ORIGINAL ARTICLE

CrossMark

Validation of the Chinese version of the Pelvic Floor Distress Inventory-20 (PFDI-20) according to the COSMIN checklist

Yidi Ma[1] · Tao Xu[1] · Ye Zhang[1] · Meng Mao[1] · Jia Kang[1] · Lan Zhu[1]

Abstract

Introduction and hypothesis The objective of this study was to translate the short version of the Pelvic Floor Distress Inventory (PFDI-20) into Chinese and to evaluate its psychometric properties in Chinese women with symptomatic pelvic floor dysfunction according to the Consensus-Based Standards for the Selection of Health Status Measurement Instruments (COSMIN) checklist.

Methods Between October 2017 and May 2018, a cross-sectional analysis of the clinical data of 126 patients who met the inclusion criteria was performed. The patients completed the questionnaires at the baseline (T1), 1–2 weeks later (T2), and 3 months after surgery (T3). Reliability testing included internal consistency, test–retest reliability, and measurement error. The methodical tests for validity were content validity, criterion validity, construct validity, and hypothesis testing. Responsiveness was also taken into consideration.

Results One hundred twenty-six patients completed all questionnaires. Internal consistency, measured by Cronbach's alpha value, was good, and the test–retest reliability was high, with an intraclass correlation coefficient (ICC) of 0.99. Construct validity was verified by factor analysis. All assumptions were confirmed, and there were no ceiling or floor effects in this study. Spearman's correlation coefficient between the PFDI-20 and the Pelvic Floor Impact Questionnaire (PFIQ-7) was 0.867, showing a significant correlation. Furthermore, the minimal important change (MIC) of 50.0 was less than the smallest detectable change (SDC) of 18.36, indicating the sufficient responsiveness.

Conclusions The Chinese version of the PFDI-20 developed in this study is a reliable and valid instrument that provides good responsiveness to clinical changes.

Keywords PFDI-20 questionnaire · Chinese validation · COSMIN checklist

Introduction

Pelvic floor dysfunction (PFD) is a clinical syndrome characterized by a variety of symptoms relating to pelvic organ prolapse (POP), urinary incontinence (UI), fecal incontinence (FI), and difficulty in emptying and sensing the lower urinary tract. The first three symptoms are the most common, with separate incidence rates of 11.4–9.56%, 30.9% [1], and 1.28% [2] in China. PFD is not fatal but significantly impacts women's quality of life (QOL).

With increasing attention paid to pelvic floor disorders, a series of validated self-administered QOL questionnaires have been developed to access individual symptoms [3]. The Pelvic Floor Distress Inventory-20 (PFDI-20) is a condition-specific short form of the Pelvic Floor Impact Questionnaire (PFDI), which was recommended by the International Continence Society (ICS) as a class A questionnaire and is currently widely used [4]. The questionnaire not only contains all items of the Urinary Distress Inventory (UDI), which was frequently used in assessing UI several years ago, but also includes items regarding POP and anorectal dysfunction. It has been divided into three subscales related to POP symptoms, colorectal symptoms, and urinary symptoms. PFDI-20 has been recently translated and validated in many different countries [5–11]. In China, the Pelvic Floor Impact Questionnaire-7 (PFIQ-7) was validated in 2011 by Zhu L et al. [12], yet there was no Chinese version of the PFDI-20. Measurement properties of the instrument were evaluated according to the Consensus-

✉ Lan Zhu
zhu_julie@vip.sina.com

1   Department of Obstetrics and Gynecology, Peking Union Medical
    College Hospital, Peking Union Medical College, Chinese Academy
    of Medical Sciences, No.1 Shuai Fu Road, Dongcheng District,
    Beijing, China

Based Standards for the Selection of Health Status Measurement Instruments (COSMIN) checklist [13]. The checklist is based on an international Delphi study in 2010, and it is used to evaluate the methodological quality of studies on health status measurement instruments [14].

The objective of this study was to translate the short version of the PFDI-20 into Chinese and to evaluate its psychometric properties in Chinese women with symptomatic PFD according to the COSMIN checklist.

Materials and methods

Questionnaires

In addition to the PFDI-20, the following questionnaires were used in this study:

Pelvic Floor Impact Questionnaire

The PFIQ-7, including three corresponding subscales (urinary, colorectal–anal and POP) assesses the impact of the condition on four aspects of patient QOL (physical activity, travel, social/relationships, and emotional health). The Chinese version was validated by Zhu et al. [12] and has high reliability and validity in the Chinese population.

Subjective assessment

The Patient Global Impressions of Improvement (PGI-I) questionnaire is a one-item questionnaire that asks patients to rate the perceived change in response to therapy. Patients are asked to check the one number that best describes how their symptoms are now compared with how they were before surgery. A 7-point scale score was given as very much better, much better, a little better, no change, a little worse, much worse, or very much worse. It has been validated in clinical studies with stress urinary incontinence (SUI) [15] and urogenital prolapse [16].

