International Urogynecology Journal (2019) 30:115–122
https://doi.org/10.1007/s00192-018-3724-1

ORIGINAL ARTICLE

The minimal important difference of the Australian Pelvic Floor Questionnaire

Kaven Baessler[1] · Alexandra Mowat[2] · Christopher F. Maher[3]

Received: 15 April 2018 / Accepted: 13 July 2018 / Published online: 7 August 2018
© The International Urogynecological Association 2018

Abstract

Introduction and hypothesis The aim of this study was to establish the minimal important difference (MID) of the Australian Pelvic Floor Questionnaire (APFQ) in women undergoing surgery for stress urinary incontinence or symptomatic pelvic organ prolapse. A further aim was to estimate dysfunction scores dependent on the bothersomeness in a community cohort.

Methods The APFQ was completed before and 6 weeks after pelvic floor surgery by 183 women ($n = 80$ suburethral tape insertion; $n = 103$ laparoscopic sacrocolpopexy). Distribution and anchor-based methods were used to establish the effect size, standardised response mean and MID (calculated as the difference between women who stated no change or a little better in the Patient Global Impression of Improvement [PGI-I]).

In a community cohort of 470 women aged 42–80 years, the APFQ was analysed according to disclosed bothersomeness.

Results For the suburethral tape group, the effect size in the bladder domain was 1.5 and the PGI-I-based MID 1.3. For the POP surgery, group the effect size in the prolapse domain was calculated at 2.2 and the PGI-I-based MID at 1.0. The domain scores for women who declared no bother were significantly different from those who were a little bothered (bladder domain 2.2 vs 4.0, bowel 0.6 vs 1.7, POP 0.1 vs 3.2, sex 1.8 vs 3.0) with wide variations.

Conclusions The MID of the APFQ ranged from 1.0 to 1.3 in the domains after POP or continence surgery respectively. This is corroborated by the differences in domain scores from community-based women who were bothered versus not bothered by pelvic floor symptoms.

Keywords Minimal important difference · Pelvic floor questionnaire · Pelvic floor dysfunction · Pelvic floor surgery

Introduction

The Australian Pelvic Floor Questionnaire (APFQ) has four domains to assess bladder, bowel, prolapse and sexual symptoms, their severity, impact on quality of life and bothersomeness in women with pelvic floor dysfunction (PFD) [1]. It has been designed to be used in research and to facilitate routine urogynaecological assessment. Most items are scored from 0 to 3 employing different descriptions such as never, occasionally, frequently and daily to assess severity/frequency, and not at all, slightly, moderately, and greatly to estimate bothersomeness. The APFQ has been demonstrated to be valid, reliable (absolutely and relatively) and the self-administered version is also responsive to change [1, 2]. There is also a validated version for community-dwelling women available [3]. According to the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) [4], the APFQ is a health-related measurement tool with published details on content, construct and criterion validity and reliability.

Patient-related outcomes (PROs) have gained more importance because objective and functional results after treatment of PFDs do not necessarily correlate. The minimal important difference (MID) of a questionnaire indicates the change in

✉ Kaven Baessler
kaven.baessler@franziskus-berlin.de

[1] Pelvic Floor Centre Franziskus and St Joseph Hospitals, Budapester Strasse 15–17, 12787 Berlin, Germany

[2] Greenslopes Private Hospital and University of Queensland, Brisbane, QLD, Australia

[3] Royal Brisbane & Women's, Wesley and Mater Hospitals Brisbane, Suite 86, Sandford Jackson Building, 30 Chasely Street, Auchenflower, QLD 4066, Australia

symptoms that makes a meaningful difference to the patient [5, 6]. It can be determined using distribution-based methods including effect sizes. Effect sizes are considered equivalent to the minimal important difference [7] and have been published for the APFQ when responsiveness of the self-administered version was tested. However, the MID for the APFQ has not been established using anchor-based methods. The so-called anchor is an external patient-related outcome measure and therefore the MID is determined by the patients. The Patient Global Impression of Improvement (PGI-I) [8, 9] is frequently used in trials assessing pelvic floor surgery and has also served as an external anchor in studies determining the MID of the Pelvic Floor Distress Inventory (PFDI) and International Consultation on Incontinence Questionnaire-Urinary Incontinence Short Form [10, 11]. The PGI-I is a single-item, patient-related outcome questionnaire with a seven-point scale that assesses whether a symptom or condition has improved (very much better; much better; a little better), not changed or worsened (a little worse; much worse; very much worse) after treatment. It has been validated for use in women with urinary incontinence and pelvic organ prolapse (POP) [8, 9].

