Thai Journal of Obstetrics and Gynaecology
September 2022, Vol. 30, No. 5, pp. 321-328.

GYNAECOLOGY

Reliability of the Cervamet and POPstix in Pelvic Organ Prolapse Quantification Measurement

Payada Nutsati, M.D.*,
Suvit Bunyavejchevin, M.D.*,
Purim Ruanphoo, M.D.*,
Noppadol Chaiyasit, M.D.*,
Pairat Tangpornprasert, PhD**,
Chanyaphan Virulsri, PhD**

** Department of Obstetrics and Gynecology, Faculty of medicine, Chulalongkorn University, Bangkok, Thailand*
*** Department of Mechanical Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand*

ABSTRACT

Objectives: The primary objective was to evaluate the reliability of Cervamet and POPstix in pelvic organ prolapse quantification (POP-Q) measurement by 2 different groups of experienced physicians. The secondary objectives were to study time used for examination and patients' discomfort scores.

Materials and Methods: 156 Thai pelvic organ prolapse women attending a gynecology clinic at King Chulalongkorn Memorial Hospital from October 2019 to July 2020 participated in the study. The participants who underwent POP-Q measurement were divided into 4 groups: group1: two staffs using Cervamet, group 2: two staffs using POPstix, group 3: two residents using Cervamet, and group 4: two residents using POPstix. The results were blinded to one another. The time used for examination and the patients' discomfort score were recorded.

Results: The reliabilities (intraclass correlation coefficient) were good in both resident groups: POPstix (0.75-0.98) and Cervamet (0.76-0.98) and in the staffs group: POPstix (0.75-0.95) and Cervamet (0.78-0.96). The median of the time used by the resident group was higher than in the staff group (POPstix: 1.71 vs 1.45 minutes, Cervamet: 3.04 vs 1.99 minutes) ($p < 0.01$). There was no statistical difference in the patients' discomfort scores in all groups.

Conclusion: The Cervamet was a non-disposable tool which equivalent reliability when compared to the POPstix that could be considered as an alternative for POP-Q measurement.

Keywords: pelvic organ prolapse, POP-Q, POPstix, cervamet, reliability.

Correspondence to: *Suvit Bunyavejchevin, M.D., MHS, Department of Obstetrics and Gynecology, Faculty of medicine, Chulalongkorn University, Bangkok, Thailand. E-mail: suvit.b@chula.ac.th*

Received: 26 November 2020, **Revised:** 22 June 2021, **Accepted:** 14 Jyly 2021

การศึกษาความเที่ยงในการประเมินภาวะอุ้งเชิงกรานหย่อนด้วยเครื่องมือ Cervamet and POPstix

ปาญาดา นุตสติ, นายแพทย์สุวิทย์ บุณยะเวชชีวิน, ปุริม เรือนภู่, นพพดล ไชยสิทธิ์, ไพรัช ตั้งพรประเสริฐ, ชัญญาพันธ์ วิรุฬห์ศรี

บทคัดย่อ

วัตถุประสงค์: *วัตถุประสงค์หลักเพื่อศึกษาความเที่ยงของเครื่องมือ Cervamet และ POPstix ในการประเมินตำแหน่ง และระดับความรุนแรงของอวัยวะอุ้งเชิงกรานหย่อนด้วยวิธี pelvic organ prolapse quantification (POP-Q) ระหว่างกลุ่มแพทย์ที่มีประสบการณ์ต่างกันสองกลุ่มในการตรวจ POP-Q วัตถุประสงค์รองเพื่อศึกษาเวลาที่ใช้ตรวจ และความไม่สบายตัวในการตรวจ POP-Q โดยเครื่องมือทั้ง 2 ชนิด*