Translation process

To maintain as much original meaning as possible, the PFDI-20 translation contains two dependent forward and backward translations [17]. First, the English version was translated into Chinese independently by two bilingual translators. These two versions were reviewed by a group of researchers to increase the face validity of the questionnaire. Second, the resultant translation was back-translated into English by two other bilingual experts. Finally, a consensus was established by a panel of bilingual translators and clinical experts. A pretest of 50 eligible PFD patients was performed to determine whether the questionnaire had unclear or vague items and whether its

completion time was appropriate. We then synthesized and developed the final version, the translation process of which was modified based on cross-cultural adaption (see Appendix for the final Chinese version).

Validation study

A cross-sectional study was conducted between October 2017 and May 2018 to evaluate the reliability, validity, and responsiveness of the PFDI-20. Inclusion criteria were a diagnosis of PFD or UI and sufficient reading and comprehension abilities. Patients were excluded if they had chronic inflammation and organ lesions. All patients who completed the PFDI-20 and PFIQ-7 in an outpatient setting were grouped as T1 patients. After 1–2 weeks, (T2) they were asked to complete the questionnaire again by telephone if there were no symptomatic changes and no interventions were taken. In the third round, patients undergoing surgeries were asked to complete the PFDI-20 and PGI-I to evaluate responsiveness (T3). The validation process is shown in Fig. 1.

Statistical analysis

Statistical analysis was performed using the SPSS software package (version 23.0, SPSS Inc., Chicago, IL, USA). Descriptive data are presented as means ± standard deviation (SD) or medians (25th percentile, 75th percentile). The chi-squared test was used for univariate associations, and the Mann–Whitney U test was used for comparisons of independent groups. $P<0.05$ was considered significant. Psychometric properties were evaluated as recommended in the COSMIN checklist. Methodological testing, including reliability, validity, and responsiveness, was assessed.

Reliability (internal consistency, test–retest reliability)

Reliability, free from measurement error, concerns the ability to distinguish patients from each other. We calculated internal consistency, test–retest reliability, and measurement error to evaluate reliability. Cronbach's alpha was calculated for PFDI-20 scores as a measure of internal consistency. A value of ≥0.70 WAS considered adequate [14]. Test–retest reliability was evaluated with intraclass correlation coefficients (ICCs) to quantify agreement of total and subscale scores. The preferable range for ICC is >0.70 [14].

Measurement error

Measurement error, the systematic and random error of a patient's score that is not attributed to true changes in the
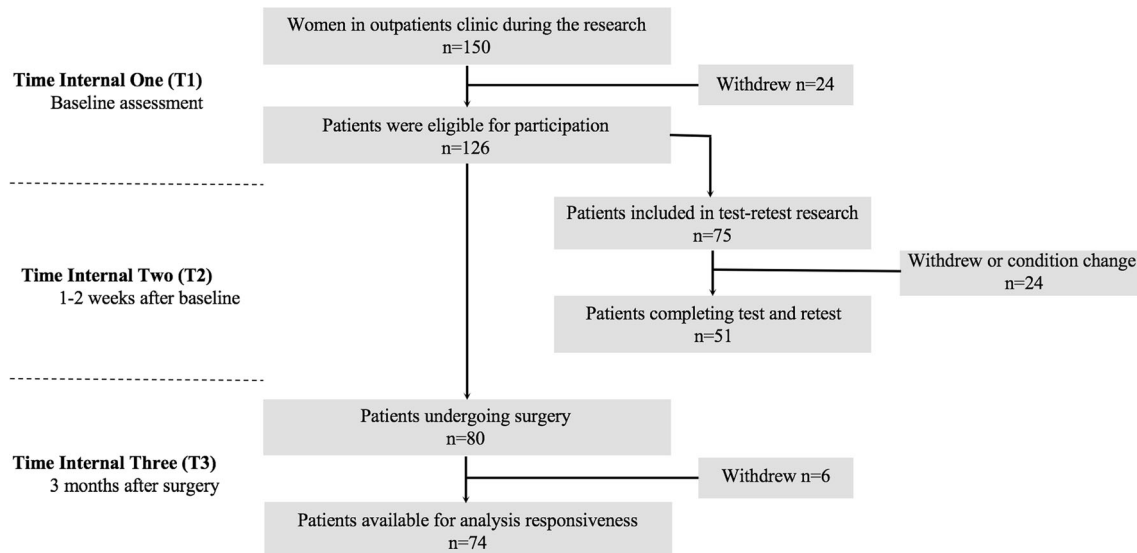
Fig. 1 Verification process of the Pelvic Floor Distress Inventory-20 (PFDI-20)

construct, can be expressed as the standard error of measurement (SEm) and the smallest detectable change (SDC). Data from T1 and T2 were used to determine measurement error. We assumed there would be no real change in a patient's level of function in a 1- to 2-week interval, and change scores presented as normally distributed and close to zero. The SEm represents the SD of repeated measures in one patient and was calculated using the square root of the error variance [13]. The SDC represents the smallest individual change that a patient needs to show on the scale to ensure that the observed change is real. The SDC was calculated using the formula $SDC = 1.96 \times \sqrt{2} \times SEM/\sqrt{n}$ [13].