The APFQ [2] and its validated German versions [12, 13] have been used to assess the effectiveness of conservative and surgical treatment [14–17], and to estimate the prevalence and incidence of PFDs [18, 19]. Apart from the MID, "normal values" of bladder, bowel and sexual function have not yet been estimated. However, these may be helpful in interpreting treatment effects.

The aim of this study was to establish the MID of the Australian Pelvic Floor Questionnaire in women undergoing surgery for stress urinary incontinence or pelvic organ prolapse using anchor-based methods. A further aim was to estimate dysfunction scores dependent on the bothersomeness in a community-based cohort of women.

Materials and methods

We recruited 183 consecutive women who opted for pelvic floor surgery from urogynaecological clinics. The APFQ was completed before and 6 weeks after suburethral tape (SUT group, $n = 80$) operation for stress urinary incontinence (SUI) or POP surgery (POP group, $n = 103$). Retropubic and transobturator tension-free vaginal tapes were applied. POP surgery included laparoscopic sacrocolpopexy with or without Burch colposuspension, paravaginal defect repair or posterior repair, which were performed as required. The local institutional Ethics Committee (Uniting Care Health Human Research Ethics Committee 2017.27.241) approved this study.

Distribution-based methods were used to establish the effect size (mean change of score/standard deviation of baseline score) [20] and standardised response mean (mean change of score/standard deviation) [21]. Effect sizes and standardised response means of 0.2–0.49 were regarded as small, 0.5–0.8 as medium and more than 0.8 as large [20]. For the MID calculation based on an external anchor, the PGI-I was completed by all women after surgery. An at least moderate correlation (i.e. $r \geq 0.3$) of the PGI-I and APFQ is considered a prerequisite for the usefulness of the PGI-I as a chosen external anchor. Further analyses were performed after establishing this correlation. The minimal important difference was calculated as the difference in score changes between women who stated no change and a little difference in the PGI-I.

The sample size estimation was based on previously determined APFQ scores [2]. To demonstrate a score change of 1 from a mean bladder domain score of 3.1 (SD 1.7), with a power of 80% and alpha = 0.05, a total of 46 women were required. Recruitment was continued beyond this number until 10% of women ($n = 5$) indicated "no change" or "worse" after SUT insertion. Similarly after POP surgery, to demonstrate a score change of 1 from a mean prolapse domain score of 2.7 (SD 1.9), with a power of 80% and alpha 0.05, a total of 57 women were required. Recruitment was continued until 10% ($n = 6$) of women indicated "no change" or "worse" postoperatively. Receiver operating characteristics were performed to establish that the APFQ scores were able to distinguish women who reported improvement of symptoms from those who did not using the PGI-I.

To find a threshold score of the APFQ of bothersome pelvic floor symptoms, in a community cohort of 470 women aged 42–80 years the APFQ was analysed according to disclosed bothersomeness (Likert scale: not at all bothered—slightly—moderately—greatly bothered) in each APFQ domain. This cohort was part of a 5-year longitudinal study of ageing in women (LAW study, $n = 511$, aged 40–80 years), which was an extensive research programme of the Betty Byrne Henderson Centre at the Royal Brisbane & Women's Hospital. We excluded women with dementia, neurological diseases, bladder or bowel diversion and pelvic surgery in the study period, leaving 470 women for analysis.

For statistical analyses the "Statistical Package for the Social Sciences" (SPSS version 22) program was used. Distribution- and anchor-based methods included paired t tests to determine standardised response means (SRMs) and effect sizes (ESs). Receiver operating characteristic curves were generated after dividing women into those who reported improvement and those who considered themselves unchanged or worse (dichotomisation of the PGI-I). ANOVA was used to analyse differences in scores depending on bothersomeness in the community cohort. Demographic data were compared using non-parametric tests. We performed linear regression analyses to evaluate associations between the PGI-I and bladder and prolapse domain scores, change in score, bothersomeness and possibly interacting variables such as age, parity and body mass index (BMI).

Table 1 Demographic data in the different groups

	SUT group $N = 80$	POP surgery group $N = 103$	Community cohort $N = 470$	p*
Age (years)	59.0 (33–87)	67.0 (46–86)	60 (42–80)	<0.001
BMI	26.7 (19.1–45.8)	26.5 (17.3–36.7)	26.4 (16.2–51)	0.629
Parity	2 (0–7)	2 (0–6)	3 (0–10)	0.003

Values are median and range

BMI body mass index, *SUT* suburethral tape, *POP* pelvic organ prolapse

*Kruskal–Wallis test

Results

Demographic characteristics for the different study groups are summarised in Table 1.

