

Validation of the pelvic organ prolapse quantification (POP-Q) system in left lateral position

G. Alessandro Digesu · Stavros Athanasiou ·
Linda Cardozo · Simon Hill · Vik Khullar

Received: 19 December 2008 / Accepted: 26 March 2009 / Published online: 22 April 2009
© The International Urogynecological Association 2009

Abstract

Introduction and hypothesis The purpose of this study is to examine the inter-observer reliability of the pelvic organ prolapse quantification (POP-Q) system in left lateral position.

Methods Women attending urogynaecology outpatient clinics were examined in the left lateral position using a digital examination and POP-Q. This was repeated separately by a second blinded clinician. The inter-observer agreement was calculated using the Cohen's kappa coefficient. The POP-Q examination was then performed with a woman lying in dorsal lithotomy position. The POP-Q findings in the two positions were compared.

Results Two hundred and eighteen women were recruited. The digital examination had a moderate inter-observer reliability with a kappa value of 0.54. The POP-Q showed a high degree of reliability (0.88). There was a high degree

of correlation between the POP-Q findings in left lateral and lithotomy position ($\rho > 0.95$, $p < 0.001$).

Conclusion The POP-Q in the left lateral position is reliable, easy to perform, acceptable for patients, and is not a time-consuming examination.

Keywords POP-Q · Sims · Position · Pelvic organ prolapse · Vaginal examination · Women

Introduction

Pelvic organ prolapse (POP) is a complex condition that requires a multidisciplinary management approach [1–3].

The communication between physicians is a critical issue, and it plays an important role. Communication is helped by having a reliable method of description and examination of the pathology. To get this aim, it has been recommended that during a vaginal examination, the examiner uses a reliable method to visualise, describe, and quantify the maximum uterovaginal protrusion reported by women during their daily activities and specifies the position of the subject, the type of examination table or chair used, the type of vaginal specula, the retractors used, the type of straining used to develop the prolapse maximally (e.g., Valsalva manoeuvre, cough), the fullness of the bladder, and the contents of the rectum [4]. This minimises misunderstanding between clinicians and allows a more accurate clinical evaluation of POP, which is essential for planning appropriate surgery as well as monitoring treatment outcomes or disease progression.

In 1996, the international continence society (ICS) introduced the pelvic organ prolapse quantification (POP-Q) [4].

The POP-Q system is a descriptive system that contains a series of site-specific measurements of the anterior, apical,

G. A. Digesu · V. Khullar
Department of Urogynaecology, St. Mary's Hospital,
London, UK

S. Athanasiou
Department of Obstetrics and Gynaecology I, Urogynaecology
Unit Alexandra Hospital,
Athens, Greece

L. Cardozo
Department of Urogynaecology, King's College Hospital,
London, UK

S. Hill
Department of Urogynaecology, Royal Blackburn Hospital,
Blackburn, UK

G. A. Digesu (✉)
Via Bottalico 59,
Bari 70124, Italy
e-mail: digesualex@hotmail.com

and posterior pelvic organ support. Prolapse is measured in centimetres relative to the hymeneal ring in relation to six defined points. Points proximal to the hymeneal ring are denoted as negative and points distal, positive. Other measurements that complete the examination are the widths of genital hiatus, perineal body, and total vaginal length [4].

This standard system, which represents a reliable and internationally accepted tool for describing the anatomic position of the pelvic organs [4], has been validated in the dorsal lithotomy, standing and upright position [5–7].

However, at the present time, vaginal examination for urogenital prolapse in the UK is performed either as a digital examination in the supine position or using a Sim's speculum in the left lateral position. There is nothing in the literature about whether this method of evaluation is reliable or valid. The American literature refers to examination in the dorsal lithotomy position and thus is not relevant to the UK [4, 8].

Therefore, the aim of our study was to examine the inter-observer reliability of the POP-Q in the left lateral position.

Material and methods

Women with symptoms of POP and/or lower urinary tract symptoms referred to urogynaecology outpatient clinics of two tertiary referral teaching hospitals were studied. They were asked to void, and the post-micturition residual was checked using trans-abdominal ultrasound [9].

Each woman was then examined with an empty bladder lying in the left lateral position while performing a maximum Valsalva twice by two different clinicians (AD, VK).

Prior to assessment, all women were instructed how to perform a Valsalva manoeuvre until they could reproducibly performed it. They were taught to inhale deeply and bear down as if they were constipated and trying to have a bowel movement. Each clinician initially performed a digital examination in the left lateral position using a four-grading system: none, slight, moderate, and severe [10]. Digitally, the

prolapse was defined as slight, moderate, and severe if the leading edge of the prolapse with maximum Valsalva was felt above, at, or below the introitus, respectively. No speculum was used during digital vaginal examination since this involved palpation with the index finger, of the most caudal leading vaginal wall edge or vault/cervix prolapse.