Validity (face validity, construct validity, criterion validity, and hypothesis testing)

Face/content validity was assessed by an expert panel and/or patient focus group during development of the original questionnaire and the Chinese translation version. To evaluate criterion validity, Spearman's correlation coefficient was calculated between total/subscale scores of the PFDI-20 and a related criterion standard, the PFIQ-7 questionnaire, which had already been validated in our country for the assessment of PFD [12]. Corrected item-total correlations >0.70 were considered evidence of validity. Factor analysis was used as a tool for estimating construct validity. We hypothesized that patients who had UI and lower urinary tract symptoms (LUTS) would have higher UDI-6 scores than those who did not and that patients who experienced POP would have higher POPDI-6 scores than those who did not. Hypothesis testing was adequate if 75% of these hypotheses were confirmed, and the sample size of each group was required to be >50 [14]. In

addition, floor and ceiling effects and percentage of patients obtaining minimum and maximum scores were calculated, and >15% was considered problematic.

Responsiveness

Responsiveness is the sensitivity of the PFDI-20 to clinically significant changes. A comparison between SDC and MIC was carried out to demonstrate responsiveness. SDC less than MIC was considered adequate. An anchor-based method was used to calculate the MIC. T3 patients were asked to complete the PGI-I and were grouped as the answered rate. Comparisons between different groups were conducted to demonstrate that the instrument can detect clinical changes.

Results

Study population and protocol

Between October 2017 and May 2018, 150 patients were invited to participate in the survey. A total of 126 completed all questionnaires, and data analysis was conducted. Mean patient age was 58.7 ± 10.5 years. Thirty-nine (29.8%) had symptoms of UI, one of whom had anal incontinence. Sixty-four participants (48.9%) had LUTS. A total of 89 patients (67.9%) felt vaginal/uterus prolapse when performing physical work or in the resting state. Among all participants, POP-Q III was the most prevalent finding (57.3%). Demographic data and score results are shown in Table 1. Seventy-five patients were selected randomly for retest analysis to complete the questionnaires again 1–2 weeks later (this interval is considered

Table 1 Characteristics of the 126 participants

Variables	Characteristics	Statistics
Age (years)		58.7 ± 10.5
BMI (kg/m^2)		24.7 ± 3.1
Symptom	Urinary incontinence (%)	39
	Lower urinary tract syndrome (%)	64
	Pelvic organ prolapse (%)	89
	Fecal incontinence (%)	1
	Missing	5
POP-Q stage[a]	1st degree	26 (19.8%)
	2nd degree	7 (0.1%)
	3rd degree	75 (57.3%)
	4th degree	12 (0.1%)
	Missing	3 (0.0%)
Category of POP stage II and higher [b]	Cystocele (anterior compartment)	85
	Rectocele (posterior compartment)	81
	Apical prolapse (middle compartment)	31
Questionnaire scores	PFDI-20	69.9 (32.75, 93.75)
	POPDI-6	29.8 (12.50, 42.75)
	CRADI-8	9.9 (0.00, 12.50)
	UDI-6	29.5 (8.25, 42.75)

The data are presented as the means ± standard deviation, number (%), or medians (25%,75%)

POP pelvic organ prolapse, *POP-Q* Pelvic Organ Prolapse Quantification system, *PFD* pelvic floor disorders, *BMI* body mass index, *PFDI-20* Pelvic Floor Distress Inventory, *POPDI-6* Pelvic Organ Prolapse Distress Inventory, *CRADI-8* Colorectal Anal Distress Inventory, *UDI-6* Urinary Distress Inventory

[a] Several patients had POP in more than one compartment, and the highest stage reported in any compartment was recorded

[b] Several patients may have had organ prolapse in multiple compartments, and the most serious compartment was shown

short enough to avoid changes in presenting symptoms and long enough for patients to forget their previous responses), while 24 of them were excluded for undergoing conservative treatment, such as Kegel's exercises or drug therapy. Of the original respondents, 80 patients underwent PFD surgery, including procedures such as vaginal hysterectomy, anterior/posterior colporrhaphy, laparoscopic sacrocolpopexy, colpocleisis, Total Prolift System surgery, and tension-free vaginal tapes; 74 required a follow-up visit after 3 months (response rate = 92.5%). Questionnaires containing missing items or unclear individual information were excluded. Baseline information of six patients who were withdrawn in T3 were secondarily analyzed, and there was no significant difference between them and the original respondents. All missing data occurred randomly.

