250

Hayward L[1], Wong V[2], Tomlinson L[1], Smalldridge J[1]
1. *Middlemore Hospital , Dept Womens Health,* ***2.*** *Middlemore Hospital, Dept Womens Health*

A PROSPECTIVE INTEROBSERVER STUDY USING THE POPSTIX DEVICE, A MEASURING TOOL TO SIMPLIFY POPQ MEASUREMENT.

Hypothesis / aims of study

The ICS POPQ system is the recommended standard for reporting of pelvic organ prolapse. The POPQ system consists of 9 points that are measured to the nearest 0.5cm. When practising urogynaecologists were surveyed, it was found that many clinicians estimated POPQ points, used modified spatulas or swabs to perform POPQ measurements. We have developed a measuring tool, POPstix (*Fig 1)* to improve the speed and accuracy of POPQ scoring. This tool is evaluated in a prospective interobserver study using measured versus estimated POPQ scores.

Fig 1.

Study design, materials and methods

Women attending a urogynaecology clinic were examined by 2 clinicians, each blinded to the results of the other. Clinicians were randomised to either first guess POPQ points or measure POPQ points using the POPstix device. All POPQ points except total vaginal length (TVL) were recorded with women in a modified dorsal lithotomy position at maximum valsalva.

Outcomes measured included the reproducibility of both methods (estimated and measured) in assessing POPQ points. The agreement between each examiners estimated POPQ points and measured POPQ points (using 0.5 cm as an acceptable degree of accuracy) was calculated. POPQ points were converted to ordinal stages and the same analysis was repeated. Taking the measured POPQ stages as the most accurate, we calculated the percentage of time POPQ stages were altered (under or overestimated) by the estimation of POPQ points.

Results

38 patients consented to participate in this study. Interobserver correlation co-efficient was higher amongst measured values compared to estimated values. Results of agreement between the two observers' assessments were analysed using the Spearman ranking correlation coefficient, whilst mean difference and limit of agreement was analysed using Bland-Altman plot (*Tables 1&2*). Staging and agreement of staging were determined using Kappa statistics, results are summarised in *Table 3*.

POPQ point	**Spearman ranking correlation coefficient**		**Mean difference**		**Limits of agreement**	
	Estimated	**Measured**	**Estimated**	**Measured**	**Estimated**	**Measured**
AA	0.88	0.93	-0.5(0.8)	-0.5(0.7)	-2.0,1.5	-2.0,1.5
BA	0.96	0.96	0.5(1.0)	-0.5(1.0)	-2.0,2.5	-2.5,2.0
AP	0.85	0.9	0.5(0.8)	0.5(0.7)	-1.5,2.0	-1.5,2.0
BP	0.85	0.88	1.0(2.3)	1.0(2.0)	-4.0,5.5	-3.5,5.0
C	0.74	0.98	-1.5(4.1)	-0.5(1.1)	-9.5,7.5	-2.5,2.5
D	0.75	0.92	0.5(1.1)	-0.5(1.1)	-2.0,3.0	-2.0,3.0
TVL	0.41	0.83	1.0(1.7)	0.5(0.9)	-3.0,4.0	-2.0,2.0
GH	0.8	0.91	0.5(1.0)	0.5(0.8)	-2.0,2.5	-1.5,2.0
PB	0.63	0.76	0.5(0.7)	0.5(0.6)	-1.5,1.5	-1.5,2.0

Table 1.Interobserver estimated and measured POPQ points.

POPQ point	**Spearman ranking correlation coefficient**	**Mean difference**	**Limits of agreement**
AA	0.97	0.1(0.5)	-1.0,1.5
BA	0.99	0.5(0.5)	-1.0,1.0
AP	0.94	-0.5(0.5)	-1.5,1.0
BP	0.99	-0.5(0.5)	-1.0,1.0
C	0.92	-0.5(2.0)	-4.0,3.0
D	0.89	-0.5(1.5)	-3.0,2.0
TVL	0.89	-0.5(1.0)	-2.5,1.5
GH	0.92	0.5(0.7)	-1.5,2.0
PB	0.88	0.5(0.5)	-1.0,1.0

Table 2. Estimated POPQ points compared to measured POPQ points

Staging	*% of time POP-Q stages differed between Estimated compared to Measured*	*% of time overestimation*	*% of time underestimation*	*Kappa statistics for agreement of staging*
Anterior compartment	28.9	54.5	45.5	0.74
Posterior compartment	21.1.	50	50	0.76
Apical compartment	18.4	100	0	0.70

Table 3 Staging differences between estimated and measured POPQ

Interpretation of results

Agreement between all POPQ points was better when points were measured rather than estimated. Correlation between observers in the estimation of POP-Q points was worst in the apical compartment and also with TVL. The percentage of time a difference of more than 0.5cm measurement was obtained between estimated values compared to measured values were 35.5% in anterior compartment, 32.9% in posterior compartment whilst 47.4% in the apical compartment. A staging change occurred when POPQ points were measured in 28.9% in the anterior compartment, 21% in the posterior compartment and 18% in the apical compartment. This represents an important difference if reporting surgical outcomes or planning surgical intervention.

Concluding message

Estimation of POPQ points is inaccurate. Use of the POPstix measuring device improved accuracy of measurements. The surgical outcomes of papers utilising estimated POPQ points should be questioned.

Specify source of funding or grant	**no disclosures**
Is this a clinical trial?	**Yes**
Is this study registered in a public clinical trials registry?	**No**
What were the subjects in the study?	**HUMAN**
Was this study approved by an ethics committee?	**Yes**
Specify Name of Ethics Committee	**northern Y ethics committee, New Zealand**
Was the Declaration of Helsinki followed?	**Yes**
Was informed consent obtained from the patients?	**Yes**