

**Author Posting. © The Authors 2009 This is the author's version of the work. It is posted here for personal use, not for redistribution. The definitive version was published in British Journal of Obstetrics & Gynaecology, 116(1);25-31. [http://dx.doi.org/DOI: 10.1111/j.1471-0528.2008.01903.x](http://dx.doi.org/DOI:10.1111/j.1471-0528.2008.01903.x)**

**TITLE Psychometric properties of the Pelvic Organ Prolapse Symptom Score (POP-SS)**

**RUNNING TITLE The pelvic organ prolapse symptom score (POP-SS)**

**AUTHORS**

Dr Hagen, Suzanne (**corresponding author**)

Urogenital Disorders Programme Leader

Nursing, Midwifery and Allied Health Professions Research Unit

Glasgow Caledonian University

Cowcaddens Road

Glasgow G4 0BA

Telephone: 0141 331 8104 E-mail: [s.hagen@gcal.ac.uk](mailto:s.hagen@gcal.ac.uk)

Dr Glazener, Cathryn

Gynaecologist & Reader in Health Services Research

Health Services Research Unit, Health Sciences Building

Foresterhill

Aberdeen AB25 2ZD

Ms Sinclair, Lesley (*details given are for post held at time of writing*)

Urogenital Disorders Research Fellow

Nursing, Midwifery and Allied Health Professions Research Unit

Glasgow Caledonian University

Cowcaddens Road

Glasgow G4 0BA

Mrs Stark, Diane

Superintendent Physiotherapist

Southern General Hospital

1345 Govan Road

Glasgow G51 4TF

Dr Bugge, Carol

Senior Lecturer

Department of Nursing and Midwifery

University of Stirling

Stirling FK9 4LA

## **ABSTRACT**

**Objective;** to assess the internal consistency, construct validity and sensitivity to change of a pelvic organ prolapse symptom score (POP-SS).

**Design;** analysis of data from three prolapse studies, including symptomatic and asymptomatic women, who completed the POP-SS.

**Setting;** 1) a community setting in New Zealand; 2) two gynaecology outpatient departments in Scotland; 3) a gynaecological surgery department in Scotland.

**Population or sample;** 1) participants from a survey of post-natal women at 12 year follow up, invited to complete a prolapse questionnaire and have prolapse assessment; 2) new gynaecology outpatients presenting with prolapse symptoms, randomised to pelvic floor muscle training (PFMT) or control; 3) women having anterior and/or posterior prolapse surgery, randomised to mesh insert or no mesh.

**Methods;** data were analysed to assess internal consistency, construct validity and sensitivity to change of the POP-SS.

**Main outcome measures;** Cronbach's alpha, significance of differences in POP-SS scores between studies, significance of difference in POP-SS scores pre- to post-intervention.

**Results;** For internal consistency, Cronbach's alpha ranged from 0.723 to 0.828. Women having surgery had higher POP-SS scores than those having conservative management (mean difference 5.0, 95% confidence interval 3.1 to 6.9), who in turn had higher scores than the asymptomatic women (mean difference 5.9, 95% CI 4.4 to 7.4). Significant differences in POP-SS score were detected after surgery and PFMT. The improvement due to surgery was significantly greater than that associated with PFMT ( $z = -3.006$ ,  $p = 0.003$ ).

**Conclusions;** The POP-SS has good internal consistency and construct validity, and is sensitive to change. (250 words)

**Keywords;** pelvic organ prolapse, outcome measure, psychometric properties, internal consistency, construct validity, sensitivity to change

## INTRODUCTION

Pelvic organ prolapse (POP), a common female condition, is symptomatic descent, from the normal anatomical position, of the vaginal walls, apex or vault <sup>1</sup>. Women with prolapse present with a variety of symptoms (vaginal, urinary, bowel, back, abdominal and sexual symptoms). Some of these symptoms are specifically associated with the descending pelvic organs protruding into the vaginal canal, for example, feeling of a bulge or something coming down. Others, such as urinary and bowel problems, can co-exist and may be related to or independent of the prolapse. It is important in research and clinical practice that we quantify such symptoms using standardised instruments with known psychometric properties.

