

## FURTHER PROPERTIES OF THE PELVIC ORGAN PROLAPSE SYMPTOM SCORE: MINIMALLY IMPORTANT CHANGE AND TEST-RETEST RELIABILITY

### Hypothesis / aims of study

The pelvic organ prolapse symptom score (POP-SS) was developed to meet the need for a brief but robust instrument to measure key prolapse symptoms. This 7-item scale (each item scored 0 (never) to 4 (all of the time)) is summed to give a total score ranging from 0 to 28 where higher scores indicate more severe symptoms. To date to our knowledge the POP-SS has been used as the continuous primary outcome measure in three trials, and in one longitudinal study. Data supporting its internal consistency, construct validity and sensitivity to change have been published (1). The aim of this research is to report on two other important properties of this instrument: the minimally important change (MIC), and the test-retest reliability. Identifying the MIC of the POP-SS aids interpretation of the scores and facilitates its use in trials. Analysis of test-retest reliability establishes its stability, an important feature of an outcome measure.

### Study design, materials and methods

#### *Minimally Important Change*

51 out of a possible 66 women who participated in a factorial trial comparing prolapse surgery with and without mesh, and with two different suture materials, had POP-SS scores at baseline and two year follow-up and contributed to analyses. An anchor based approach, incorporating the Receiver Operator Curve (ROC) method (2), was applied to these data to estimate the MIC. The change in POP-SS score was anchored to women's opinion at two years as to whether they were satisfied with the results of their operation (i.e. improved) or not. The MIC was determined as the maximum value of Youden's Index (sensitivity+(specificity-1)) (3) calculated from the coordinates of the ROC.

#### *Test-retest reliability*

As part of a longitudinal study of pelvic floor dysfunction in women after childbirth, a questionnaire was distributed at 12 year follow-up which included the POP-SS. Women at one site were invited to complete the POP-SS on a second occasion a short time later in order to assess test-retest reliability of the instrument. Scores from the two occasions were compared using the percentage agreement and the mean difference and standard deviation.

### Results

#### *Minimally Important Change*

Women who were satisfied with their surgery had a significantly greater improvement in their POP-SS score compared to women who were not ( $t=-4.446$ ,  $df=49$ ,  $p<0.001$ ; mean diff in POP-SS change between groups  $-10.5$ , 95% CI  $-15.2$ ,  $-5.7$ ) (table). The ROC derived from the POP-SS change scores is shown in figure 1. The maximum value of Youden's Index indicated that a decrease of 1.5 was the minimum improvement in POP-SS that was important to these women (figure 2). 100% of women who reported they were satisfied had a decrease in POP-SS of 1.5 or more, whereas only 40% of women who were not satisfied fell into this category. Some uncertainty is attached to the MIC estimate due to the small sample size. For example, MIC cut-off values from  $-0.5$  to  $-5.5$  yielded similar values of Youden's Index, that is, it is possible that it is a larger decrease in POP-SS which indicates important change.

Table. Change in POP-SS\* by satisfaction with operation at 2 year follow up

Satisfaction at 2 years	N	Mean	SD
Change in POP-SS from satisfied with operation baseline to 2 years	41	-11.28	6.20
not satisfied with operation	10	-.80	8.51

\* Negative change values indicate decrease in POP-SS from baseline to 2 year follow up i.e. symptoms becoming less severe

Figure 1. Receiver operator curve for POP-SS change score for identification of women satisfied with their operation at 2 years

