

Article Type: Clinical Article

Subheading: Gynecology

**Comparison between the Valsalva maneuver and intraoperative traction measurements in pelvic organ prolapse assessment**

Stergios K. Doumouchtsis <sup>1,2</sup>, Nivedita Gauthaman <sup>3</sup>, Azar Khunda <sup>4</sup>, Maya Basu <sup>5</sup>, Kiran Dadhwal <sup>2</sup>, Yasmin V. Gayle <sup>2</sup>, Constantin M. Durnea <sup>1,\*</sup>

<sup>1</sup> Department of Obstetrics and Gynaecology, Epsom and St Helier University Hospitals, Epsom, UK

<sup>2</sup> St. George's University of London, London, UK

<sup>3</sup> Hampshire Hospitals NHS Foundation Trust, Basingstoke, UK

<sup>4</sup> St Thomas' Hospital, London, UK

<sup>5</sup> Medway NHS Foundation Trust, Gillingham, UK

\* Correspondence

Constantin M. Durnea, Department of Obstetrics and Gynaecology, Rowan House, Dorking Road, Epsom, KT18 7EG, UK.

Email: [costea.durnea@gmail.com](mailto:costea.durnea@gmail.com)

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ijgo.12319

This article is protected by copyright. All rights reserved.

**Keywords:** Difference; Measurement; POP-Q; Traction; Valsalva maneuver

**Synopsis:** For all POP-Q points, a higher degree of pelvic organ prolapse was diagnosed intraoperatively with mechanical traction than preoperatively with the Valsalva maneuver.

#### **Abstract**

**Objective:** To compare the assessment of pelvic organ prolapse (POP) between the Pelvic Organ Prolapse Quantification (POP-Q) system with Valsalva maneuver and intraoperative measurement with mechanical traction.

**Methods:** A prospective observational study included 100 women with POP attending a tertiary urogynecology clinic in the UK and undergoing vaginal prolapse surgical procedures between October 2011 and October 2014. The women were examined in the clinic using POP-Q with the Valsalva maneuver and in the operating theater under general anesthesia with mechanical traction. The two sets of measurements were compared.

**Results:** All POP-Q measurements obtained with traction demonstrated significantly higher descent as compared with those measured by Valsalva maneuver (mean differences: Aa 0.64 cm; Ap 1.32 cm; Ba 0.96 cm; Bp 1.34 cm; C 3.57 cm; D 3.40 cm; all  $P \leq 0.003$ ). The perineal body and total vaginal lengths did not differ significantly.

**Conclusion:** Measurements of six POP-Q points obtained with traction showed a higher grade of POP than those assessed with Valsalva maneuver. On this basis, surgeons might decide on the extent of surgical procedure after examination under

anesthesia; however, preoperative patient counselling would be essential to obtain consent for this approach. The clinical significance of the findings requires further evaluation.

## 1 INTRODUCTION

Pelvic organ prolapse (POP) affects up to 40% of postmenopausal women according to the results of the Women's Health Initiative study in the USA [1]. The lifetime risk of undergoing at least one operation for POP is 11% [2]. The presentation of POP is variable, depending on the affected compartment, severity of prolapse, presence of symptoms, and associated impact on quality of life. Clinical symptoms may or may not be specific to the affected compartment.

Management options range from conservative measures—e.g. pelvic floor exercises and use of pessaries—to surgical procedures with or without the use of synthetic meshes or biological grafts. However, the choice of treatment depends on several factors, including the patient's desires, surgeon's preferences and recommendations, site and stage of prolapse, compartments affected, presence of symptoms, and impact on quality of life. A clinical assessment documenting the exact location and stage of prolapse determines possible indication for further investigations and management options.

Clinical assessment systems have been developed and adjusted with time to achieve the most optimal clinical description of POP. The Baden–Walker Halfway system measures the extent of prolapse using six points on the anterior, apical, and posterior compartments of the vagina, with the hymen as the fixed anatomic reference point. POP is classified into four grades: zero indicates no prolapse and four indicates the maximal prolapse of the site

measured [3]. The Baden–Walker Halfway system provides a comprehensive assessment of POP, but lacks specific locations of fascial defects that might be responsible for the prolapse.

