

Psychometric properties of the pelvic organ prolapse symptom score

Hagen, Suzanne; Glazener, Cathryn M.A.; Sinclair, L.; Stark, D.; Bugge, C.

Published in:
BJOG: An International Journal of Obstetrics and Gynaecology

DOI:
[10.1111/j.1471-0528.2008.01903.x](https://doi.org/10.1111/j.1471-0528.2008.01903.x)

Publication date:
2009

Document Version
Author accepted manuscript

[Link to publication in ResearchOnline](#)

Citation for published version (Harvard):
Hagen, S, Glazener, CMA, Sinclair, L, Stark, D & Bugge, C 2009, 'Psychometric properties of the pelvic organ prolapse symptom score', *BJOG: An International Journal of Obstetrics and Gynaecology*, vol. 116, no. 1, pp. 25-31. <https://doi.org/10.1111/j.1471-0528.2008.01903.x>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please view our takedown policy at <https://edshare.gcu.ac.uk/id/eprint/5179> for details of how to contact us.

1 **TITLE PAGE**

2

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

5

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

7

8 **AUTHORS**

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

10 Urogenital Disorders Programme Leader

11 Nursing, Midwifery and Allied Health Professions Research Unit

12 Glasgow Caledonian University

13 Cowcaddens Road

14 Glasgow G4 0BA

15 Telephone: 0141 331 8104 E-mail: s.hagen@gcal.ac.uk

16

17 Dr Glazener, Cathryn

18 Gynaecologist & Reader in Health Services Research

19 Health Services Research Unit, Health Sciences Building

20 Foresterhill

21 Aberdeen AB25 2ZD

22

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

24 Urogenital Disorders Research Fellow

25 Nursing, Midwifery and Allied Health Professions Research Unit

26 Glasgow Caledonian University

27 Cowcaddens Road

28 Glasgow G4 0BA

29

30 Mrs Stark, Diane

31 Superintendent Physiotherapist

32 Southern General Hospital

33 1345 Govan Road

34 Glasgow G51 4TF

35

36 Dr Bugge, Carol

37 Senior Lecturer

38 Department of Nursing and Midwifery

39 University of Stirling

40 Stirling FK9 4LA

41

42

43 **ABSTRACT**

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

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

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

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

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

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

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

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

69

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

72

73

74

75 **INTRODUCTION**

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

85

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

97

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

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

108

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

115

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

119

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

124

125 **METHODS**

126

127 ***The Pelvic Organ Prolapse Symptom Score***

128 The POP-SS consists of seven items, each with a 5-point Likert response set
129 (0=never, 1=occasionally, 2=sometimes, 3=most of the time, 4=all of the time) (Table
130 1). The question format and response set were modelled on those used by the ICIQ

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

141

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

153

154 The POP-SS has to date been used in three studies ^{14,15,16}, undertaken by the same
155 research group, described below.

156

157

158 **Datasets**

159 Study 1: PROlapse and incontinence: LONG-term research (ProLong) ¹⁴

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

167

168 Study 2: Pelvic Organ Prolapse PhysiotherapY (POPPY) feasibility study ¹⁵

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

175

176 Study 3: Insertion of Mesh or sutures for PRolapsE Surgery Success (IMPRESS) ¹⁶

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

180

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

183

184 ***Psychometric properties***

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

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

189

190 A valid scale is one which measures what it intends to, and this is best assessed by
191 comparison with a “gold standard” measure of the same quantity (criterion validity) ¹⁹.

192 When no gold standard measure exists, as is the case for prolapse symptoms, it is
193 appropriate to assess construct validity instead. Hypotheses or constructs can be
194 established regarding the responses to the scale, and if the hypotheses are
195 supported by the data this provides evidence of construct validity. A form of
196 construct validity known as *trait validity* was investigated via the hypothesis that
197 scores at baseline (i.e. prior to any treatment) would be lowest in an asymptomatic
198 group of women (Study 1), followed by a conservative management group (Study 2),
199 and highest in a surgical intervention group (Study 3).