Scores in all domains improved significantly after continence (Table 2) or POP surgery (Table 3). Tables 2 and 3 also display mean domain score changes, ESs and SRMs in both treatment groups. The external anchor (PGI-I) correlated significantly with the postoperative bladder domain score ($r = -0.73$), the mean score changes (pre- to postoperatively) in the bladder domain ($r = 0.53$) and global APFQ score ($r = 0.38$; all $p < 0.001$) in the SUT group. In a linear regression model we assessed the possible relation between PGI-I and bladder domain score, change in bladder score, bothersomeness, age, parity and BMI. The bladder domain score ($r = -0.73$), the change in the bladder score ($r = 0.59$) and the bother score ($r = -0.69$) remained in the model and significantly explained 62% of the PGI-I variances ($R^2 = 0.621$; $p < 0.001$).

For the POP surgery group, the correlation coefficients between PGI-I and the postoperative prolapse domain score, the mean change and the global APFQ score were 0.31, 0.41 and 0.38 respectively ($p < 0.001$). Regression analysis demonstrated that the change in POP score ($r = 0.41$), POP bother score ($r = -0.31$) and BMI ($r = -0.30$), postoperative bladder domain bother ($r = 0.28$) and bowel domain bother ($r = -0.22$) explained 39% of the PGI-I variances.

In the SUT group, the effect size in the bladder domain was 1.5 and expectedly lower in the other domains (Table 2). The mean change in the global APFQ score was 4.3 (SD 3.4), the effect size was 1.2. The PGI-based MID was 1.3: in 7 women with "no change" the mean difference in bladder score was 0.2 (SD 0.7) in contrast to the mean change of 1.5 (SD 0.9) in 16 women who specified themselves as being "a little better" (Table 4). Figure 1a displays the ROC curve that distinguished well (area under the curve 0.96) women who reported improvement versus no change or deterioration of symptoms.

In the POP group, the effect size in the prolapse domain was 2.2, but also lower in the other domains (Table 3). The effect size for the global APFQ score was 1.7 and the mean change 7.1 (SD 3.6). The PGI-based MID was calculated at 1.0 (Table 4): 6 women with "no change" had a mean prolapse score change of 2.6 (SD 3.4), whereas 19 women who were "a little better" had a mean change of 3.6 (SD 1.8). Of the women with no change or deterioration in the global PGI-I estimation, none had an increase in their prolapse score after POP surgery; all of them showed an improvement. These women documented some worsening bladder function scores (−0.04, Table 4). The ROC curve for distinguishing women who reported improvement versus no change in or deterioration of symptoms is shown in Fig. 1b.

In the community-based cohort, in women who declared no bother, all APFQ domain scores were significantly different from those who were a little bothered (bladder domain 2.2 vs 4.0, bowel 0.6 vs 1.7, POP 0.1 vs 3.2, sex 1.8 vs 3.0. Table 5 displays all APFQ domain scores dependent on bothersomeness. Although all APFQ domain scores were significantly higher in the treatment groups (SUT and POP group) compared with the community-based cohort (data not presented), the bladder domain scores were dependent on bothersomeness in the SUT group (slightly 2.5—moderately

Table 2 Minimum important difference (MID) according to distribution-based methods in the SUT group ($n = 80$)

Domains	Score Preoperatively	Score Postoperatively	Mean change	ES	SRM	Paired t test
Bladder	4.1 ± 1.4	1.9 ± 1.3	2.1 ± 1.5	1.5	1.4	<0.001
Bowel	1.9 ± 1.2	1.5 ± 0.9	0.4 ± 0.9	0.3	0.4	<0.001
POP	1.1 ± 1.9	0.5 ± 1.0	0.7 ± 1.6	0.4	0.4	<0.001
Sex	1.7 ± 1.4 ($N = 54$)	1.1 ± 0.8 ($N = 37$)	0.7 ± 1.2	0.5	0.6	0.002
Global PFD	9.3 ± 3.7	5.0 ± 2.0	4.3 ± 3.4	1.2	1.3	<0.001