The Pelvic Organ Prolapse Quantification (POP-Q) system provides more site-specific information, while also grading POP into four stages and using fixed anatomic reference points. Additionally, POP-Q provides measurements of total vaginal length (TVL), genital hiatus (GH) and perineal body (PB) [4]. However, it has been observed that the stage of POP detected is greater in an upright position than in a lithotomy position for many women, even with maximum straining [5]. Furthermore, previous research has suggested that there is a significant difference between POP-Q measurements made in clinic with the Valsalva maneuver and those made under general anesthesia in theater with applied traction [6]. However, this study reported limited data in terms of different POP-Q points. The accuracy of preoperative assessment of POP could therefore be suboptimal, potentially leading to incorrect choice of a specific surgical intervention after clinical examination.

The aim of the present study was to compare POP-Q measurements obtained in the clinic using the Valsalva maneuver with assessments of POP made under anesthesia with mechanical traction.

## **2 MATERIALS AND METHODS**

The present prospective observational study included women with symptoms of POP attending a consultant urogynecology clinic in the UK between October 1, 2011, and October 31, 2014. Women who subsequently underwent vaginal prolapse surgical procedures were deemed eligible. Women who attended the clinic but did not undergo

Accepted Article

prolapse surgery, those with previous hysterectomy, and those undergoing abdominal prolapse surgery were excluded. Because the women were registered in the audit database in a prospective manner, the selection process was completed when 100 women meeting the study criteria were entered in the database. Consecutive women meeting the study criteria were included to avoid selection bias. Because the aim of the study was to evaluate multiple measurements of different vaginal and pelvic floor compartments and women with different stages and types of POP over a realistic time period, a sample of 100 women was judged suitable. The study was registered with the local audit department (Audit Registration 1511). Ethical approval and informed consent were not required because the study was based on data obtained during standard practice.

All women were initially assessed using the POP-Q system with a Valsalva maneuver in the left lateral position, as per the clinic's standard assessment practice. The women's bladders were empty to moderately filled. Measurements were performed using custom-made rulers and wooden smear spatulas.

In the operating theater, all surgery was performed under general anesthesia. The Aa, Ba, Ap, Bp, C, and D points were assessed using the same POP-Q system intraoperatively under general anesthesia in the lithotomy position, with mechanical traction applied by the surgeon before commencing the surgical procedure for the treatment of prolapse. TVL, GH, and PB measurements were undertaken without traction. Measurements were performed using sterile rulers. The surgeons in theater were not masked to the outpatient measurements because both assessments were part of standard practice.

All sets of variables were recorded and entered prospectively into both an Excel 2010 spreadsheet (Microsoft Corporation, Redmond, WA, USA) and the patient's clinical notes.

POP-Q points were documented to the nearest 0.5 cm. The maximum extent of prolapse was confirmed in the clinic during the third maximal Valsalva maneuver [7], as well as intraoperatively using an Allis clamp applied to the relevant POP-Q point and performing mechanical traction until no further descent occurred.

The primary outcome measure was the correlation between the two types of assessment. Outliers were identified and removed, and cases with missing data were excluded from the analysis.

Statistical analysis was performed using SPSS version 21 (IBM, Armonk, NY, USA). The variance between measurements obtained by the two techniques was assessed by calculating differences between individual POP-Q variables and comparing values by paired *t* tests. Bland–Altman plots were also used to compare the two methods of assessment. For each variable, a Spearman rank-order correlation test was calculated to measure the strength of association between the average POP-Q measurement and the difference between POP-Q measurements made by both techniques.  $P < 0.05$  was considered to be statistically significant.

### **3 RESULTS**

The demographic characteristics of the study women and the surgical procedures are presented in Table 1. The POP-Q measurements with Valsalva maneuver and traction, and the differences between the two sets of measurements are presented in Table 2. For all POP-Q landmarks apart from TVL and PB, the measurements made by the two approaches

were significantly different ( $P<0.05$ ). The largest differences between the two types of measurement were noted for cervical descent (point C), posterior vaginal fornix (point D), and posterior vaginal wall (points Ap and Bp).