วัสดุและวิธีการ: *ผู้หญิงภาวะอุ้งเชิงกรานหย่อน จำนวน 156 คน ที่มารับบริการตรวจในคลินิกนรีเวชกรรม ณ โรงพยาบาลจุฬาลงกรณ์ ในช่วงเดือน ตุลาคม 2562 – กรกฎาคม 2563 จะถูกแบ่งออกเป็น 4 กลุ่ม [กลุ่ม 1 = อาจารย์แพทย์ 2 คน ใช้เครื่องมือ Cervamet, กลุ่ม 2 = แพทย์ประจำบ้าน 2 คน ใช้เครื่องมือ POPSTIX, กลุ่ม 3 = อาจารย์แพทย์ 2 คน ใช้เครื่องมือ POPstix, และ กลุ่ม 4 = แพทย์ประจำบ้าน 2 คน ใช้เครื่องมือ POPSTIX] โดยที่ผู้ตรวจจะไม่ทราบผลการตรวจของอีกคน โดยมีการบันทึกเวลาที่ใช้ตรวจ และให้ผู้ป่วยประเมินคะแนนความไม่สบายตัว*

ผลการศึกษา: *ค่าสัมประสิทธิ์สหสัมพันธ์ภายในของเครื่องมืออยู่ในเกณฑ์ดี ทั้งในกลุ่มแพทย์ประจำบ้าน: POPstix (0.75-0.98), Cervamet (0.76-0.98) และในกลุ่มอาจารย์แพทย์ : POPstix (0.75-0.95), Cervamet (0.78-0.96) ค่ามัธยฐานของระยะเวลาที่ใช้ในแพทย์ประจำบ้านมากกว่าในกลุ่มอาจารย์แพทย์ (POPstix: 1.71 vs 1.45 นาที, CERVAMET: 3.04 vs 1.99 นาที) อย่างมีนัยสำคัญทางสถิติ ($p < 0.001$) คะแนนความไม่สบายตัวของ 4 กลุ่มไม่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติ*

สรุป: *เครื่องมือ Cervamet เป็นอุปกรณ์ที่สามารถนำมาใช้ซ้ำได้ และมีความเที่ยงเช่นเท่ากับ POPstix จึงเหมาะสมที่จะเป็นอีกหนึ่งทางเลือกที่จะนำมาใช้ในการประเมินตำแหน่ง และระดับความรุนแรงของอวัยวะอุ้งเชิงกรานหย่อนด้วยวิธี POP-Q*

คำสำคัญ: *ภาวะอุ้งเชิงกรานหย่อน, POP-Q, POPstix, Cervamet, ความเที่ยง*

Introduction

Pelvic organ prolapse quantification (POP-Q) system is an assessment of pelvic organ descent. It has been accepted by the American College of Obstetricians and Gynecologists (ACOG), the International Continent Society (ICS), the American Uro-gynecologic Society (AUGS), and the Society of Gynecologic Surgeons (SGS) as a standard protocol for classification and staging of the disease which is used for the communication between physicians and for research purpose[1]. POP-Q examination contains the 6 anatomical points; Aa = length of anterior vaginal wall 3 cm proximal to external urethral meatus, Ba = length of most distal/dependent part of any portion of anterior vaginal wall just anterior to vaginal cuff/anterior lip of cervix, C = length of most distal edge of cervix, Ap = length of posterior wall 3 cm proximal to hymen, Bp = length of most distal/dependent part on posterior vaginal wall, D = length of posterior fornix or pouch of Douglas which are measured in centimeters by using the hymen lining as a reference point and 3 anatomical markers; genital hiatus (GH), perineal body (PB), total vaginal length (TVL). The leading points that are superior to the hymen will be recorded as the negative numbers. The point below the hymen will be recorded as the positive numbers. The 6 anatomical points were grouped in 3 compartments: anterior, middle, and posterior. Then each compartment was classified as the stage 1-4 according to the distance of prolapse from hymen[2]. These 6 anatomical points and 3 anatomical landmarks are used for staging and for follow-up after treatments. There was a report of good interobserver and intraobsever reliability for POP-Q measurement (intraclass correlation 0.7 - 0.9) using the ring forceps and the instrument for POP-Q measurement[3].