Many instruments exist for measuring urinary symptoms and associated quality of life, including 17 questionnaires which the International Consultation on Incontinence (ICI) classed as Grade A (i.e. having established reliability, validity and responsiveness demonstrated in one or more datasets) <sup>1</sup>. Far fewer are available for the specific symptoms of prolapse. The ICI in 2005 <sup>1</sup> concluded that questionnaires in this area were “poorly developed to date and required encouragement”; two questionnaires of Grade B (validity and reliability established with rigour, or validity, reliability and responsiveness indicated) were identified (Pelvic Floor Disorder Inventory (PFDI) <sup>2</sup> and Pelvic Floor Impact Questionnaire (PFIQ) <sup>2</sup>), and an additional five which were in early development (Grade C) (e.g. P-QoL, ICIQ Vaginal Symptoms Questionnaire).

Since then work has been published on the above prolapse measures (short-form versions of the PFDI and PFIQ <sup>3</sup>; the P-QoL <sup>4</sup>; ICIQ Vaginal Symptoms Questionnaire <sup>5</sup>). The most prominent of these measures <sup>2,4</sup> are fairly lengthy, cover a range of symptoms and include a number of subscales, for example relating to urinary and bowel symptoms. It could be argued that these commonly co-existing symptoms are

better measured using validated, condition-specific instruments such as those developed by the International Consultation on Incontinence Questionnaire (ICIQ) group <sup>6</sup>, and that there remains a need for a brief symptom index which encapsulates the presence and extent of key prolapse symptoms. We report here on a scale which fulfils this need.

At the start of a programme of work on prolapse in 2000, when we sought a brief validated prolapse symptom scale, no suitable scale was available. We thus developed a simple set of key questions covering the symptoms caused or exacerbated specifically by prolapse which could serve as the primary outcome measure for subsequent randomised controlled trials of various interventions for POP. The key questions formed the basis for a POP symptom scale (POP-SS).

Our intention was to supplement the POP-SS with a number of existing validated scales aimed specifically at urinary (ICIQ-UI SF <sup>7</sup>), bowel (ICIQ-BS <sup>6</sup>) and sexual symptoms (PISQ-12 <sup>8</sup>), so that these functions could be assessed independently.

We administered the POP-SS to women in a number of research studies in order to generate data on its acceptability and performance. This article presents the findings regarding psychometric properties of the POP-SS, including internal consistency, construct validity and sensitivity to change.

## **METHODS**

### ***The Pelvic Organ Prolapse Symptom Score***

The POP-SS consists of seven items, each with a 5-point Likert response set (0=never, 1=occasionally, 2=sometimes, 3=most of the time, 4=all of the time) (Table 1). The question format and response set were modelled on those used by the ICIQ

group to standardise outcome measures in pelvic floor dysfunction research and clinical practice <sup>6</sup>. The items were developed from reviewing the literature in the course of undertaking a number of prolapse-related Cochrane systematic reviews<sup>9,10,11</sup>, and from discussion with gynaecologists, physiotherapists and women with prolapse. Some of the items are similar to those in other instruments since they target universally acknowledged symptoms associated with prolapse (e.g. a feeling of something coming down in the vagina). A total score (range 0 to 28) is calculated by summing the seven individual symptom responses to derive the POP-SS score. In addition, women indicate which one of the seven symptoms causes them most bother (Table 1).

At an early stage the POP-SS was assessed in qualitative interviews with 10 women (mean age 49 years) during which they completed the seven questions as part of a larger questionnaire. Women, who had either stage I (n=5) or II (n=5) prolapse, were purposively selected to represent the range of prolapse types (4 rectocele, 3 cystocele, 2 rectocele+cystocele, 1 uterine prolapse). The “think aloud” method <sup>12</sup> was used to encourage women to make explicit their understanding of the questions and rationale for responses chosen. Women were also asked to comment on the comprehensiveness and acceptability of the questionnaire. This approach provided evidence of content validity and acceptability, since women could understand the questions, and found them acceptable and relevant to the symptoms that troubled them in relation to their prolapse <sup>13</sup>.

The POP-SS has to date been used in three studies <sup>14,15,16</sup>, undertaken by the same research group, described below.

### ***Datasets***

### Study 1: PROlapse and incontinence: LONG-term research (ProLong) <sup>14</sup>

In New Zealand in 2005, 435 women were followed up 12 years after giving birth, at which time they had responded to a survey investigating postnatal urinary and faecal incontinence <sup>17</sup>. All women completed the POP-SS, and a sub-group of 166 women agreed to have objective prolapse assessment using the Pelvic Organ Prolapse – Quantification system (POP-Q) <sup>18</sup>. Women were not known to be symptomatic of prolapse: they were selected entirely on the basis of their involvement in the earlier survey.