Bland–Altman analysis demonstrated that there is little agreement between the two measurement techniques (Figures S1–S9). Similarly, very weak correlation was demonstrated on Spearman rank-order correlation for all POP-Q points (Table 3).

The traction and Valsalva maneuver measurements were also assessed by grade of POP. The differences were more pronounced for low-stage POP than for high-stage POP (Table 4).

#### **4 DISCUSSION**

The present study investigated the correlation between POP-Q scores measured in the clinic with Valsalva maneuver and those measured in the operating theater under general anesthesia with mechanical traction. Of the nine POP-Q measurements, all six intravaginal points (Aa, Ap, Ba, Bp, C, and D) were higher when assessed intraoperatively by mechanical traction under anesthesia as compared with the original values recorded in the clinic using the Valsalva maneuver. The greatest differences between measurements were observed for points C and D, both representing the apical compartment. This is consistent with data reported previously for point C [6]. The present findings are also in agreement with other studies [8,9]; but the prolapse observed on traction for all intravaginal POP-Q points was greater in the present study than in previous investigations.

Accepted Article

It is possible that the general anesthetic agents were responsible for the differences observed: it has been suggested that neuromuscular blockade and denervation of the pelvic floor are associated with the development of POP [10]. During anesthesia, the patient's levator ani muscles become relaxed, which might contribute to the increased degree of descent measured with traction under anesthesia. This has been demonstrated previously, when assessment in the theater showed significantly greater POP-Q examination measurements compared with the clinic ones [11]. In another study [9], a significant intraoperative increase in measurements was demonstrated only in the case of grade 1 apical prolapse; the authors attributed this to relaxation of the tissues under general anesthesia. By contrast, Crosby et al. [12] concluded that use of a neuromuscular blockade agent did not account for the difference between point C measurements demonstrated with the two techniques.

Collectively, the practical implication of these observations is the potential amendment of the surgical plan in view of the different examination findings obtained during anesthesia [13]. For example, a decision to proceed to anterior vaginal repair or to apical resuspension—or indeed to perform only a posterior vaginal repair—might be made on the basis of an examination made under anesthesia, as long as this has previously been discussed and agreed with the patient. Overall, the present findings indicate a need for further research on the cause of the differences in measurements and the impact of neuromuscular blockade on prolapse assessment.

On a simplistic basis, the differences might be due to the higher mechanical traction forces relative to the patient's downward force during the Valsalva maneuver. Almost all patients can achieve an optimal Valsalva maneuver on push [14]; however, sometimes they might not perform it correctly even after being taught how to do so. The patient might be too



inhibited to provide an effective Valsalva maneuver due to fear of urinary, fecal, or flatus incontinence, or might simply be embarrassed in relation to the nature of this examination.

Co-activation of the levator ani muscles might also affect the efficacy of a Valsalva maneuver [15,16].

The present results demonstrate a small decrease in GH length when measured intraoperatively under anesthesia as compared with preoperatively in the clinic with Valsalva maneuver. Although the difference was significant, it was below the precision limit of the measurement (0.5 cm) and might arise because GH is measured during Valsalva maneuver in the clinic but measured without traction or tension in the theater. This finding is consistent with previous studies [9].

The TVL and PB length measurements demonstrated the lowest differences, which were not significant, between the two examination methods. To our knowledge, no studies have reported that either of these measurements differs significantly. Notably, the paired *t* test did not demonstrate a difference in TVL or PB between the two measurement techniques in the present study, and the Bland–Altman analysis and Spearman rank-order test did not confirm high agreement or correlation; however, there was some agreement between the two techniques for TVL and PB measurements. This observation might be explained by the fact that both variables are measured without mechanical traction. Additionally, PB has lower distensibility on Valsalva maneuver.

Thus, the question is which of the two assessments correlates better with the real clinical problem? On one hand, Crosby et al. [12] demonstrated that patients with a lower grade of POP are more likely to have higher discrepancies between measurements of up to 5 cm, in

agreement with the present results. On the other hand, another study [17] suggested that, on the basis of examination at the 3-month follow-up, there is no benefit in using intraoperative assessment to amend the surgical plan. A long-term follow-up study might shed more light on this matter. Additionally, not all types of POP are symptomatic and the research question would ideally need to be validated against the presence of symptoms [18].