Values represent mean ± standard deviation

PFD pelvic floor dysfunction, *ES* effect size, *SRM* standardised response mean

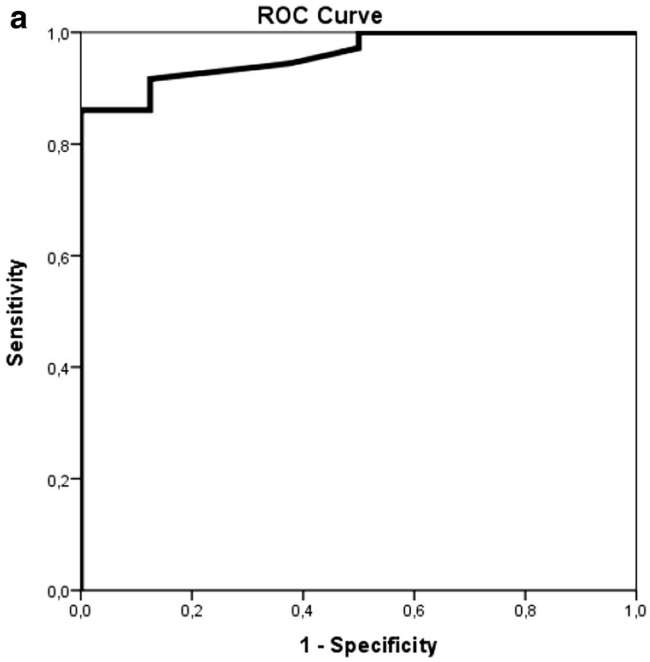
Table 3 MID according to distribution-based methods in the POP surgery group ($N = 103$)

Domains	Score Preoperatively	Score Postoperatively	Mean change	ES	SRM	Paired *t* test
Bladder	3.1 ± 1.6	2.0 ± 1.5	1.0 ± 1.4	0.6	0.7	<0.001
Bowel	2.7 ± 1.6	2.1 ± 1.2	0.5 ± 1.3	0.3	0.4	<0.001
POP	5.5 ± 2.2	0.7 ± 1.1	4.8 ± 2.3	2.2	2.1	<0.001
Sex	1.8 ± 1.3 ($N = 72$)	1.1 ± 0.8 ($N = 53$)	0.8 ± 1.3	0.6	0.6	<0.001
Global PFD	12.5 ± 4.1	5.4 ± 2.7	7.1 ± 3.6	1.7	2.0	<0.001

Values represent mean ± standard deviation

Table 4 Displayed are the mean differences in domain scores (pre- to postoperative score) ± SD according to the Patient Global Impression of Improvement (PGI-I) in the SUT and POP surgery groups. The MID is calculated as the difference in score changes between women with PGI-I = 0 and 1 in the bladder domain for the SUT group and prolapse domain for the POP group (marked in bold)

Domain scores		SUT group			POP surgery group		
	PGI	*n*	Score change	ANOVA	*n*	Score change	ANOVA
Bladder	−3	0		<0.001	0		0.023
	−2	1	−2		1	−2.0	
	−1	0			1	1.1	
	0	**7**	**0.2 ± 0.7**		6	−0.04 ± 1.6	
	1	**16**	**1.5 ± 0.9**		19	1.0 ± 1.3	
	2	19	2.3 ± 1.0		28	0.7 ± 1.3	
	3	37	2.8 ± 1.5		48	1.4 ± 1.4	
Bowel	−3	0		0.193	0		0.219
	−2	1	−0.3		1	0.9	
	−1	0			1	−1.5	
	0	7	0.4 ± 0.5		6	0.05 ± 0.7	
	1	16	0.5 ± 0.5		19	0.62 ± 1.2	
	2	19	0.7 ± 0.9		28	0.2 ± 1.1	
	3	37	0.2 ± 0.9		48	0.8 ± 1.4	
POP	−3	0		0.158	0		0.001
	−2	1	−0.7		1	0	
	−1	0			1	4	
	0	7	0.7 ± 0.9		**6**	**2.6 ± 3.4**	
	1	16	−0.1 ± 1.2		**19**	**3.6 ± 1.8**	
	2	19	0.7 ± 1.4		28	5.1 ± 2.1	
	3	37	1.0 ± 1.8		48	5.5 ± 2.1	
Sex	−3	0		0.694	0		0.089
	−2	1	0		1	1.4	
	−1	0			0		
	0	0			3	2.7 ± 2.0	
	1	9	0.8 ± 1.1		13	0.9 ± 1.1	
	2	9	0.4 ± 1.2		11	0.4 ± 1.1	
	3	15	0.9 ± 1.4		25	0.7 ± 1.3	
Global PFD	−3	0		0.007	0		0.007
	−2	1	−2.9		1	0.3	
	−1	0			1	3.6	
	0	7	1.9 ± 2.3		6	4.1 ± 5.1	
	1	16	2.9 ± 2.9		19	6.1 ± 2.9	
	2	19	4.4 ± 2.6		28	6.8 ± 2.9	
	3	37	4.6 ± 2.8		48	8.3 ± 3.6	