There had been several different instruments introduced for POP-Q measurement such as the fingers, ring forceps, spatula, cotton swabs, etc. but there were differences in the methodology design and parameter used for outcome measurement[3, 4]. In a comparative study of using spatula versus POPstix in POP-Q measurement, POP-Q points using POPstix had a good to excellent reliability except for PB and TVL. The interobserver reliability and the agreement for the staging of POP-Q were better in the POPstix group than in the spatula group[5]. The authors suggested using POPstix device to improve the accuracy of POP-Q measurement[5]. Another study in 2016, comparative study of using digital assessment and quantification of pelvic organ prolapse (DPOP-Q) versus POPstix® in POP-Q measurement, result shown DPOP-Q had high interobserver and intraobserver reliability and significant quicker and less uncomfortable than POPstix®[6]. The International Urogynecological Association (IUGA) has recommended choosing the POPstix as the standard instrument for POP-Q measurement; however, this was not widely adopted because of its relatively high cost and non-reusable materialization. Hence, the reusable equipment can be rather helpful in lowering the long-term cost in POP-Q measurement.

The research team from the Department of Obstetrics and Gynecology, Faculty of Medicine, Chulalongkorn University and the Prosthetic and Orthopedic Implant center, Mechanical Engineering Department, Faculty of Engineering, Chulalongkorn University have developed the Cervamet, a new device to be used in POP-Q evaluation. This instrument is designed as a rod shape with measurement in centimeters and millimeters (ruler-like scale), and with the sliding sleeve bar. This instrument which is made from aluminum, has very light weight, and it can be reused many times. Therefore, in order to implement this new measuring reusable equipment, the psychometric study (reliability) should be conducted and published.

The primary objective of this study was to evaluate the interrater reliability of Cervamet and POPstix in POP-Q measurement. The secondary objectives were to evaluate time used to perform the POP-Q examination and the patients' discomfort between the Cervamet and POPstix.

Materials and Methods

This cross-sectional study was carried out at the Department of Obstetrics and Gynecology, King Chulalongkorn Memorial Hospital, Faculty of Medicine, Chulalongkorn University, Bangkok,

Thailand from October 2019 to July 2020. The research protocol was approved by the Institutional Review Board of the Faculty of Medicine, Chulalongkorn University. Thai women with pelvic organ prolapse stage 1-4 participated in the study. Inclusion criteria were women aged 35 - 65 years and with body mass index (BMI) 18.5 - 30 kg/m^2. Exclusion criteria were pregnancy, pelvic mass, and history of pelvic surgery.

The women were randomly divided by using the block of four technique into 4 groups (Fig. 1), based on examiners and instrument types. Each patient was examined by two physicians consecutively and only one instrument to prevent discomfort from examination; group 1: two staffs using Cervamet and, group 2: two staffs using POPstix, group 3: two residents using Cervamet, and group 4: two residents using POPstix. (Fig. 2) Standardization of the POP-Q measurement technique for all examiners were done before beginning the study follow by Bump et al[7]. The patients were examined in lithotomy position for POP-Q measurement while straining. The POP-Q points were measured by using the POPstix by the standard technique described by Bump et al[7]. The POP-Q anatomical points measured by Cervamet were done by placing the tip of Cervamet (Fig. 3) at each different POP-Q points, then the sliding knob was moved to the hymen level. Switching of the location of the tip of Cervamet and the sliding knob was required in case of anatomical POP-Q point protruding beyond the hymen. The distance from POP-Q point to the hymen was displayed on the scale at lower part of measurement rod (in millimeters).

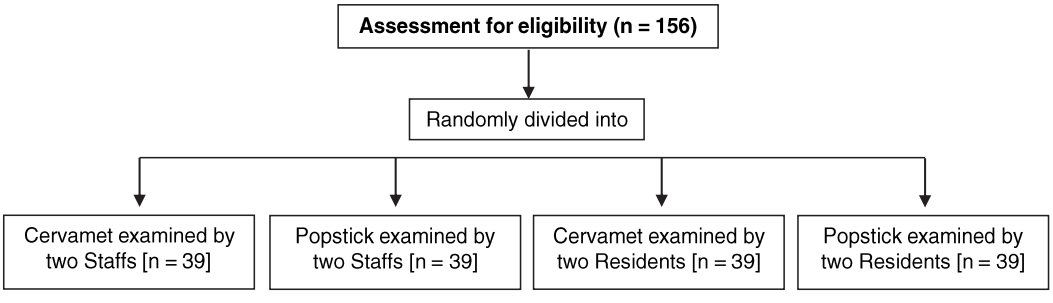

Fig. 1. Flow diagram.