Reliability

The value of Cronbach's alpha was adequate for internal consistency in the PFDI-20 ($\alpha = 0.88$), POPDI-6 ($\alpha = 0.77$), UDI-6 ($\alpha = 0.80$), and CRADI-8 ($\alpha = 0.84$) (Table 2). Meanwhile,

Table 2 Internal consistency and reproducibility of the PFDI-20

	Internal consistency (Cronbach's alpha)	Test-retest ($n = 55$) ICC	P value
PFDI-20	0.88	0.997	<0.05
POPDI-6	0.77	0.994	<0.05
CRADI-8	0.84	0.997	<0.05
UDI-6	0.80	0.995	<0.05

PFDI-20 Pelvic Floor Distress Inventory, *POPDI-6* Pelvic Organ Prolapse Distress Inventory, *CRADI-8* Colorectal Anal Distress Inventory, *UDI-6* Urinary Distress Inventory

Table 3 Spearman's correlation coefficients between PFDI-20 and PFIQ-7

Correlation between questionnaires		Correlation coefficient (r)	P value
The whole scale	PFDI-20 and PFIQ-7	0.87	<0.05
Each subscale	PFDI-20 POPDI-6 and PFIQ-7 POPIQ-7	0.70	<0.05
	PFDI-20 CRADI-8 and PFIQ-7 CRAIQ-7	0.56	<0.05
	PFDI-20 UDI-6 and PFIQ-7 UIQ-7	0.81	<0.05
Internal subscale	PFDI-20 and POPDI-6	0.73	<0.05
	PFDI-20 and CRADI-8	0.69	<0.05
	PFDI-20 and UDI-6	0.84	<0.05

PFDI-20 Pelvic Floor Distress Inventory, *PFIQ-7* Pelvic Floor Impact Questionnaire, *POPIQ-7* Pelvic Organ Prolapse Impact Questionnaire, *CRAIQ-7* Colorectal Anal Impact Questionnaire, *UIQ-7* Urinary Impact Questionnaire, *POPDI-6* Pelvic Organ Prolapse Distress Inventory, *CRADI-8* Colorectal Anal Distress Inventory, *UDI-6* Urinary Distress Inventory

there was no sign of growth by deleting any item. In the test–retest analysis, the instrument showed good reliability. The total PFDI-20 showed an ICC of 0.997, and a range from 0.994 to 0.997 was found in its subscales (Table 2). The SEm was 49.1, and the SDC indicating the smallest individual change was 18.36.

Validity

Content validity

The floor/ceiling effect is also an important component of content validity. Three patients (0.02%) scored the maximum score of 300, which rejected the presence of a floor effect for PFDI-20. There was no relevant ceiling effect, because no patients scored zero (0.00%).

Criterion validity

The assessment of criterion validity was analyzed by the correlation between scores on the PFDI-20 and the PFIQ-7 (Table 3). Spearman's correlation coefficient between the two questionnaires was 0.87 and ranged from 0.56 to 0.81 on the subscales, demonstrating good criterion validity for PFDI-20. Meanwhile, in this instrument, the total score correlated well with its respective subscales.

Construct validity

Confirmed factor analysis (CFA) with Varimax rotation was employed to assess construct validity. As shown in Table 4, data were suitable for factor analysis by the Kaiser-Meyer-Olkin (KMO) and Bartlett test (KMO measure 0.821). CFA provided five factors explaining 69.55% of the variance accumulatively (cutoff point eigenvalue >1.0), which indicates a good factor extraction. In Table 5, the five factors explained 22.12%,

40.22%, 53.67%, 62.84%, and 69.55% of the variance. Table 5 gives factor loading of the Varimax-rotated five-factor solution. Questions 7–13 had high factor loadings on the first factor, which could be explained as colorectal–anal distress. Questions 1–3, 5, 15, 19, and 20 loaded high on the second factor, which could be classified into direct feelings of organ prolapse and lower urinary tract obstruction or irritation symptoms. Questions 16–18 had high loading on the third factor and could be classified as various types of UI. Questions 4 and 6 belonged to a fourth factor: excretion with external force. Only question 14 independently belonged to the fifth factor: rectocele symptoms. All factor contributions of the variance ranged from 45 to 90%. Although this structure is not exactly the same as the original version, its logical structure indicated that this instrument has good construct validity.

Hypothesis testing

We hypothesized that patients with POP would have higher POPDI-6 scores than those without these symptoms. Patients with UI or LUTS had higher UDI-6 scores than those without those symptoms. Hypothesis testing was adequate if 75% of the hypotheses were confirmed. All predefined hypotheses were confirmed, as shown in Fig. 2. Validity analysis of the CRADI subscale was not

Table 4 Kaiser-Meyer-Olkin (KMO) and Bartlett test of confirmed factor analysis

Factor analysis		
KMO measure of sampling adequacy		.821
Bartlett test of sphericity	Approximately χ^2	1340.65
	df	190
	Significance	.00