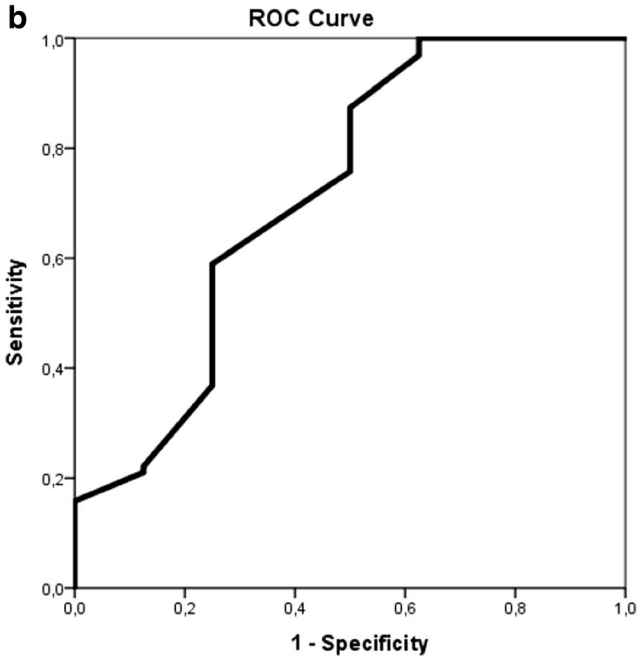


Fig. 1 Receiver operating characteristic (ROC) curve: the changes in the Australian Pelvic Floor Questionnaire (APFQ) **a** bladder and **b** prolapse domain scores distinguish well between women who improved and those who did not (dichotomised Patient Global Impression of Improvement [PGI-I]) after a suburethral tape or pelvic organ prolapse operation. Area under the curve bladder score (**a**) 0.96; 95% confidence interval 0.91–1.00; $p < 0.001$; pelvic organ prolapse score (**b**) 0.72; 95% confidence interval 0.50–0.94; $p = 0.04$

3.8—greatly bothered 4.6) and were similar to the bother scores in the community-based cohort. Similarly, for the POP surgery group, POP domain scores for those who were slightly bothered (2.3), moderately bothered (4.9) and greatly bothered (6.7) were similar to the scores in the community-based cohort. Figure 2 displays the box plots for all groups for the bladder, bowel and prolapse APFQ domains.

Discussion

The self-administered APFQ is a comprehensive instrument that was specifically designed for use in clinical practice and in research. The detection of clinically important changes over time or after treatment is important [22]. For women who have undergone continence surgery, a change of 1.3 in the bladder

Table 5 Community cohort of 470 women aged 42–80 years: mean domain scores depending on self-declared bothersomeness in each domain

Domain	Bothersomeness	*N*	Mean scores	95%CI	Range	ANOVA *p*
Bladder	0	362	0.7	0.6–0.7	0–5	<0.001
	1	84	2.2	2.0–2.5	0.6–6.4	
	2	16	4.0	3.3–4.7	1.9–6.1	
	3	8	5.2	3.6–6.8	2.5–7.8	
Bowel	0	360	0.6	0.6–0.7	0–2.9	<0.001
	1	66	1.7	1.5–1.9	0.6–4.7	
	2	24	3.0	2.5–3.4	1.5–5.9	
	3	60	3.7	3.2–4.2	2.4–5.6	
Prolapse	0	438	0.1	0.05–0.1	0–3.3	<0.001
	1	23	3.2	2.6–3.9	2.2–7.8	
	2	5	5.8	3.1–8.4	3.3–8.9	
	3	4	8.6	5.3–12.0	5.6–10	
Sex	0	98	1.8	1.6–2.0	0–6.7	<0.001
	1	54	3.0	2.7–3.3	1.3–6.0	
	2	44	4.8	4.4–5.1	2.7–6.7	
	3	7	6.5	5.5–7.4	5.3–8.0	