Fig. 2. POPstix.

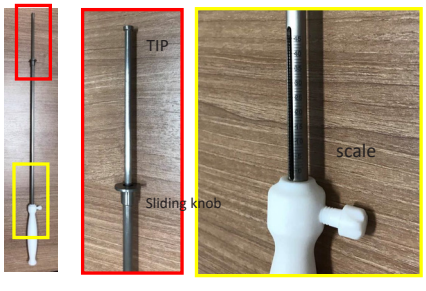

Fig. 3. Cervamet.

Patients' demographic data, such as age, weight, height, menopause status, and the numbers of children, were recorded. The time used by each examiner and the patients' discomfort score (Likert scale 1-5; 1 = no discomfort, 2 = mild, 3 = moderate, 4 = severe, 5 = worst) were recorded by the nurse assistant and such results were blinded to other examiner.

The sample size for the reliability test study was estimated according to the formula for sample size and optimal designs for reliability studies[8]. We used the intraclass correlation coefficient (ICC) of the TVL (POP-Q anatomical distance) with the ICC of 0.8 from our pilot study. From our pilot study (n = 10), the PB had the lowest ICC that was expected for the highest sample sizes. The p1 was 0.8, p0 (minimally acceptable level of reliability) was 0.6, α was 0.05, and β was 0.2. The value of p0 was estimated from acceptable error (20%) of p1 (0.8×0.5 = 0.16 - 0.2). The sample size estimation was 39 cases per group. The total sample was 156 cases.

Statistical analysis was carried out by SPSS software version 22.0. Categorical variables were presented as frequency and percentages, whereas continuous variables were presented as mean and standard deviation. ICC was used to calculate the reliability of Cervamet and POPstix in POP-Q measurement. Student's t test was used for comparison of the discomfort scores and duration of examinations. A p value < 0.05 was considered statistically significant.

Results

One hundred and fifty-six patients participated in this study (39 patients in each group) as shown in Fig. 1. The demographic data of the study population are presented in Table 1. Most patients were aged between 50 to 60 years old (Table 1). The ICC values in the resident group are shown in Table 2. Values were similar in both Cervamet and POPstix (in the range of 0.75-0.98 and 0.76-0.98, respectively). The lowest value of ICC for each instrument were 0.75 (PB) in POPstix and 0.76 (TVL) in Cervamet, respectively.

Table 1. Patients' characteristics (n = 156).

Variable	Total
Age (years)	58.23 ± 7.98
Weight (kg)	58.67 ± 8.66
Height (cm)	154.27 ± 6.70
Body mass index (kg/m^2)	24.60 ± 3.20
Menopause (years)	49.85 ± 4.02
Number of Child	
Nulliparous	15 (9.6)
1 child	28 (17.9)
2 children	52 (33.3)
3 children	35 (22.5)
> 3 children	26 (16.7)
POP stage	
Stage 1	43 (27.6)
Stage 2	50 (32.0)
Stage 3	38 (24.4)
Stage 4	25 (16.0)
Constipation	18 (11.5)
Lifting heavy object	17 (10.9)

Data are presented as mean ± standard deviation and n (%), POP: pelvic organ prolapse

Table 2. Patients' characteristics (n = 156).