### Study 2: Pelvic Organ Prolapse PhysiotherapY (POPPY) feasibility study <sup>15</sup>

In 2003/04, in a feasibility study at two Scottish centres, focussing on stage I or II prolapse, 47 women were randomised to either a pelvic floor muscle training (PFMT) intervention group or a control group receiving only a prolapse-related lifestyle advice leaflet. Objective quantification of prolapse type and severity was carried out at baseline and 6 months in both groups using the POP-Q <sup>18</sup>, and women completed a postal questionnaire including the POP-SS at baseline, 20 and 26 weeks.

### Study 3: Insertion of Mesh or sutures for PRolapsE Surgery Success (IMPRESS) <sup>16</sup>

In 2005 at one Scottish gynaecology centre, 66 women completed the POP-SS before and 6 months after having prolapse surgery (anterior and/or posterior repair). No POP-Q data were collected.

Analysis of the data resulting from these studies contributed information regarding internal consistency, construct validity and sensitivity to change of the POP-SS.

### ***Psychometric properties***

It is desirable for questions within a scale which are measuring the same concept, in this case extent of prolapse symptoms, to have high correlation; a property known as



“internal consistency”. Internal consistency of the POP-SS was assessed using data from Studies 1, 2 and 3.

A valid scale is one which measures what it intends to, and this is best assessed by comparison with a “gold standard” measure of the same quantity (criterion validity) <sup>19</sup>. When no gold standard measure exists, as is the case for prolapse symptoms, it is appropriate to assess construct validity instead. Hypotheses or constructs can be established regarding the responses to the scale, and if the hypotheses are supported by the data this provides evidence of construct validity. A form of construct validity known as *trait validity* was investigated via the hypothesis that scores at baseline (i.e. prior to any treatment) would be lowest in an asymptomatic group of women (Study 1), followed by a conservative management group (Study 2), and highest in a surgical intervention group (Study 3).

Ability to detect change in prolapse symptoms due to an intervention is an important scale property. Sensitivity to change of the POP-SS was assessed by testing for a significant pre- to post-intervention improvement in scores using data from Study 2 (PFMT intervention) and Study 3 (surgical intervention). The improvement in scores was expected to be greater in Study 3.

### ***Statistical analysis***

The three data sets described above were analysed separately and combined as appropriate to examine the properties of the POP-SS. The POP-SS scores were found to be non-normally distributed in several of the samples, particularly post-intervention when symptoms are likely to have resolved, thus primarily non-parametric methods were used.

Cronbach's alpha <sup>20</sup> was used to assess internal consistency of the seven item POP-SS using data from Studies 1, 2 and 3. Good internal consistency was assumed if Cronbach's alpha was between 0.7 and 0.9 <sup>21</sup>. It is undesirable for alpha to be too high as this suggests redundancy in the items of the scale.

In assessing trait validity, initially mean and median scores for the three study groups were tabulated. Non-parametric one-way analysis of variance (Kruskal-Wallis) was used to test for a significant difference between groups. Parametric analysis of variance, with post-hoc t-tests of differences between group means with Bonferroni correction, was also performed.

In terms of sensitivity to change, the Wilcoxon paired test was used to test for statistically significant differences between pre- and post-intervention POP-SS scores within studies. Differences between studies in pre- to post-intervention change in score were tested using the Mann-Whitney U test.

Analysis was undertaken using SPSS software and a 5% level of significance was used throughout.

## **RESULTS**

### *Sample characteristics*

The women in Study 3 (surgery group) were oldest and those in Study 1 (asymptomatic) were youngest, reflecting the differing study populations (Table 2).

### *Internal consistency*

The correlation amongst questions within the POP-SS was assessed in individual study datasets. Cronbach's alpha values (Table 3) indicate that the POP-SS seven

items have good internal consistency, i.e.  $\alpha > 0.7$ . The POPPY study (Study 2), which had the smallest sample size, had slightly lower Cronbach's alpha for both 20 and 26 week follow up time-points.