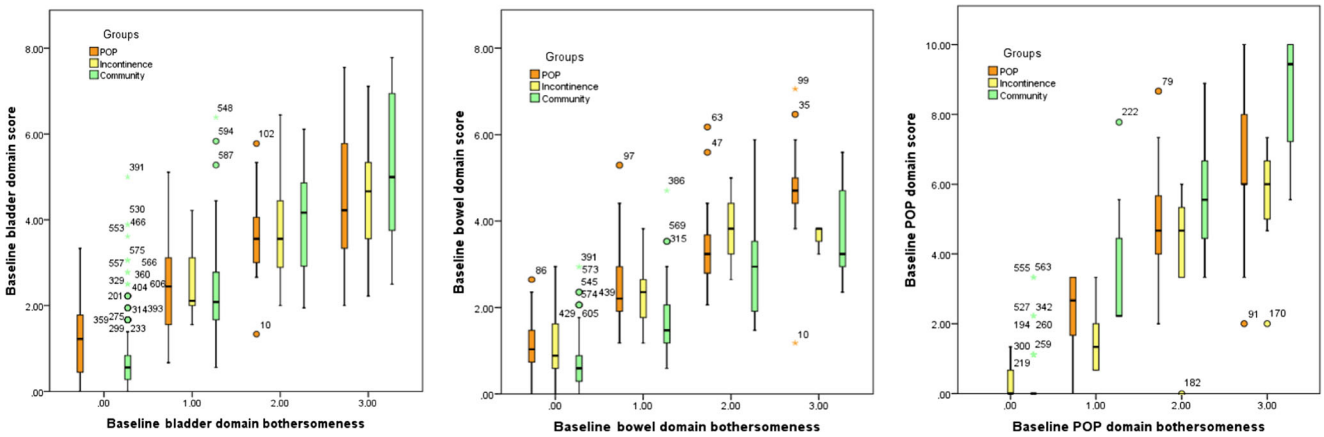


Fig. 2 Bladder, bowel and POP domain scores are dependent on perceived bothersomeness in all groups (POP, suburethral tape [SUT] and community-based cohorts). In the SUT group (*yellow bars*) all women reported bothersome stress urinary incontinence before the operation. This was similar in the POP surgery group, which included only women with symptomatic POP

domain implied a meaningful improvement. This represents a 24% change from the baseline bladder domain score, which is similar to the 23% change reported for the PFDI [23]. For women after POP surgery, a difference of 1 in the prolapse domain was calculated as an important improvement, indicating a change of 18% from the baseline prolapse score. The calculated MIDs of the APFQ ranged from 1.0 to 1.3 in the domains and from 1.2 to 2.2 for the global APFQ score depending on the primary PFD and subsequent continence or POP surgery. Therefore, changes of approximately 1 in the domain scores appear to adequately describe important differences. This is corroborated by the observed similar differences in domain scores from women who were bothered versus not bothered by pelvic floor symptoms in the community cohort. The PGI-I served well as an external anchor, which was demonstrated with adequate domain score correlations above 0.3 in both study groups. It is, however, a global score, taking more than just bladder or prolapse symptoms into account.

When interpreting the above MIDs it should be kept in mind that they may be specific to this cohort and also to the interventions. For questionnaires, the MID is a critical measure for interpreting changes after conservative or surgical treatment. It has been proposed that it makes a difference whether a woman undergoes conservative treatment such as pelvic floor muscle training for urinary incontinence or an operation [23]. Lower changes may be accepted and considered important after conservative treatment. The calculated ESs after continence and POP surgery in this study were similar to those after specific pelvic floor muscle exercises calculated previously for the German version of the APFQ [12]. In that study, distribution-based methods, including effect sizes and standardised response means were applied and an external anchor was not used. The estimation of the effect size grounds on baseline differences whereas the standardised response mean is based on differences in change. Although the effect size is considered the most adequate measurement [24], the baseline bladder domain scores differed considerably in women undergoing PFMT in Germany (2.6 ± 1.1) or continence surgery in Australia (4.1 ± 1.4). The standardised response mean, however, was also calculated at 1.3 and 1.4 respectively, supporting our observation that similar changes in scores were important for women independent of the choice of treatment. Still, whether the German and Australian cohorts are comparable may be disputed. SRM and effect size evaluations are suggested methods for detecting clinically meaningful changes [24, 25]. Conversely, COSMIN does not agree and argues that the effect of interventions are measured rather than the quality measures of the questionnaire [4].

We found some deviations in score changes after POP surgery, which explain the relatively low change in the prolapse domain dependent on the PGI-I. None of the women who described themselves as "worse" in the PGI-I after POP surgery scored higher in the prolapse domain postoperatively. Instead, they had higher (worse) scores in the bladder and bowel domains. This reflects the multidimensional effect of prolapse on bladder, bowel and sexual function: although POP surgery can effectively treat the prolapse, bladder or bowel function may deteriorate. This obviously has an impact on the overall APFQ dysfunction score and on the PGI-I as a global impression of symptoms and their impact on quality of life. In contrast, continence surgery is unlikely to severely affect bowel or prolapse symptoms and we did not find similar deviations in the continence surgery group. This issue has been recognised previously for the PFDI for example [26], and may be inherent to the complexity of PFDs.