	POPstix	Cervamet
POP-Q point, Intraclass correlation coefficient (95% confidence interval)		
Aa	0.98 (0.97, 0.99)	0.98 (0.97, 0.99)
Ba	0.96 (0.93, 0.98)	0.95 (0.92, 0.98)
Ap	0.84 (0.72, 0.91)	0.89 (0.81, 0.94)
Bp	0.77 (0.61, 0.87)	0.94 (0.90, 0.97)
C	0.97 (0.95, 0.98)	0.94 (0.89, 0.97)
D	0.95 (0.91, 0.97)	0.94 (0.89, 0.97)
TVL	0.83 (0.70, 0.91)	0.76 (0.58, 0.86)
GH	0.85 (0.73, 0.92)	0.83 (0.70, 0.91)
PB	0.75 (0.58, 0.86)	0.88 (0.79, 0.94)

POP-Q: pelvic organ prolapse quantification

ICC values in the staff group are displayed in Table 3. Values were also similar in both Cervamet and POPstix (in the range of 0.75-0.94 and 0.78-0.96 respectively). The lowest value of ICC shown 0.75 (TVL) in POPstix and 0.78 (PB) in Cervamet.

Table 3. Intraclass correlation coefficient in staff group.

	POPstix	Cervamet
POP-Q point, Intraclass correlation coefficient (95% confidence interval)		
Aa	0.94 (0.89, 0.97)	0.95 (0.92, 0.98)
Ba	0.91 (0.83, 0.95)	0.96 (0.92, 0.98)
Ap	0.92 (0.86, 0.96)	0.84 (0.72, 0.91)
Bp	0.90 (0.82, 0.95)	0.88 (0.78, 0.93)
C	0.90 (0.82, 0.95)	0.88 (0.78, 0.94)
D	0.95 (0.90, 0.97)	0.96 (0.92, 0.98)
TVL	0.75 (0.56, 0.86)	0.79 (0.63, 0.88)
GH	0.79 (0.64, 0.89)	0.79 (0.64, 0.89)
PB	0.76 (0.59, 0.87)	0.78 (0.63, 0.88)

POP-Q: pelvic organ prolapse quantification

In the resident group, the median time (interquartile range (IQR)) spent in POP-Q measurement using POPstix was significantly shorter than the time spent using Cervamet [1.71 (1.52, 2.30) vs 3.04 (2.84, 3.27) minutes, $p < 0.001$] (Table 4). In the staff group, the median time for POPstix was 1.45 (1.23, 1.70) minutes, which was also significantly lower than the time for Cervamet [1.99 (1.75, 2.37) minutes] (Table 4).

Table 4. Time (minutes) uses for POP-Q measurement.

	POPstix	Cervamet	p value (a)
Residents	1.71 (1.52, 2.30)	3.04 (2.84, 3.27)	< 0.001
Staffs	1.45 (1.23, 1.70)	1.99 (1.75, 2.37)	< 0.001
p value (b)	< 0.001	< 0.001	

Data are presented as median (interquartile range), Data were analyzed with Mann-Whitney U test (a) compare between instrument, (b) compare between physicians' group, POP-Q: pelvic organ prolapse quantification

Comparing between the staff group and resident group, the timed used for POPstix and Cervamet were significantly lower in the staff group (1.45 vs 1.71 min, 1.99 vs 3.04 min). The discomfort scores of all study groups were cross compared as shown in Table 5.

Table 5. Discomfort scores in POP-Q measurement.

	POPstix	Cervamet	p value (a)
Residents group	1 (1, 2)	2 (1, 2)	0.667
Staffs group	2 (1, 3)	2 (1, 2)	0.344
p value (b)	0.212	0.952	

Data are presented by median (interquartile range)
Data were analyzed with Mann-Whitney U test (a) compare between instrument, (b) compare between physician hierarchy
Note: score 1 = lowest discomfort, 5 = highest discomfort
POP-Q: pelvic organ prolapse quantification

Discussion

From our study, the POPstix and Cervamet yielded good to excellent reliability in both resident group and staff group, according to the classification by Koo and Li[9]. The ICC in both groups were more than 0.75 which resulted in good to excellent reliability. We found the good to excellent reliability in all POP-Q measurement points. The time used in POP-Q measurement depends on the experience. The resident spent more time for both equipments than the staff. In both staff group and resident group, the Cervamet required more time when compared to POPstix. We found difference in the discomfort scores of patients after the measurement by POPstix and Cervamet.