Subsequently, the clinician re-examined the woman using POP-Q [2]. The POP-Q recorded six defined points around the vagina: two anterior (Aa, Ba), two posterior (Ap, Bp), and two apical (C, D). Each point was expressed as distance in centimetres from the hymen, considered as landmark for reference, with the woman performing maximum Valsalva. They were defined as zero if measured at the level of the introitus and as negative or positive numbers if they were seen cranial or caudal to the introitus, respectively. Instruments used for the determination of the nine quantitative POP-Q measurements included a Sims' speculum and a 10-cm plastic ruler.

The staging system adopted for the POP-Q is shown in Table 1 as described by the ICS [4]. The prolapse is staged by the structure that protrudes the most during forceful straining.

A second experienced clinician, blinded to the findings of the other's examination, reassessed the woman in the same way. Each woman consented to participate in the study.

Finally, time needed for each examination and patient discomfort were recorded. The inter-observer agreement was calculated using the Cohen's kappa coefficient. Cohen's kappa measures the agreement between the evaluations of two raters when both are rating the same object. A value of 1 indicates perfect agreement. A value of 0 indicates that agreement is no better than chance. The strength of agreement is defined as poor, fair, moderate, good, and very good if the value of k is between <0.20 , $0.21–0.40$, $0.41–0.60$, $0.61–0.80$, and $0.81–1.00$, respectively [11].

A group of 50 consecutive women were also examined by the same clinician using a POP-Q with woman lying in lithotomy position.

Table 1 The pelvic organ prolapse ordinal staging system described by the international continence society

Stages	Characteristics
Stage 0	Points Aa, Ap, Ba, and Bp are all at -3 cm, and either point C or D is at lesser than or equal to $-(X-2)$ cm
Stage I	The criteria for stage 0 are not met, and the leading edge of prolapse is lesser than -1 cm
Stage II	Leading edge of prolapse is greater than -1 cm but lesser than $+1$ cm
Stage III	Leading edge of prolapse is greater than $+1$ cm but lesser than $+(X-2)$ cm
Stage IV	Leading edge of prolapse is greater than $+(X-2)$ cm

X=total vaginal length in centimetres in stages 0, III, and IV

Stages I through IV can be subgrouped according to which portion of the lower reproductive tract is the leading edge of the prolapse using the following qualifiers: a=anterior vaginal wall, p=posterior vaginal wall, C=vaginal cuff, Cx=cervix, and Aa, Ap, Ba, Bp, and D for the defined points of measurement

Table 2 Median and 95% confidence intervals (CI) of the POP-Q for the whole population

Pelvic organ prolapse quantification (POP-Q)		
	Median	95% CI
Aa	-1	-1.5; 0
Ba	-0.5	-1; 0
C	-5	-6; -4
Ap	-1.5	-2; -1
Bp	-1	-1.5; 0
D	-6	-7; -5
TVL	7	7; 8
GH	4	1.5; 4.5
Pb	3.5	3; 4
Stage	II	II; II

For the power analysis, we considered a previous study by Swift and Herring [7]. This study showed that since there are five stages (0–4), a one-stage difference between examinations would represent a 20% difference. Therefore, on the basis of this assumption, it was determined that 50 patients would be required to detect a one-stage difference with a power of 0.8 and a p value of 0.05. A Spearman's correlation test was used to compare the six site-specific points of the prolapse examination (Aa, Ba, C, Ap, Bp, D) between the lithotomy and left lateral positions to determine the equality of the two measures. A perfect correlation of 1.0 indicates that the questions are measuring an identical construct. A poor correlation instead suggests that the items are testing different traits. A Cronbach's alpha of ≥ 0.7 has been recommended as acceptable [12, 13]. Wilcoxon signed-rank test was used to compare the assigned stages between the lithotomy and left lateral positions.

Table 3 Agreement between clinicians using POP-Q measurements

Pelvic organ prolapse quantification (POP-Q)				
	Mean difference (cm)	Standard deviation	Limits of agreement	95% CI of the difference
Aa	0.16	0.66	-1.15 to 1.48	-0.19 to 0.22
Ba	0.12	0.62	-1.12 to 1.36	-0.3 to 0.08
C	-0.18	2	-1.18 to 1.82	-0.49 to 0.13
Ap	0.02	0.77	-1.52 to 1.56	-0.22 to 0.2
Bp	0.06	0.7	-1.34 to 1.46	0.06 to 0.7
D	0.49	0.96	-1.43 to 2.41	0.06 to 0.92
TVL	-0.14	0.83	-1.8 to 1.52	-0.42 to 0.14
GH	-0.16	0.53	-1.22 to 0.9	-0.32 to 0.01
Pb	-0.02	0.44	-0.9 to 0.86	-0.16 to 0.11

Table 4 Agreement between clinicians using the current UK prolapse assessment system

	Kappa
Anterior vaginal wall prolapse	0.38
Cervix/vault descent	0.45
Posterior vaginal wall prolapse	0.43

p values of <0.05 were reported as statistically significant. Statistical analysis was performed using t test for paired samples and SPSS version 14.0 (SPSS Inc., Chicago, USA).