Limitations of the study include that patients completed the questionnaire 6 weeks postoperatively. This may be considered too short a time to adequately assess the effect of surgery. In particular, the evaluation of sexual function is difficult, as many women had not resumed sexual activity at this stage. However, this appointment was routine and convenient and

the results show significant improvement in all pelvic floor domains. Whether the length of follow-up influences the perceived improvement remains unclear, although quality of life research in breast cancer patients, for example, suggests this [27]. A response shift due to adaptation and coping strategies [28] in addition to re-calibration [25] is likely to occur, but we would judge a 6-week follow-up too brief to allow any of these mechanisms to noticeably influence the patient-related outcomes used in our study. There may also be a re-call bias when applying the PGI-I after treatment, which is a limitation. However, again, a 6-week period appears to be too short to have a serious effect.

In view of an expected response shift in PRO assessment, the application of validated symptom-specific questionnaires and external anchors may be of more value than a global assessment.

The assessment of PFD in community-dwelling women showed a wide range of scores in each domain. There was also a large overlap between women who stated that they were not bothered versus slightly, moderately or greatly bothered. The minimum scores increased steadily, whereas the maximum scores were nearly independent of the degree of bother. This suggests that there are no "normal values" and bothersomeness appears to be the most important factor when considering treatment. All APFQ mean domain scores differed significantly between community and treatment cohorts, as there were more women who were bothered by incontinence or prolapse. This indicates that a higher score and bothersomeness has to be reached to seek treatment, which has recently been demonstrated in a survey employing the APFQ [29]. The domain scores were subject to reported bothersomeness in all of our groups. The determinants of perceived bothersomeness, however, remain unclear, and also why only some women subsequently seek care.

The degree of bother was also an important variable in the regression analyses to explore factors that influence the impression of improvement versus no change or deterioration of symptoms. The postoperative bladder domain score, the change of the bladder score from baseline and the bother score were independent variables. In the POP surgery group, the model was more complex. Apart from the prolapse domain score change, the modifiable risk factor BMI in addition to bladder and bowel domain bothersomeness remained independent factors.

Conclusion

The estimation of MIDs for the Australian Pelvic Floor Questionnaire can help to interpret symptom changes over time and after treatment. Changes of approximately 1 in the appropriate pelvic floor domain can be considered clinically important differences.

Compliance with ethical standards

Conflicts of interest None.

References

1. Baessler K, O'Neill SM, Maher CF, Battistutta D. Australian pelvic floor questionnaire: a validated interviewer-administered pelvic floor questionnaire for routine clinic and research. Int Urogynecol J Pelvic Floor Dysfunct. 2009;20(2):149–58.
2. Baessler K, O'Neill SM, Maher CF, Battistutta D. A validated self-administered female pelvic floor questionnaire. Int Urogynecol J. 2010;21(2):163–72. https://doi.org/10.1007/s00192-009-0997-4.
3. Baessler K, O'Neill SM, Maher CF, Battistutta D. An interviewer-administered validated female pelvic floor questionnaire for community-based research. Menopause. 2008;15(5):973–7.
4. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010;19(4):539–49. https://doi.org/10.1007/s11136-010-9606-8.
5. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989;10(4):407–15.
6. de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. Health Qual Life Outcomes. 2006;4:54. https://doi.org/10.1186/1477-7525-4-54.
7. Norman GR, Sridhar FG, Guyatt GH, Walter SD. Relation of distribution- and anchor-based approaches in interpretation of changes in health-related quality of life. Med Care. 2001;39(10):1039–47.
8. Srikrishna S, Robinson D, Cardozo L. Validation of the patient global impression of improvement (PGI-I) for urogenital prolapse. Int Urogynecol J. 2010;21(5):523–8. https://doi.org/10.1007/s00192-009-1069-5.
9. Yalcin I, Bump RC. Validation of two global impression questionnaires for incontinence. Am J Obstet Gynecol. 2003;189(1):98–101.
10. Barber MD, Spino C, Janz NK, Brubaker L, Nygaard I, Nager CW, et al. The minimum important differences for the urinary scales of the pelvic floor distress inventory and pelvic floor impact questionnaire. Am J Obstet Gynecol. 2009;200(5):580.e581–7. https://doi.org/10.1016/j.ajog.2009.02.007.
11. Sirls LT, Tennstedt S, Brubaker L, Kim HY, Nygaard I, Rahn DD, et al. The minimum important difference for the International Consultation on Incontinence Questionnaire-Urinary Incontinence Short Form in women with stress urinary incontinence. Neurourol Urodyn. 2015;34(2):183–7. https://doi.org/10.1002/nau.22533.
12. Baessler K, Junginger B. Validation of a pelvic floor questionnaire with improvement and satisfaction scales to assess symptom severity, bothersomeness and quality of life before and after pelvic floor therapy. Aktuelle Urol. 2011;42(5):316–22. https://doi.org/10.1055/s-0031-1271544.
13. Baessler K, Kempkensteffen C. Validation of a comprehensive pelvic floor questionnaire for the hospital, private practice and research. Gynakol Geburtshilfliche Rundsch. 2009;49(4):299–307.
14. Junginger B, Seibt E, Baessler K. Bladder-neck effective, integrative pelvic floor rehabilitation program: follow-up investigation. Eur J Obstet Gynecol Reprod Biol. 2014;174:150–3. https://doi.org/10.1016/j.ejogrb.2013.12.022.