Table 5 Results of factor analysis within the five PFDI-20 dimensions

Questions	Five-factor solution				
	1	2	3	4	5
1. Usually experience pressure in the lower abdomen	−0.02	**0.70**	−0.10	0.14	−0.11
2. Usually experience heaviness or dullness in the pelvic area	0.02	**0.82**	0.02	0.14	−0.10
3. Usually have a bulge or feeling of something falling out	0.09	**0.53**	−0.43	0.26	0.37
4. Ever have to push on the vagina or around the rectum	0.38	0.28	0.05	**0.75**	0.02
5. Usually experience a feeling of incomplete bladder emptying	0.32	**0.67**	0.22	0.11	−0.12
6. Ever have to push up on a bulge in the vaginal area	0.15	0.21	−0.09	**0.79**	0.30
7. Feel you need to strain too hard to have a bowel movement	**0.45**	0.33	0.18	0.35	−0.24
8. Feel you have not completely emptied your bowels	**0.63**	0.24	0.23	0.44	−0.29
9. Usually lose stool beyond your control if your stool...	**0.85**	0.09	0.05	0.02	−0.00
10. Usually lose stool beyond your control if your stool is loose	**0.84**	0.14	0.14	−0.02	0.10
11. Usually lose gas from the rectum beyond your control	**0.82**	0.04	0.08	−0.08	0.12
12. Usually have pain when you pass your stool	**0.74**	0.04	−0.13	0.28	−0.12
13. Experience a strong sense of urgency	**0.68**	0.21	0.06	0.01	0.34
14. Does part of your bowel ever pass...	0.20	−0.02	0.10	0.15	**0.83**
15. Usually experience frequent urination	0.25	**0.67**	0.34	0.06	0.19
16. Usually experience urine leakage	0.15	0.39	**0.73**	−0.02	0.04
17. Usually experience urine leakage	0.12	0.03	**0.89**	−0.02	0.04
18. Usually experience small amounts of urine leakage	0.20	0.03	**0.90**	−0.06	−0.04
19. Usually experience difficulty emptying bladder	0.45	**0.55**	0.23	0.03	0.08
20. Usually experience pain or discomfort	0.25	**0.68**	−0.05	−0.29	0.14
Nonrotated solution					
Eigenvalue	6.66	2.50	2.16	1.32	1.28
Variance explained (%)	**33.29**	**12.50**	**10.80**	**6.57**	**6.38**
Cumulative variance explained (%)	33.29	45.79	56.60	63.17	**69.55**

Factor loadings >0.40 are in bold. The five nonrotated factors explained 33.29%, 12.50%, 10.80%, 6.57%, and 6.38%, together representing 69.55% of item variance

included in the hypothesis testing because of the low prevalence of FI, which led to an inadequate number of patients.

Responsiveness and interpretability

In the third investigation section, 80 participants received surgical treatment, 74 of whom completed the questionnaires (response rate = 92.5%) and were grouped by PGI-I scores. As shown in Fig. 3, there was a significant difference in scores in the very much better ($p < 0.05$) and much better groups ($p < 0.05$), and there was no statistical significance in scores in the no-change group ($p = 0.10$). The group responding a little better was excluded because of small sample size. In addition, the score difference was more obvious in the very much better group than in the much better group, indicating the ability to transform a qualitative effect into a quantitative one.
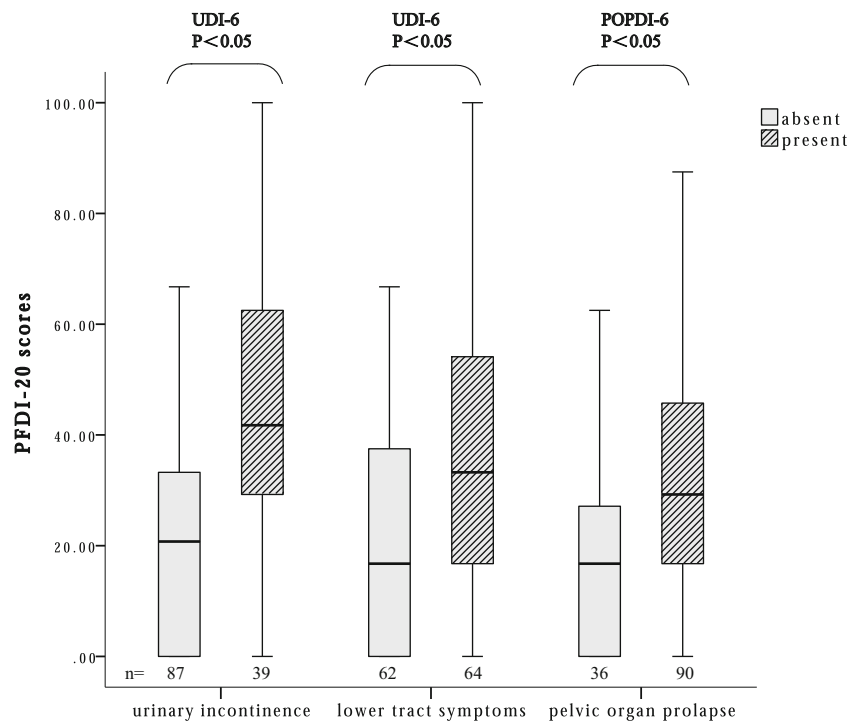
There was no gold standard for calculating the MIC according to the COSMIN checklist. MIC was estimated with an anchor-based method. The 95% confidence interval (CI) of effect size (ES) of the much better and no change groups was calculated; the cutoff point should be outside the 95% CI of the no change group and also be the smallest for the much better group. Therefore, the estimated ES in our population was 1.86 (Fig. 4), and the MIC of the PFDI-20 was 50.0. The Chinese version of the PFDI-20 showed a lower MIC value than the SDC, and the responsiveness was adequate based on Terwee et al. [14].