The present study has a few strengths. First, the measurements of prolapse were quantitative, ensuring that statistical analysis could be performed to see whether the differences were significant. Second, a validated assessment of prolapse was used (the POP-Q system).

The study also has some limitations. First, its sample size of subgroups with different stages of POP in different compartments was small for subanalyses. Second, interobserver variability might be a potential limitation. To minimize this, however, the researchers adhered to the principles of POP-Q measurements as described, and used a standardized method of assessing the prolapse under anesthesia by applying maximum traction on the specific POP-Q points using Allis forceps until no further descent was observed. Third, the degree of force used during intraoperative mechanical traction might have varied; however, the tissue has limited elasticity and the present measurements were obtained by a consistent technique. Fourth, demographic variables that might affect POP—e.g. body mass index, parity, and mode of previous deliveries—were not included. Parity and number of deliveries might affect the degree of descent due to ligament or fascial defects in a different way as compared with aging, chronic constipation, or other causes of POP. Because the two types of examinations were compared on the same patient, however, the effect of parity would theoretically be limited in the present study; additionally, such an evaluation would require a

much larger sample size. Last, the Valsalva maneuver itself might be inconsistent because it is patient-dependent, and its results might be variable.

In conclusion, points Aa, Ap, Ba, Bp, C, and D scored significantly higher when measured preoperatively under anesthesia with applied mechanical traction than when measured in clinic with Valsalva maneuver. These variations should be taken into account when planning surgical treatment, because the stage of prolapse might be higher than the one diagnosed in clinic. The clinical significance of these findings needs to be validated in studies based on long-term follow-up, with additional assessments of impact on quality of life. Future studies might also evaluate the association of the different measurements with the patient's clinical symptoms to provide more evidence on the clinical value of the different forms of POP assessment.

#### **Author contributions**

SKD contributed to project development, data collection, data analysis, and manuscript editing. NG, AK, and MB contributed to data collection, data management, and manuscript editing. KD contributed to data management, data analysis, and manuscript writing and editing. YVG contributed to data management, data analysis, and manuscript writing. CMD contributed to data analysis and manuscript writing.

#### **Conflicts of interest**

The authors have no conflicts of interest.

## References

1. Hendrix SL, Clark A, Nygaard I, Aragaki A, Barnabei V, McTiernan A. Pelvic organ prolapse in the Women's Health Initiative: gravity and gravidity. *Am J Obstet Gynecol*, 2002. 186(6): 1160-6.
2. Olsen AL, Smith VJ, Bergstrom JO, Colling JC, Clark AL. Epidemiology of surgically managed pelvic organ prolapse and urinary incontinence. *Obstet Gynecol*, 1997. 89(4): 501-6.
3. Kuncharapu, I., B.A. Majeroni, and D.W. Johnson, Pelvic organ prolapse. *Am Fam Physician*, 2010. 81(9): 1111-7.
4. Bump RC, Mattiasson A, Bø K, et al. The standardization of terminology of female pelvic organ prolapse and pelvic floor dysfunction. *Am J Obstet Gynecol*, 1996. 175(1): 10-7.
5. Barber MD, Lambers A, Visco AG, Bump RC. Effect of patient position on clinical evaluation of pelvic organ prolapse. *Obstet Gynecol*, 2000. 96(1): 18-22.
6. Chao FL, Rosamilia A, Dwyer PL, Polyakov A, Schierlitz L, Agnew G. Does pre-operative traction on the cervix approximate intra-operative uterine prolapse? A randomised controlled trial. *Int Urogynecol J*, 2012. 23(4): 417-22.
7. Tumbarello JA<sup>1</sup>, Hsu Y, Lewicky-Gaupp C, Rohrer S, DeLancey JO. Do repetitive Valsalva maneuvers change maximum prolapse on dynamic MRI? *Int Urogynecol J*, 2010. 21(10): 1247-51.
8. Vierhout, M.E., J. Stoutjesdijk, and J. Spruijt, A comparison of preoperative and intraoperative evaluation of patients undergoing pelvic reconstructive surgery for pelvic organ prolapse using the Pelvic Organ Prolapse Quantification System. *Int Urogynecol J Pelvic Floor Dysfunct*, 2006. 17(1): 46-9.