200

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

206

207 ***Statistical analysis***

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

213

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

218

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

224

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

229

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

232

233 **RESULTS**

234

235 *Sample characteristics*

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

238

239 *Internal consistency*

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

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

245

246 *Construct validity*

247 The median POP-SS score at baseline was highest in the surgery study (Study 3),
248 followed by the conservative intervention study (Study 2), and lowest in the study of
249 asymptomatic women (Study 1) (Table 2). A significant difference between groups
250 (Kruskal-Wallis $X^2 = 176.730$, $df = 2$, $p < 0.001$) was detected. The ProLong mean
251 score was significantly lower than that at baseline from POPPY (mean difference -
252 5.9, 95% CI [-7.4, -4.4]) and IMPRESS (mean difference -10.9, 95% CI [-12.2, -9.6]),
253 and the baseline POPPY mean score was significantly lower than that for IMPRESS
254 (mean difference -5.0, 95% CI [-6.9, -3.1]). That is, the POP-SS scores differed
255 between studies in a predictable way.

256

257 Table 1 highlights where differences in POP-SS scores between studies arose from.
258 In the asymptomatic group of women (Study 1) a low percentage responded
259 positively to having each of the seven symptoms. A feeling of incomplete bladder
260 (38%) and bowel (46%) emptying were the symptoms most commonly reported, and
261 the latter was the symptom which women said caused most bother. In contrast only
262 16% reported a feeling of something coming down. Percentages were consistently
263 higher (in excess of 50% for each symptom) in the conservative treatment group
264 (Study 2), with the most commonly reported symptom being a feeling of something
265 coming down (79%) (Table 1). In the surgical group (Study 3) the percentages were
266 highest of all studies, across all symptom questions. Most women in this study
267 reported a feeling of something coming down (89%): this was both the most
268 prevalent symptom (but not reported by everyone) and the one which most women
269 identified as causing most bother.

270

271 *Sensitivity to change*

272 In both the POPPY and IMPRESS studies a significant decrease in score after the
273 interventions was detected (Table 4). The average decrease in score was shown to
274 be significantly greater in the IMPRESS women than in the POPPY women ($z = -$
275 3.006 , $p = 0.003$), i.e. there was greater improvement in the surgery group than the
276 PFMT group. Thus the POP-SS was able to detect the changes brought about by
277 both types of intervention, and a difference in the magnitude of the change was
278 distinguishable between studies.

279

280 **DISCUSSION**

281 *Summary of aims*

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

290

291 *Internal consistency*

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

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

303

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

308

309 *Validity*

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

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

327

328

329 *Sensitivity to change*

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

337

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

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

351

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

361

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

365

366 **Conclusion**

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

372

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

375

376 **Acknowledgements**

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

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

381

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

383

384 **Disclosure of interests**

385 None.

386

387 **Contribution to authorship**

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

396

397 **Details of ethical approval**

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

400

401 1. **ProLong**: Lower South Regional Ethics Committee, New Zealand; Ethics ref.

402 LRS/05/04/009; approved 31 May 2005,

403 2. **POPPY**: a) Southern General Hospital Ethics Committee; Paper no. EC/02/S/115;

404 approved 25 September 2002; b) Grampian Research Ethics Committee; Project no.

405 02/0243; approved 11 March 2003

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

408

409 **Funding**

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

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

414 3. **IMPRESS**: None

415

416

417 **Table/ Figure Caption List**

418

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

421

422 Table 2 Characteristics of women from included studies

423

424 Table 3 Internal consistency of the POP-SS

425

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

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

429

430 *How often during the last four weeks have you had the following symptoms (0=never,*
 431 *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%

432

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

436 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

437 * women who had both anterior and posterior repair were assumed to have equal
438 leading edges
439 # 1 woman presenting with prolapse symptoms but found to be stage 0 on
440 examination was included. 1 woman with stage 3 prolapse was erroneously
441 included.

442 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

443

444 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

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

447

448

449

450 **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.icig.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, Menefee SA, Powell CR, Nager CW. Predictive value of prolapse symptoms: a large database study. *International Urogynecology Journal* 2005 06/01;16(3):203-209.

451

452

453

454

455

456