Discussion

The purpose of this study was to translate the PFDI-20 into Chinese and validate it in Chinese women. The psychometric properties included reliability, validity, responsiveness, and interpretability. Cronbach's coefficient of 0.875 showed a satisfactory internal consistency of the PFDI-20. Similar values were reported in Japan, Brazil, and African countries. The second round of investigation

Fig. 2 Hypotheses testing. Pelvic Floor Distress Inventory (PFDI)-20 scores with comparisons between groups. The *bold lines* present the median, the *box* represents the interquartile range, and the *whiskers* represent the minimum and maximum scores

was conducted by phone and used to assess the test–rest reliability. The PFDI-20 showed excellent test–retest reliability, with observed ICC values of 0.997.

Because there is no gold standard for PFD symptoms, we calculated correlations between PFDI-20 and the validated PFIQ-7 to estimate criterion validity. Spearman correlation coefficient of 0.867 presented an adequate result, and there were significant correlations between subscale scores on both instruments, except the FI subscales ($r = 0.559$). Moreover, adequate correlations of

questionnaires with similar structures of the UI, FI, and POP subscales further verified construct validity, similar to some validation studies, such as those conducted in Holland, Brazil, and some African countries [7, 8, 18]. In the study reported here, factor analysis was further employed to evaluate construct validity. Structure of the PFDI-20 is similar to that of the original version, the PFDI, which includes 46 items and three subscales (UDI, POPDI, CRADI). Among them, the UDI retains all three original structures (obstruction, irritation, and

Fig. 3 Responsiveness and interpretability. Comparison of Pelvic Floor Distress Inventory (PFDI)-20 scores before and after operation in each Patient Global Impression of Improvement (PGI-I) group. The *bold lines* present the median, the *box* represents the interquartile range, and the *whiskers* represent the minimum and maximum score
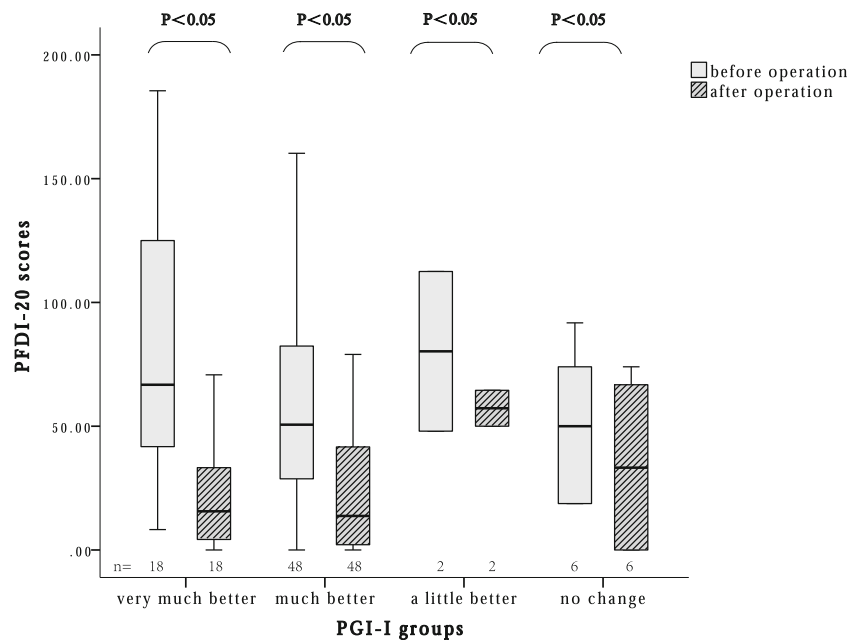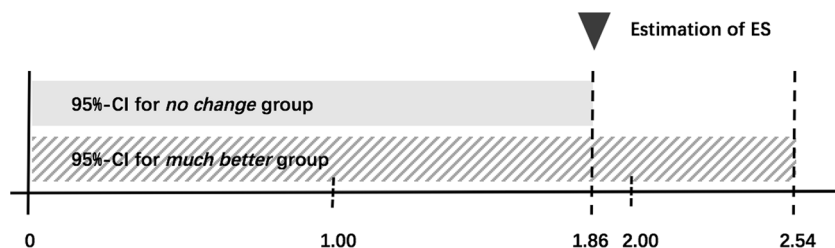
Fig. 4 Estimation of effect size by indication of the 95% confidence interval (CI) for no change and much better groups

stress) described by Shumaker et al. [19] and was expanded using nine items related to LUTS, which are common in PFD patients. The POPDI consists of 16 items divided into three parts (overall, anterior, and posterior compartment); the CRADI consists of 17 items associated with lower gastrointestinal disorders, which are divided into four parts (obstruction, incontinence, pain/irritation, and rectal prolapse). The CFA eventually explained 69.55% of the item variance, while five factors were found to largely match the logical dimensions, strongly supporting the construct validity of the PFDI-20.