### *Construct validity*

The median POP-SS score at baseline was highest in the surgery study (Study 3), followed by the conservative intervention study (Study 2), and lowest in the study of asymptomatic women (Study 1) (Table 2). A significant difference between groups (Kruskal-Wallis  $\chi^2 = 176.730$ ,  $df = 2$ ,  $p < 0.001$ ) was detected. The ProLong mean score was significantly lower than that at baseline from POPPY (mean difference -5.9, 95% CI [-7.4, -4.4]) and IMPRESS (mean difference -10.9, 95% CI [-12.2, -9.6]), and the baseline POPPY mean score was significantly lower than that for IMPRESS (mean difference -5.0, 95% CI [-6.9, -3.1]). That is, the POP-SS scores differed between studies in a predictable way.

Table 1 highlights where differences in POP-SS scores between studies arose from. In the asymptomatic group of women (Study 1) a low percentage responded positively to having each of the seven symptoms. A feeling of incomplete bladder (38%) and bowel (46%) emptying were the symptoms most commonly reported, and the latter was the symptom which women said caused most bother. In contrast only 16% reported a feeling of something coming down. Percentages were consistently higher (in excess of 50% for each symptom) in the conservative treatment group (Study 2), with the most commonly reported symptom being a feeling of something coming down (79%) (Table 1). In the surgical group (Study 3) the percentages were highest of all studies, across all symptom questions. Most women in this study reported a feeling of something coming down (89%): this was both the most prevalent symptom (but not reported by everyone) and the one which most women identified as causing most bother.

### *Sensitivity to change*

In both the POPPY and IMPRESS studies a significant decrease in score after the interventions was detected (Table 4). The average decrease in score was shown to be significantly greater in the IMPRESS women than in the POPPY women ( $z = -3.006$ ,  $p = 0.003$ ), i.e. there was greater improvement in the surgery group than the PFMT group. Thus the POP-SS was able to detect the changes brought about by both types of intervention, and a difference in the magnitude of the change was distinguishable between studies.

## **DISCUSSION**

### *Summary of aims*

Our objective was to investigate the psychometric properties of a brief prolapse symptom scale (POP-SS) which might be used as an outcome measure in future trials of various prolapse interventions, and in clinical practice. No suitable validated scale of this nature was available at the onset of our programme of work. There are now a number of published prolapse instruments which are reported to be valid and reliable, however their length and complexity may make them impractical for some purposes. To our knowledge a reliable, valid and sensitive scale such as the POP-SS is still lacking in the literature.

### *Internal consistency*

Good internal consistency was confirmed across the three studies, and the POP-SS compared favourably with other instruments in this respect. Digesu et al <sup>4</sup> found Cronbach's alpha to be in excess of 0.80 in their assessment of the P-QoL, and Barber et al <sup>2</sup> reported Cronbach's alpha of 0.82 for the Pelvic Organ Prolapse Distress Inventory and 0.97 for the Pelvic Organ Prolapse Impact Questionnaire, which are the relevant subscales of the PFDI and PFIQ. It is reassuring that all POP-

SS items appear to be measuring the same trait, that is, there is homogeneity of the items within the scale. The value of Cronbach's alpha did not exceed 0.9 which would have suggested that the questions were too highly correlated and that some items were redundant. The findings suggest that a simple summation of scores over the seven symptom questions makes a reasonable index <sup>19</sup>.

It is interesting that the internal consistency of the POP-SS is good (Cronbach's alpha 0.823) in a sample of women selected without knowledge of their status with regards to prolapse (ProLong). This is encouraging if the POP-SS were to be used in trials of interventions to prevent prolapse.

#### *Validity*

The three study populations were representative of women with differing profiles of prolapse. Study 1 comprised a group of women who had participated in a post-natal survey 12 years previously, and for whom prolapse status was therefore unknown. Study 2 included women opting for conservative treatment, predominantly with stage I or II prolapse. Finally, Study 3 included women with prolapse of stage II or greater, having prolapse repair surgery. These groups of women would be expected to have different symptoms leading to their differing treatment choices, or in the case of Study 1, to no treatment for prolapse being sought. The ability of the POP-SS to differentiate between these groups, as indicated by the significant difference in scores, supports the trait validity of the scale. The predicted ordering in average group scores was observed in the data, providing additional evidence of validity. In a similar analysis the P-QoL domain scores were also found to differ significantly between symptomatic and asymptomatic women <sup>4</sup>. Other studies have investigated validity in terms of the relationship between symptom scores and prolapse severity, however, to date, it is not clear whether increasing symptoms are correlated with

increasing prolapse severity <sup>22</sup>. Analysis of the relationship between the POP-SS and the POP-Q is underway, and will contribute information to this debate.