9. Krissi H, Eitan R, Ram E, Peled Y. How accurate is preoperative evaluation of pelvic organ prolapse in women undergoing vaginal reconstruction surgery? *PLoS One*, 2012. 7(10): e47027.
10. Allen RE, Hosker GL, Smith AR, Warrell DW. Pelvic floor damage and childbirth: a neurophysiological study. *Br J Obstet Gynaecol*, 1990. 97(9): 770-9.
11. Brazell HD, Claydon CS, Li J, Moore C, Dereska N, Hudson S, Swift S. Does neuromuscular blockade affect the assessment of pelvic organ prolapse? *Int Urogynecol J*, 2012. 23(11): 1599-603.
12. Crosby EC1, Sharp KM, Gasperut A, Delancey JO, Morgan DM. Apical descent in the office and the operating room: the effect of prolapse size. *Female Pelvic Med Reconstr Surg*, 2013. 19(5): 278-81.
13. Smith, P., S. Swift, and J.O. DeLancey, To pull or not to pull, that is the question...how should we define prolapse? *Int Urogynecol J*, 2013. 24(12): 1995-6.
14. Mulder, F.E., K.L. Shek, H.P. Dietz. What's a proper push? The Valsalva manoeuvre revisited. *Aust N Z J Obstet Gynaecol*, 2012. 52(3): 282-5.
15. Dietz, H.P., B.T. Haylen, J. Broome, Ultrasound in the quantification of female pelvic organ prolapse. *Ultrasound Obstet Gynecol*, 2001. 18(5): 511-4.
16. Shek, K.L, H.P. Dietz. Assessment of pelvic organ prolapse: a review. *Ultrasound Obstet Gynecol*, 2016. 48(6): 681-692.
17. Foon R, Agur W, Kingsly A, White P, Smith P. Traction on the cervix in theatre before anterior repair: Does it tell us when to perform a concomitant hysterectomy? *Eur J Obstet Gynecol Reprod Biol*, 2012. 160(2): 205-9.

18. Dietz, H.P. K.P. Mann, What is clinically relevant prolapse? An attempt at defining cutoffs for the clinical assessment of pelvic organ descent. *Int Urogynecol J*, 2014. 25(4): 451-5.

Supporting information legends

**Figure S1** Bland–Altman plot of point Aa to analyze the agreement between average Aa measurements by both techniques and the difference between measurements by both techniques.

**Figure S2** Bland–Altman plot of point Ap to analyze the agreement between average Ap measurements by both techniques and the difference between measurements by both techniques.

**Figure S3** Bland–Altman plot of point Ba to analyze the agreement between average Ba measurements by both techniques and the difference between measurements by both techniques.

**Figure S4** Bland–Altman plot of point Bp to analyze the agreement between average Bp measurements by both techniques and the difference between measurements by both techniques.

**Figure S5** Bland–Altman plot of point C to analyze the agreement between average C measurements by both techniques and the difference between measurements by both techniques.

**Figure S6** Bland–Altman plot of point D to analyze the agreement between average D measurements by both techniques and the difference between measurements by both techniques.

**Figure S7** Bland–Altman plot of total vaginal length to analyze the agreement between average measurements of total vaginal length by both techniques and the difference between measurements by both techniques.

**Figure S8** Bland–Altman plot of genital hiatus to analyze the agreement between average genital hiatus measurements by both techniques and the difference between measurements by both techniques.

**Figure S9** Bland–Altman plot of point PB to analyze the agreement between average PB measurements by both techniques and the difference between measurements by both techniques.