Evaluation of treatment efficiency is considered extremely important and, as the basis of modern evidence-based medicine, shows strong potential in clinical work. SDC and the MIC are the basis for patient-reported outcome measures (PROM) interpretability to determine if the observed changes can benefit patients. To determine the clinical significance of score changes at the individual level, measurement error needs to be assessed and should not be greater than the MIC. Otherwise, the observed change cannot be determined to be a real change, as the risk of measurement error is >5%. One method of reducing SDC is to decrease the measurement error by averaging the testing values of repeated measurements, which leads to an additional burden on patients and increases the change of recall bias. Improving the quality of questionnaires by adding or improving items may be an alternative approach.

There is currently no consensus on choosing an anchor point for calculating the MIC. The slightly improved group was considered to effectively reflect the smallest important clinical change. Some studies have used the 15-point scale to apply the average change in the almost the same or slightly better or worse group to represent the MIC. We used the much better group of the 7-point PGI-I scale in this responsiveness analysis because only one patient was in the slightly improved group. Finally, adequate responsiveness was demonstrated with a higher MIC value (50.0) compared with SDC (18.36), indicating that we can determine 95% of the change caused by nonmeasurement error when the patient's score change is greater than the MIC. Grouping comparisons by the PGI-I was also performed to assess preoperative and postoperative score changes. Results showed a statistically significant decreasing trend in the significant improvement and much better groups, but especially in the significant improvement group.

No statistically significant difference in the no change group before and after operation was observed.

There were some limitations in this study. First, although there are no studies showing differences in MIC in distinct surgical or nonsurgical groups, some people believe that analysis and assessment between different groups is needed [20]. In our study, only patients undergoing surgical interventions were analyzed, and T3 patients should be divided into groups by intervention methods for future research. Second, although the anchor technique is considered to be the best method for evaluating MIC, its effectiveness and the best calculation methods remain controversial [21]. Furthermore, due to the small sample size of the slightly improved group, we chose the much better group as the second-best method of calculating MIC, which means it was inevitably greater than the true value. Finally, sample sizes of the UI and POP groups in the hypothesis testing were <50, meaning that those groups did not strictly meet inclusion criteria. There were not enough patients with FI to evaluate psychometric properties of the CRADI-8 subscale, meaning that validity of that criterion was unsatisfactory. Future research should focus on popularity of the PGI-I and enlargement of the sample size.

To conclude, the Chinese version of the PFDI-20 is a reliable and valid instrument and can considerably contribute to improvement in Chinese PFD patients.

Compliance with ethical standards

Conflicts of interest   None.

Appendix (Chinese questionnaires)

盆底窘迫量表简易问卷(Pelvic Floor Distress Inventory-20, PFDI-20).

回答一下调查问卷的所有问题,涉及最近3个月的膀胱、直肠和盆腔的症状,分列为POPDI-6,CRADI-8、UDI-6三个栏目。 如果您有以下症状,请选择影响程度。 每项选择的分值标在□ (0~4分),分数越高对生活质量影响越大。.

盆腔器官窘迫量表 6（POPI-6）

Q1. 经常体验到下腹腹压吗?

□0，没有；□有；如果有，对您的影响如何:

□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

Q2. 经常感到盆腔坠胀吗?

□0，没有；□有；如果有，对您的影响如何:

□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

Q3. 经常看到或感到阴道有肿物脱出吗?

□0，没有；□有；如果有，对您的影响如何:

□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

Q4. 曾经需要推压阴道或直肠周围来协助排便吗?

□0，没有；□有；如果有，对您的影响如何:

□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

Q5. 经常有膀胱排尿不尽的感觉吗?

□0，没有；□有；如果有，对您的影响如何:

□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

Q6. 经常不得不用手指托起阴道的膨出部分来协助排尿吗?