This study is a sub-analysis of a project that was approved by the local institutional human research ethics committee.

Results

Two hundred and eighteen women were studied. The mean age was 61 years (range, 41–79 years). The mean weight was 67 kg (range, 45–105 kg), and the median parity was 2 with ranges of 1 to 7. In two patients who were complaining of pelvic pain and had a narrow pubic arch, examination with a ruler was impossible, thus, they were excluded from the study.

Discomfort during the examination was reported as minimal and moderate by 93% (201/216) and 7% (17/216) of women, respectively. Examination was more difficult in the presence of scarring due to previous surgery (43/201) and/or small vaginal introitus ≤ 2.5 cm (15/201) which were associated to more discomfort and pelvic pain (p value, 0.03; independent t test).

The POP-Q examination lasted longer than the digital examination but never exceeded 3.5 min for either examiner.

Table 5 Stages of POP-Q examination performed in left lateral and dorsal lithotomy position

POP-Q stages (<i>n</i> =50)	Left lateral	Lithotomy
0	1	1
1	10	10
2	27	28
3	11	10
4	1	1
Median stage	2	2

Values are expressed as number of patients. (*p* value, 0.6; Wilcoxon signed-rank test)

The findings of the POP-Q assessment are displayed in Table 2. The agreement between examination findings of the two blinded clinicians are displayed in Tables 3 and 4.

The digital examination had a poor inter-observer reliability with a kappa value of 0.54. The POP-Q showed a high degree of reliability (0.88), with 95% confidence interval, for the difference between examiners being 0.92 cm for point D.

The POP-Q stages in the two positions are shown in Table 5. There was disagreement between stages for one patient only. The difference was one stage lower assigned in dorsal lithotomy position. However, the difference was not statistically significant between the two examinations (*p* value, 0.3).

The Spearman's rank correlation analysis confirmed that there was a high degree of correlation between POP-Q findings in left lateral and in dorsal lithotomy position (Table 6, *p*<0.001, ρ >0.95).

Discussion

In 1996, the ICS, the society of gynecologic surgeons, and the American urogynecologic society introduced the first standardised, objective, site-specific system (POP-Q) for

Table 6 Correlation between POP-Q examination findings in left lateral and dorsal lithotomy position

POP-Q points	Spearman correlation coefficient (ρ)	<i>p</i> value
Aa	0.965	0.001
Ba	0.974	0.001
C	0.949	0.001
Ap	0.946	0.001
Bp	0.975	0.001
D	0.999	0.001

describing, quantitating, and staging pelvic organ support in women [4].

It has been shown that this system is easy to learn and teach, takes only 2 to 3 min to perform, and has good intra as well as inter-observer reproducibility. The reliability of the measurements has also been demonstrated by independent examiners [14]. Finally, this method of examination has been reported to improve the clinical and scientific communication regarding POP [7] due to the calibrated and precise nature of the measurements.

Different patient positions during POP-Q examination have been assessed for the POP-Q systems [6, 7, 14].

Some authors have proposed to evaluate POP with a woman either standing or sitting upright in a birth chair [6, 7]. This is based on the assumption that the maximum hip flexion in the upright position, straightening and enlarging the pelvic outlet, allows the pelvic organs to protrude to a greater extent than the lithotomy position even with maximum Valsalva. Barber et al. [6] compared the POP-Q measurements in both dorsal lithotomy and upright sitting position and found a greater stage of prolapse when women were upright.

Although the increase of the measurements was statistically significant in the upright position, it was not clinically important, varying between 0.20 and 0.60 cm. However, there was a moderate to good correlation between the POP-Q measurements made in each position [6].

Unfortunately, a birth chair is not universally available, not widely used in some countries such as UK, and takes up a considerable amount of space.

Therefore, some other authors have suggested examination in the standing position [7, 14–16]. Swift and Herring found no significant differences between points and stages when the POP-Q was performed in either supine or standing position [7], but due to limitations of examining in the standing position, the authors were unable to complete two of the nine POP-Q measurements, genital hiatus, and perineal body length.

Nevertheless, many clinicians continue to prefer a digital examination [17]. Although the position during a vaginal examination has not been standardised, supine and left lateral positions are the most common positions during a digital examination in the UK.