#### *Sensitivity to change*

POP-SS could detect change due to both conservative and surgical interventions, and as expected the improvement in symptoms was greater in women who had surgery. This is an important property for a scale which is to be used in trials establishing the effectiveness of interventions for treatment of prolapse. The sensitivity to change of the P-QoL and PFDI/PFIQ has not been reported. The short forms of the PFDI and PFIQ were however found to have moderate to excellent responsiveness 3 to 6 months after surgery <sup>3</sup>.

#### *Implications for further research/use of POP-SS*

Our aim was to develop a scale which was brief and contained only the key symptoms important in obtaining a view of how prolapse is affecting a woman. It could be argued that the three questions within the POP-SS relating to bladder and bowel are not symptoms experienced exclusively by women with prolapse. Generally we avoided in our scale such questions, however these symptoms, more than others, are linked frequently with prolapse and were regarded to be worth including. The feeling of incomplete emptying of the bladder and bowel were the symptoms most commonly reported in the ProLong study (Study 1) in which women did not necessarily have prolapse. This perhaps reflects the fact that these symptoms are experienced generally by women other than those with prolapse. The prevalence of these symptoms was, however, far higher in the studies of women with confirmed prolapse.

The POP-SS was developed from a wide perspective, drawing on published research, clinical expertise and qualitative data from women with prolapse. It would be desirable to undertake further qualitative work investigating how women with different profiles respond to POP-SS items, and how well changes in scores reflect important modifications in their symptoms. Data are being gathered currently on the test-retest reliability of the POP-SS, and on its relationship with the observed POP-Q measure. Examination of the psychometric properties of the POP-SS in other treatment groups, for example women being fitted with a vaginal pessary, is also warranted.

In prolapse research the choice of an appropriate measure is still the subject of debate. There is a need to review and produce recommendations on the currently available prolapse questionnaires.

## **Conclusion**

It has been shown that the POP-SS is a measure with good internal consistency; it is valid as a measure of prolapse symptoms as scores differed predictably between groups of women known to differ in their prolapse symptoms; finally, it is sensitive to the change brought about by treatment for prolapse, specifically surgical repair and PFMT.

The POP-SS is a brief questionnaire which is acceptable to women and lends itself to both the research and the clinical environment.

## **Acknowledgements**

Collaborators in each of the studies: Dr Ian Ramsay, Dr Stewart Pringle, Dr Robert Hawthorn, Dr John Tierney, Dr Christine Bain, Dr Kevin Cooper, Ms Lynne Swan,

Professor Don Wilson, Professor Peter Herbison, Dr Nicola Dean, Ms Gaye Ellis, Dr Sabeena Allahdin.

The Chief Scientist Office, Scottish Government, funds Suzanne Hagen's post.

### **Disclosure of interests**

None.

### **Contribution to authorship**

SH was principal investigator on one study, carried out the main data analysis and drafted the manuscript. CG was principal investigator on one study, assisted with interpretation of the findings, commented on draft manuscripts and approved the final submission. LS was involved in the analysis of the POP-SS data, commented on draft manuscripts and approved the final submission. DS was co-investigator on one study, was involved in data collection, commented on draft manuscripts and approved the final submission. CB carried out the qualitative research on the POP-SS, drafted material for the manuscript and approved the final submission.

### **Details of ethical approval**

For each study the procedures received ethical approval from the relevant research ethics committee.

1. **ProLong**: Lower South Regional Ethics Committee, New Zealand; Ethics ref. LRS/05/04/009; approved 31 May 2005,
2. **POPPY**: a) Southern General Hospital Ethics Committee; Paper no. EC/02/S/115; approved 25 September 2002; b) Grampian Research Ethics Committee; Project no. 02/0243; approved 11 March 2003



3. **IMPRESS**: Grampian Research Ethics Committee; Project no. 04/MRE10/72;  
approved 9 May 2005

### **Funding**

1. **ProLong**: University of Otago Postgraduate Scholarship in Obstetrics and  
Gynaecology

2. **POPPY**: Health Services Research Committee grant, Chief Scientist Office,  
Scottish Government

3. **IMPRESS**: None

Table 1 Pelvic Organ Prolapse Symptom Scale: percent of women responding positively to symptom questions in each study

*How often during the last four weeks have you had the following symptoms (0=never, 1=occasionally, 2=sometimes, 3=most of the time, 4=all of the time):*

