al. Diagnostic evaluation of children with daytime incontinence. J Urol. 2010; 183: 699–703.
9. Norgaard JP, Van Gool JD, Hjalmas K, Djurhuus JC, Hellstrom AL. Standardization and definitions in lower urinary tract dysfunction in children. International Children's Continence Society. Br J Urol. 1998; 81 (Suppl. 3): 1-16
10. C. H. Fry, M. Hussain, C. McCarthy, Y. Ikeda, G.-P. Sui & C. Wu. Recent advances in detrusor muscle function, Scandinavian Journal of Urology and Nephrology. 2004; 38: 215, 20-5.
11. Schaefer W, Abrams P, Liao L, et.al. Good urodynamic practices: uroflowmetry, filling cystometry and pressure-flow studies. In: Cardozo L, Staskin D. Textbook of Female Urology and Urogynecology. CRC Press; 2017. p. 1364-8.
12. Alyami F, Farhat W, Figueroa VH, Romao RL. Utility and cost-effectiveness of uroflowmetry in a busy pediatric urology practice. Can Urol Assoc J. 2014; 8(9-10): 615-8.
13. Venhola M, Reunanen M, Taskinen S, Lahdes-vasama T, Uhari M. Interobserver and Intra-observer agreement in interpreting urodynamic measurements in children. 2003; 169(6): 2344–6.
14. Chang S, Yang SSD. Inter-observer and intra-observer agreement on interpretation of uroflowmetry curves of kindergarten children. J Pediatr Urol. 2008; 422–7.
15. Van de Beek C, Stoevelaar HJ, McDonnell J, Nijs HG, Casparie AF, Janknegt RA. Interpretation of uroflowmetry curves by urologists. J Urol 1997; 157: 164-8.
16. Gacci M, Del PG, Artibani W, Tubaro A, Palli D, Vittori G, et al. Visual assessment of uroflowmetry curves: description and interpretation by urodynamists. World J Urol. 2007; 25: 333-7.
17. Altunay S, Telatar Z, Erogul O, and Aydur E. A new approach to urinary system dynamics problems: Evaluation and classification of uroflowmeter signals using artificial neural networks. Expert Syst. 2009; 4891-5.