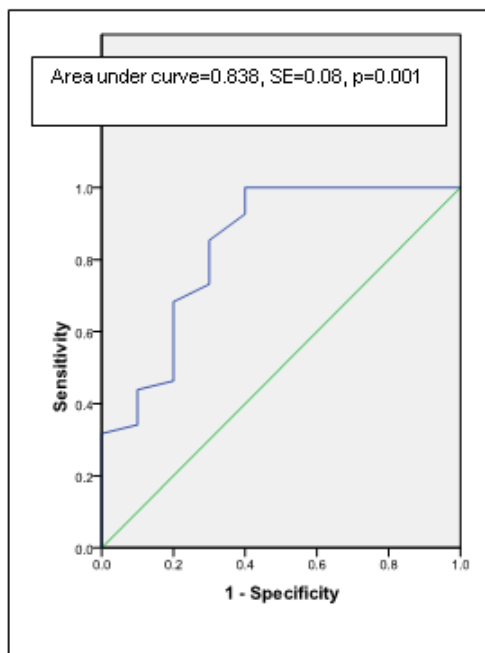
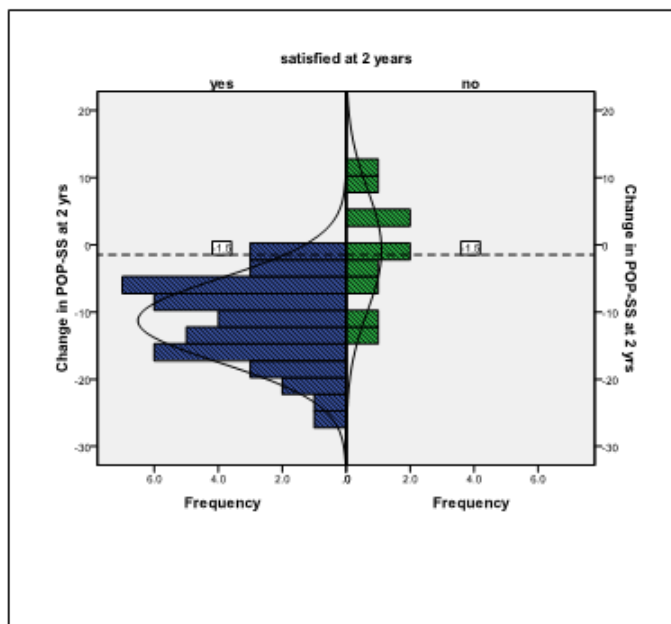


Figure 2. Changes in POP-SS at 2 years grouped by satisfaction with operation, showing normal distributions and MIC derived from ROC (-1.5)



#### Test-retest reliability

110 out of 192 women completed the POP-SS on both occasions, with a mean of 36 days between test and retest. The agreement between scores on the two occasions was moderately high (69% agreed to within 2 points) and the difference in score between occasions was small (mean difference 0.4, SD 2.8).

#### Interpretation of results

The method used of establishing MIC aids interpretation of the magnitude of change in the POP-SS which is meaningful to women, which makes the instrument more useful in measuring improvements due to treatment within clinical trials. In this sample of women who underwent surgical repair of their prolapse a decrease in the POP-SS of 1.5 or more corresponded best with satisfaction. This corresponds to a change in one of the domains (e.g. a feeling of something coming down) from “all of the time” at baseline to “sometimes” at follow-up. The POP-SS has been shown to give repeatable scores over a short period of no clinical change giving further confidence that changes detected in clinical trials will be due to intervention rather than an unstable instrument.

#### Concluding message

These findings provide preliminary information on how to interpret scores from the POP-SS and also evidence of the measure’s stability, allowing it to be more widely used in future prolapse research. Larger studies of the POP-SS MIC in varying clinical populations may be helpful to assess how the MIC might differ for women with varying levels of morbidity.

#### References

1. Hagen, S; Glazener, C; Sinclair, L; Stark, D; Bugge, C. Psychometric properties of the pelvic organ prolapse symptom score. BJOG : an international journal of obstetrics and gynaecology, 2009 Jan, 116(1):25-31.
2. de Vet, HCW; Terluin, B; Knol, DL; Roorda, LD; Mokkink, LB; Ostelo, RWJG; Hendriks, EJM; Bouter, LM; Terwee, CB. Three ways to quantify uncertainty in individually applied “minimally important change” values. Journal of clinical epidemiology, 2010 Jan, 63(1):37-45.
3. Fluss, R; Faraggi, D; Reiser, B. Estimation of the Youden Index and its Associated Cutoff Point. Biometrical Journal, 2005 Aug, 47(4):458-472.

#### Specify source of funding or grant

S Hagen's post funded by the Chief Scientist Office, Scottish Government Health Directorates. Data analysed are from the IMPRESS study, a self-funded study, and the Prolong study, funded by Wellbeing of Women, RCOG.

#### Is this a clinical trial?

No

#### What were the subjects in the study?

HUMAN

#### Was this study approved by an ethics committee?

Yes

#### Specify Name of Ethics Committee

IMPRESS study: Grampian Research Ethics Committee, project no. 04/MRE10/72, approved 9 May 2005.  
Prolong study: NHS MREC, project no. 06/MRE10/25, approved 18 April 06

#### Was the Declaration of Helsinki followed?

Yes

#### Was informed consent obtained from the patients?

Yes