Our study is the first to evaluate the inter-observer reliability of the POP-Q in left lateral position.

We demonstrated that the POP-Q in the left lateral position is reliable and easy to perform. The women find it acceptable, and it is not time consuming. Finally, our study also showed that the digital clinical evaluation of the POP using a four-grade assessment system with women lying in left lateral position, as it is currently commonly used in UK, is unreliable and has only moderate agreement between two experienced examiners.

The fact that we did not repeat the vaginal examination in the same women after 2 weeks (intra-observer reliability) or measure the intensity of straining, evaluating the vesical or rectal pressure, represented the weaknesses of our study.

In conclusion, on the light of our data, we might conclude that POP-Q in left lateral position is a useful research and clinical tool for the comparison of published series and evaluation of corrective surgery for vaginal prolapse.

Conflicts of interest The authors declare the following for research, consultancy, and/or advisory work. Linda Cardozo: Astellas, Lilly, UCB, Pharma, Pfizer, Gynecare, Plethora, Cook, Organon, and Bioxell; Vik Khullar: Astellas, Lilly, Allergan, Pfizer, Gynecare, Cook, and Bioxell; G.A. Digesu: Astellas and Bioxell; Stavros Athanasiou: Astellas, Pfizer, and Gynecare; Simon Hill: Pfizer, Astellas, Ardana, and UCB Pharma.

References

- Swift SE, Pound T, Dias JK (2001) Case-control study of etiologic factors in the development of severe pelvic organ prolapse. *Int Urogynecol J Pelvic Floor Dysfunct* 12(3):187–192
- Swift SE (2000) The distribution of pelvic organ support in a population of female subjects seen for routine gynecologic health care. *Am J Obstet Gynecol* 183(2):277–285
- Kapoor DS, Sultan AH, Thakar R, Abulafi MA, Swift RI, Ness W (2008) Management of complex pelvic floor disorders in a multidisciplinary pelvic floor clinic. *Colorectal Dis* 10(2):118–123
- Bump RC, Mattiasson A, Bo K, Brubaker LP, DeLancey JO, Klarskov P et al (1996) The standardization of terminology of female pelvic organ prolapse and pelvic floor dysfunction. *Am J Obstet Gynecol* 175:10–17
- Silva WA, Kleeman S, Segal J, Pauls R, Woods SE, Karram MM (2004) Effects of a full bladder and patient positioning on pelvic organ prolapse assessment. *Obstet Gynecol* 104(1):37–41
- Barber MD, Lambers A, Visco AG, Bump RC (2000) Effect of patient position on clinical evaluation of pelvic organ prolapse. *Obstet Gynecol* 96:18–22
- Swift SE, Herring M (1998) Comparison of pelvic organ prolapse in the dorsal lithotomy compared with the standing position. *Obstet Gynecol* 91:961–964
- Swift S, Morris S, McKinnie V, Freeman R, Petri E, Scotti RJ, Dwyer P (2006) Validation of a simplified technique for using the POPQ pelvic organ prolapse classification system. *Int Urogynecol J Pelvic Floor Dysfunct* 17(6):615–620
- Poston GJ, Joseph AE, Riddle PR (1983) The accuracy of ultrasound in the measurement of changes in bladder volume. *Br J Urol* 55:361–363
- Digesu GA, Khullar V, Cardozo L, Robinson D (2008) Inter-observer reliability of digital vaginal examination using a four-grade scale in different patient positions. *Int Urogynecol J Pelvic Floor Dysfunct* 19(9):1303–1307
- Altman DG (1999) *Practical statistics for medical research*. Chapman & Hall/CRC, Boca Raton
- Nunnally JC (1978) *Psychometric Theory*. McGraw Hill, New York
- Carmines E, Zeller R (1979) *Reliability and validity assessment: quantitative applications in the social science*. Sage, Beverly Hills, California
- Hall AF, Theofrastous JP, Cundiff GW, Harris RL, Hamilton LF, Swift SE et al (1996) 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* 175:1467–1470
- Grody MHT (1997) Posterior compartment defects. In: Rock J, Thompson T (eds) *Telinde's operative gynecology*, 8th edn. Lippincott-Raven Press, Philadelphia, pp 1044–1056
- Walters MD, Karram MM (1992) Enterocoele. *American urogynecologic society quarterly report*. American Urogynecologic Society, Chicago, 10
- Auwad W, Freeman RM, Swift S (2004) Is the pelvic organ prolapse quantification system (POPQ) being used? A survey of members of the international continence society (ICS) and the American urogynecologic society (AUGS). *Int Urogynecol J Pelvic Floor Dysfunct* 15(5):324–327