		Study 1 ProLong N = 435	Study 2 POPPY N = 47	Study 3 IMPRESS N = 66
A1	a feeling of something coming down from or in your vagina?	16.2%	78.7%	89.2%
A2	an uncomfortable feeling or pain in your vagina which is worse when standing?	13.0%	67.4%	70.8%
A3	a heaviness or dragging feeling in your lower abdomen / tummy?	27.0%	63.8%	81.5%
A4	a heaviness or dragging feeling in your lower back?	23.7%	59.6%	66.2%
A5	a need to strain (push) to empty your bladder?	24.1%	56.5%	72.3%
A6	a feeling that your bladder has not emptied completely?	38.1%	63.8%	87.7%
A7	a feeling that your bowel has not emptied completely?	46.4%	63.8%	76.9%
*A8	which of the symptoms above (questions A1 to A7) causes you most bother?	A7 39.3%	N/A	A1 40.0%

\* The symptom most often identified as causing most bother is shown, with the percentage of respondents which chose this symptom. This question was used only in Study 1 and Study 3.

Table 2 Characteristics of women from included studies

Variable	Study 1	Study 2	Study 3
Maximum sample size	435 (166 with POP-Q)	47	66
Median age in years (range)	40 (28, 57)	57 (31, 72)	61 (43, 84)
POP-Q at baseline n (%):			
Stage 0	3 (2)	#1 (2)	<i>all women were stage II, III or IV</i>
I	59 (35)	13 (29)	
II	101 (61)	30 (67)	
III	3 (2)	#1 (2)	
Leading edge POP type n (%):			
anterior	86 (52)	17 (70)	30 (48)
posterior	32 (19)	4 (16)	13 (20)
anterior=posterior	43 (26)	2 (8)	*19 (30)
superior	3 (1)	1 (4)	
Median POP-SS (range)			
Baseline	1 (0, 16)	8 (0, 21)	13 (3, 28)
20 wks/6 mnths post-intv	n/a	7.5 (2, 21)	3 (0, 22)
26 weeks post-intv	n/a	6 (1, 21)	n/a

\* women who had both anterior and posterior repair were assumed to have equal leading edges

# 1 woman presenting with prolapse symptoms but found to be stage 0 on examination was included. 1 woman with stage 3 prolapse was erroneously included.

Table 3 Internal consistency of the POP-SS

Study	Cronbach's alpha	N
Study 1	0.823	421
Study 2:		
Baseline	0.798	45
20 weeks	0.737	38
26 weeks	0.723	39
Study 3:		
Pre-op	0.819	65
6-mnths post-op	0.828	62

Table 4 Sensitivity to change of the POP-SS: paired tests

Study	mean pre-intervention	mean post-intervention	mean difference (pre – post)	n	Wilcoxon	p value
Study 2	9.05	6.11	3.47	*17	-2.308	0.021
Study 3	13.52	4.34	9.20	61	-6.069	<0.001

\* only data from intervention women are included: control women received only a lifestyle leaflet, and no significant change in POP-SS score was detected