**Table 1** Demographic and surgical characteristics (n=100).

<b>Characteristic</b>	<b>Value<sup>a</sup></b>
Demographics	
Age, y	60.7 ± 12.9
Ethnicity	
White	60 (60)
Asian	8 (8)
Black	9 (9)
Other	4 (4)
Weight, kg	74.0 ± 13.6
Surgical procedures	
Anterior vaginal repair	72 (72)
Posterior vaginal repair	88 (88)
Vaginal hysterectomy	14 (14)
Sacrospinous ligament fixation	17 (17)
Perineorrhaphy	70 (70)

<sup>a</sup> Values are given as mean ± SD or number (percentage).



**Table 2** POP-Q point measurements and differences (n=100).<sup>a</sup>

Landmark	No. of patients	Measurement, cm		Difference (95% CI), cm	P value
		By Valsalva maneuver	By traction		
Aa	89	-0.44 ± 1.52	0.20 ± 1.38	0.64 ± 1.29 (0.37 to 0.91)	<0.001
Ap	90	-0.73 ± 1.30	0.59 ± 1.28	1.32 ± 1.40 (1.03 to 1.62)	<0.001
Ba	90	-0.31 ± 1.82	0.64 ± 1.80	0.96 ± 1.39 (0.67 to 1.25)	<0.001
Bp	90	-0.71 ± 1.30	0.63 ± 1.48	1.34 ± 1.53 (1.02 to 1.67)	<0.001
C	88	-4.38 ± 2.98	-0.8 ± 2.87	3.57 ± 2.66 (3.01 to 4.14)	<0.001
D	62	-6.13 ± 2.02	-2.73 ± 2.81	3.40 ± 2.89 (2.67 to 4.14)	<0.001
TVL	88	8.82 ± 1.18	8.93 ± 1.42 <sup>b</sup>	0.11 ± 1.63 (-0.23 to 0.46)	0.515
GH	88	4.72 ± 1.02	4.36 ± 0.89 <sup>b</sup>	-0.35 ± 1.06 (-0.58 to -0.13)	0.003
PB	89	3.06 ± 0.92	3.00 ± 0.80 <sup>b</sup>	-0.06 ± 0.99 (-0.27 to 0.15)	0.6

Abbreviations: POP-Q, Pelvic Organ Prolapse Quantification; CI, confidence interval; TVL, total vaginal length; GH, genital hiatus; PB, perineal body.

<sup>a</sup> Values are given as mean ± SD unless indicated otherwise.

<sup>b</sup> No traction applied in theatre for this measurement.

**Table 3** Agreement of measurements between the two types of POP-Q assessment (n=100).

Landmark	No. of patients	Mean difference (limits of agreement), cm	Spearman correlation, $r_s$
Aa	89	0.64 (-1.94 to 3.22)	-0.152
Ap	90	1.32 (-1.48 to 4.12)	-0.131
Ba	90	0.96 (-1.76 to 3.68)	0.029
Bp	90	1.34 (-1.66 to 4.34)	0.041
C	88	3.57 (-1.64 to 8.78)	0.007
D	62	3.40 (-2.26 to 9.06)	0.335
TVL	88	0.11 (-3.08 to 3.30)	0.157
GH	88	-0.35 (-2.43 to 1.73)	-0.128
PB	89	-0.06 (-2.43 to 1.73)	-0.223

Abbreviations: POP-Q, Pelvic Organ Prolapse Quantification; TVL, total vaginal length; GH, genital hiatus; PB, perineal body.

**Table 4** Difference between traction and Valsalva maneuver for prolapse grades 1–2 and 3–4.

Point	Prolapse grade 1–2				Prolapse grade 3–4			
	Valsalva maneuver	Traction	Difference	<i>P</i> value	Valsalva maneuver	Traction	Difference	<i>P</i> value
Aa	-1.7	-0.5	-1.2	0.001	0.8	0.9	0.1	0.394
Ap	-1.7	0.2	-1.9	<0.001	0.5	1.1	0.6	0.005
Ba	-1.8	-0.5	-1.3	<0.001	0.9	1.6	0.7	0.006
Bp	-1.7	0.2	-1.9	<0.001	0.5	1.2	0.7	0.008
C	-5.2	-1.5	-3.7	<0.001	1.9	2	0.1	0.448
D	-6.4	-2.9	-3.5	<0.001	0.5	2	1.5	0.343