□0，没有；□有；如果有，对您的影响如何:

□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

肛肠窘迫量表 8（CRADI-8）

Q7. 便秘，排便困难

□0，没有；□有；如果有，对您的影响如何：

　　　　　　□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

Q8. 无法排尽大便

□0，没有；□有；如果有，对您的影响如何：

　　　　　　□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

Q9. 在大便成形的情况下，经常不能控制排便

□0，没有；□有；如果有，对您的影响如何：

　　　　　　□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

Q10. 当大便松散时，经常不能控制排便

□0，没有；□有；如果有，对您的影响如何：

　　　　　　□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

Q11. 经常不能控制肛门排气

□0，没有；□有；如果有，对您的影响如何：

　　　　　　□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

Q12. 经常在排便时感到疼痛

□0，没有；□有；如果有，对您的影响如何：

□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

Q13.排便急迫，不得不奔向卫生间去排便

□0，没有；□有；如果有，对您的影响如何：

□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

Q14.在排便时或之后感到有肠管从直肠脱出

□0，没有；□有；如果有，对您的影响如何：

□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

尿窘迫量表 6（UDI-6）

Q15.经常感到尿频吗？

□0，没有；□有；如果有，对您的影响如何：

□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

Q16.经常有与排尿急迫相关的漏尿吗？急迫就是必须立刻去卫生间排尿的强烈感觉

□0，没有；□有；如果有，对您的影响如何：

□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

Q17.经常有咳嗽、打喷嚏或大笑引起的漏尿吗

□0，没有；□有；如果有，对您的影响如何：

□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

Q18. 经常有少量漏尿吗（点滴漏尿）？

□0，没有；□有；如果有，对您的影响如何：

□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

Q19. 经常排空膀胱有困难吗？

□0，没有；□有；如果有，对您的影响如何：

□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

Q20. 经常感到下腹或生殖道不适吗？

□0，没有；□有；如果有，对您的影响如何：

□1，没有影响；□2，轻度影响；□3，中度影响；□4，重度影响

得出每项栏目的平均分(0~4) × 25(0~100),相加得出总评分(0~300). 总评分：

References

1. Zhu L, et al. The epidemiological study of women with urinary incontinence and risk factors for stress urinary incontinence in China. Menopause. 2009;16(4):831–6.
2. Chan SS, et al. Prevalence of urinary and fecal incontinence in Chinese women during and after their first pregnancy. Int Urogynecol J. 2013;24(9):1473–9.
3. Barber MD. Questionnaires for women with pelvic floor disorders. Int Urogynecol J Pelvic Floor Dysfunct. 2007;18(4):461.
4. Barber MD, et al. 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;193(1):103.
5. Treszezamsky AD, et al. Spanish translation and validation of four short pelvic floor disorders questionnaires. Int Urogynecol J. 2013;24(4):655–70.
6. Yoshida M, et al. Reliability and validity of the Japanese version of the pelvic floor distress inventory-short form 20. Int Urogynecol J. 2013;24(6):1039–46.
7. Utomo E, et al. Validation of the pelvic floor distress inventory (PFDI-20) and pelvic floor impact questionnaire (PFIQ-7) in a Dutch population. Int Urogynecol J. 2014;25(4):531–44.
8. Arouca MA, et al. Validation and cultural translation for Brazilian Portuguese version of the pelvic floor impact questionnaire (PFIQ-7) and pelvic floor distress inventory (PFDI-20). Int Urogynecol J. 2016;27(7):1097–106.

9. Teig CJ, et al. Norwegian translation, and validation, of the pelvic floor distress inventory (PFDI-20) and the pelvic floor impact questionnaire (PFIQ-7). Int Urogynecol J. 2017;28(7):1005–17.

10. Teleman P, et al. Validation of the Swedish short forms of the pelvic floor impact questionnaire (PFIQ-7), pelvic floor distress inventory (PFDI-20) and pelvic organ prolapse/urinary incontinence sexual questionnaire (PISQ-12). Acta Obstet Gynecol Scand. 2011;90(5): 483–7.

11. Due U, et al. Validation of the pelvic floor distress Inventory-20 and the pelvic floor impact questionnaire-7 in Danish women with pelvic organ prolapse. Acta Obstet Gynecol Scand. 2013;92(9):1041–8.

12. Zhu L, et al. Chinese validation of the pelvic floor impact questionnaire short form. MENOPAUSE. 2011;18(9):1030–3.

13. Mokkink LB, 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.

14. Terwee CB, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34–42.

15. Yalcin I, et al. Validation of two global impression questionnaires for incontinence. Am J Obstet Gynecol. 2003;189(1):98–101.

16. Srikrishna S, et al. Validation of the patient global impression of improvement (PGI-I) for urogenital prolapse. Int Urogynecol J. 2010;21(5):523–8.

17. Acquadro C, et al. Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. Value Health. 2008;11(3):509–21.

18. Henn EW, et al. Validation of the PFDI-20 and PFIQ-7 quality of life questionnaires in two African languages. Int Urogynecol J. 2017;28(12):1883–90.

19. Shumaker SA, et al. Health-related quality of life measures for women with urinary incontinence: the incontinence impact questionnaire and the urogenital distress inventory. Continence program in women (CPW) research group. Qual Life Res. 1994;3(5):291–306.

20. King MT. A point of minimal important difference (MID): a critique of terminology and methods. Expert Rev Pharmacoecon Outcomes Res. 2011;11(2):171–84.

21. Terwee CB, et al. Mind the MIC: large variation among populations and methods. J Clin Epidemiol. 2010;63(5):524–34.