## Reference List

- 1 Abrams P, Cardozo L, Khoury S, Wein A editors. Incontinence: 3rd International Consultation on Incontinence. Third ed. Paris: Health Publications Ltd.; 2005.
- 2 Barber MD, Kuchibhatla MN, Pieper CF, Bump RC. Psychometric evaluation of 2 comprehensive condition-specific quality of life instruments for women with pelvic floor disorders. *Am.J.Obstet.Gynecol.* 2001 Dec;185(6):1388-1395.
- 3 Barber MD, Walters MD, Bump RC. Short forms of two condition-specific quality-of-life questionnaires for women with pelvic floor disorders (PFDI-20 and PFIQ-7). *Am.J.Obstet.Gynecol.* 2005 Jul;193(1):103-113.
- 4 Digesu GA, Khullar V, Cardozo L, Robinson D, Salvatore S. P-QOL: a validated questionnaire to assess the symptoms and quality of life of women with urogenital prolapse. *Int.Urogynecol.J.Pelvic Floor Dysfunct.* 2005 May-Jun;16(3):176-81.
- 5 Price N, Jackson SR, Avery K, Brookes ST, Abrams P. Development and psychometric evaluation of the ICIQ Vaginal Symptoms Questionnaire: the ICIQ-VS. *BJOG* 2006 Jun;113(6):700-712.
- 6 Abrams P, Avery K, Gardener N, Donovan J. The International Consultation on Incontinence Modular Questionnaire: [www.iciq.net](http://www.iciq.net). *The Journal of Urology*, 2006 3;175(3):1063-1066.
- 7 Avery K, Donovan J, Peters TJ, Shaw C, Gotoh M, Abrams P. ICIQ: A brief and robust measure for evaluating the symptoms and impact of urinary incontinence. *Neurourol.Urodyn.* 2004;23(4):322-330.
- 8 Rogers RG, Coates KW, Kammerer-Doak D, Khalsa S, Qualls C. A short form of the Pelvic Organ Prolapse/Urinary Incontinence Sexual Questionnaire (PISQ-12). *Int.Urogynecol.J.Pelvic Floor Dysfunct.* 2003 Aug;14(3):164-8.
- 9 Adams E, Thomson A, Maher C, Hagen S. Mechanical devices for pelvic organ prolapse in women. *Cochrane Database of Systematic Reviews* 2004, Issue 2. Art. No.: CD004010. DOI: 10.1002/14651858.CD004010.pub2.
- 10 Hagen S, Stark D, Maher C, Adams E. Conservative management of pelvic organ prolapse in women. *Cochrane Database of Systematic Reviews* 2004, Issue 2. Art. No.: CD003882. DOI: 10.1002/14651858.CD003882.pub3.
- 11 Maher C, Baessler K, Glazener CMA, Adams EJ, Hagen S. Surgical management of pelvic organ prolapse in women. *Cochrane Database of Systematic Reviews* 2004, Issue 4. Art. No.: CD004014. DOI: 10.1002/14651858.CD004014.pub3.
- 12 Gilhooly K, Green C. Protocol analysis: theoretical background. In: Richardson JTE, editor. *Handbook of Qualitative Research Methods for Psychology and the Social Sciences*. First ed.: The British Psychological Society; 1996. p. 43-54.
- 13 Bugge C, Hagen S, Glazener C. The POPPY study: A qualitative evaluation of a pelvic organ prolapse outcome questionnaire. *ICS UK 12th Annual*

- Scientific Meeting; March 2005; Glasgow, UK.
- 14 Dean N, Martin L, Herbison P, Wilson D, Glazener C, MacArthur C, Hagen S, Grant A, Toozs-Hobson P, Bain C. Assessment of Pelvic Floor Dysfunction in Nulliparous Women: A Comparative Study. International Continence Society, 36th Annual Meeting; 27 Nov – 1 Dec 2006; Christchurch, New Zealand.
  - 15 Hagen S, Stark D, Sinclair L, Glazener C, Ramsay I. A feasibility study for a randomised controlled trial of a pelvic floor muscle training intervention for women with pelvic organ prolapse. International Continence Society, 36th Annual Meeting; 27 Nov – 1 Dec 2006; Christchurch, New Zealand.
  - 16 Allahdin S, Glazener CMA, Bain C. Feasibility study for a randomised controlled trial evaluating the use of absorbable mesh, polydioxanone and polyglactin sutures for anterior and posterior vaginal wall prolapse repairs. International Continence Society, 36th Annual Meeting; 27 Nov – 1 Dec 2006; Christchurch, New Zealand.
  - 17 Glazener C, Herbison G, MacArthur C, Lancashire R, McGee M, Grant A, et al. New postnatal urinary incontinence: obstetric and other risk factors in primiparae. *BJOG: An International Journal of Obstetrics and Gynaecology* 2006 02/32;113(2):208-217.
  - 18 Bump RC, Mattiasson A, Bo 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 Jul;175(1):10-17.
  - 19 Streiner DL, Norman GR. Health measurement scales, a practical guide to their development and use. Second ed. Oxford: Oxford University Press; 1995.
  - 20 Cronbach LJ. Coefficient Alpha and the internal structure of a tests. *Psychometrika* 1951;16:297-297-334.
  - 21 Nunnally JC. *Psychometric Theory*. Second ed. New York: McGraw-Hill; 1978.
  - 22 Tan JS, Lukacz ES, Meneffee SA, Powell CR, Nager CW. Predictive value of prolapse symptoms: a large database study. *International Urogynecology Journal* 2005 06/01;16(3):203-209.

## **Table/Figure Caption List**

Table 1 Pelvic Organ Prolapse Symptom Scale: percent of women responding positively to symptom questions in each study

Table 2 Characteristics of women from included studies

Table 3 Internal consistency of the POP-SS

Table 4 Sensitivity to change of the POP-SS: paired tests