The good to excellent reliability of POPstix in our study was similar to the previous study by Hayward et al[5]. They also reported similar lowest reliability of the point of PB and TVL as in this study. This can be explained by the curve surface of the perineum (for PB) and different force to stretch the vagina when measuring the TVL. We found that the Cervamet had excellent reliability similar to POPstix in both staff group and resident group in overall point of POP-Q. We also found that Cervamet needed more time for measuring than the POPstix, but there was no difference in patients' discomfort scores. The Cervamet can be cleaned and is reusable. It is also ecofriendly (while POPstix is made of wood and cannot be reused). Although Cervamet is a time-consuming instrument in POP-Q measurement, it is reusable and having comparable reliability when compared to the internationally recommended POPstix. Moreover, it is ecofriendly and can help lowering the long-term cost of utilizing disposable materials. Therefore, we advocate choosing Cervamet as a reusable option for POP-Q measurement for clinical and research purpose.

The strengths in this study

Our study included the two different experienced groups of examiners (residents, staffs) in order to confirm the reliability in different experienced operators. The Cervamet is the new tool, developed by the engineering staffs that are specialized in medical equipment development from Chulalongkorn University. This instrument prototype was finalized after the design corrections, and it was tested in the models before its application in the human for safety purpose.

The limitation of this study

There was no comparison of the accuracy of POP-Q measurement using Cervamet and POPstix in the same patient due to ethical consideration. The IRB for this study only allowed for one equipment for one patient by two doctors.

Conclusions

With equivalent reliability when compared to the internationally recommended POPstix, the Cervamet

could be considered as an alternative, non-disposable tool for POP-Q measurement.

Acknowledgement

The authors would like to thank Mom Rajawongse Ying Rossalin Kukkanang Foundation for the fund support for Cervamet invention.

Potential conflicts of interest

The authors declare no conflicts of interest.

References

1. Boyd SS, O'Sullivan D, Tulikangas P Use of the Pelvic Organ Quantification System (POP-Q) in published articles of peer-reviewed journals. Int Urogynecol J 2017;28:1719-23.
2. Haylen BT, de Ridder D, Freeman RM, Swift SE, Berghmans B, Lee J, et al. An International Urogynecological Association (IUGA)/International Continence Society (ICS) joint report on the terminology for female pelvic floor dysfunction. Int Urogynecol J 2010;21:5-26.
3. Hall AF, Theofrastous JP, Cundiff GW, Harris RL, Hamilton LF, Swift SE, et al. Interobserver and intraobserver reliability of the proposed International Continence Society, Society of Gynecologic Surgeons, and American Urogynecologic Society pelvic organ prolapse classification system. Am J Obstet Gynecol 1996;175:1467-70.
4. Muir TW, Stepp KJ, Barber MD Adoption of the pelvic organ prolapse quantification system in peer-reviewed literature. Am J Obstet Gynecol 2003;189:1632-5.
5. Hayward L, Wong V, Tomlinson L, Smalldridge J. A prospective interobserver study using the POPstix device, a measuring tool to simplify POPQ measurement. ICS 2009 San Francisco. (Abstract).
6. Thiagamoorthy G, Zacchè M, Cardozo L, Naidu M, Giarenis I, Flint R, et al. Digital assessment and quantification of pelvic organ prolapse (DPOP-Q): a randomised cross-over diagnostic agreement trial. Int Urogynecol J 2016;27:433-7.
7. Bump RC, Mattiasson A, Bø K, Brubaker LP, DeLancey JO, Klarskov P, et al. The standardization of terminology of female pelvic organ prolapse and pelvic floor dysfunction. Am J Obstet Gynecol 1996;175:10-7.
8. Walter SD, Eliasziw M, Donner A Sample size and optimal designs for reliability studies. Stat Med 1998;17:101-10.
9. Koo TK, Li MY A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med 2016;15:155-63.