15. Lin KY, Frawley HC, Granger CL, Denehy L. The Australian Pelvic Floor Questionnaire is a valid measure of pelvic floor symptoms in patients following surgery for colorectal cancer. Neurourol Urodyn. 2017;36(5):1395–402. https://doi.org/10.1002/nau.23122.
16. Farthmann J, Watermann D, Niesel A, Funfgeld C, Kraus A, Lenz F, et al. Lower exposure rates of partially absorbable mesh compared to nonabsorbable mesh for cystocele treatment: 3-year follow-up of a prospective randomized trial. Int Urogynecol J. 2013;24(5):749–58. https://doi.org/10.1007/s00192-012-1929-2.
17. Ong TA, Khong SY, Ng KL, Ting JR, Kamal N, Yeoh WS, et al. Using the Vibrance Kegel device with pelvic floor muscle exercise for stress urinary incontinence: a randomized controlled pilot study. Urology. 2015;86(3):487–91. https://doi.org/10.1016/j.urology.2015.06.022.
18. Durnea CM, Khashan AS, Kenny LC, Tabirca SS, O'Reilly BA. The role of prepregnancy pelvic floor dysfunction in postnatal pelvic morbidity in primiparous women. Int Urogynecol J. 2014;25(10):1363–74. https://doi.org/10.1007/s00192-014-2381-2.
19. Schoenfeld M, Fuermetz A, Muenster M, Ennemoser S, von Bodungen V, Friese K, et al. Sexuality in German urogynecological patients and healthy controls: is there a difference with respect to the diagnosis? Eur J Obstet Gynecol Reprod Biol. 2013;170(2):567–70. https://doi.org/10.1016/j.ejogrb.2013.08.002.
20. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care. 1989;27(3 Suppl):S178–89.
21. Katz JN, Harris TM, Larson MG, Krushell RJ, Brown CH, Fossel AH, et al. Predictors of functional outcomes after arthroscopic partial meniscectomy. J Rheumatol. 1992;19(12):1938–42.
22. Kirshner B, Guyatt G. A methodological framework for assessing health indices. J Chronic Dis. 1985;38(1):27–36.
23. Wiegersma M, Panman CM, Berger MY, De Vet HC, Kollen BJ, Dekker JH. Minimal important change in the pelvic floor distress inventory-20 among women opting for conservative prolapse treatment. Am J Obstet Gynecol. 2017;216(4):397.e391–7. https://doi.org/10.1016/j.ajog.2016.10.010.
24. Norman GR, Wyrwich KW, Patrick DL. The mathematical relationship among different forms of responsiveness coefficients. Qual Life Res. 2007;16(5):815–22.
25. Verdam MGE, Oort FJ, Sprangers MAG. Structural equation modeling-based effect-size indices were used to evaluate and interpret the impact of response shift effects. J Clin Epidemiol. 2017;85:37–44. https://doi.org/10.1016/j.jclinepi.2017.02.012.
26. Barber MD, Walters MD, Cundiff GW, Group PT. Responsiveness of the Pelvic Floor Distress Inventory (PFDI) and Pelvic Floor Impact Questionnaire (PFIQ) in women undergoing vaginal surgery and pessary treatment for pelvic organ prolapse. Am J Obstet Gynecol. 2006;194(5):1492–8. https://doi.org/10.1016/j.ajog.2006.01.076.
27. Ousmen A, Conroy T, Guillemin F, Velten M, Jolly D, Mercier M, et al. Impact of the occurrence of a response shift on the determination of the minimal important difference in a health-related quality of life score over time. Health Qual Life Outcomes. 2016;14(1):167. https://doi.org/10.1186/s12955-016-0569-5.
28. Allison PJ, Locker D, Feine JS. Quality of life: a dynamic construct. Soc Sci Med. 1997;45(2):221–30.
29. Tinetti A, Weir N, Tangyotkajohn U, Jacques A, Thompson J, Briffa K. Help-seeking behaviour for pelvic floor dysfunction in women over 55: drivers and barriers. Int Urogynecol J. 2018; https://doi.org/10.1007/s00192-018-3618-2.