The Fecal Incontinence Quality of Life Scale (FIQL) and Fecal Incontinence Severity Index (FISI): Validation of the Dutch Versions

Lisette A. 't Hoen, Elaine Utomo, Willem R. Schouten, Bertil F.M. Blok and Ida J. Korfage.

Neurourol Urodyn. 2017 Mar; 36(3):710-715

ABSTRACT

Aims

Fecal incontinence (FI) is known to have a major impact on quality of life. The Fecal Incontinence Quality of Life scale (FIQL) and Fecal Incontinence Severity Index (FISI) have been developed to assess this impact. The aim of this study was to validate the FIQL and FISI in the Dutch language.

Methods

The study population consisted of women and men experiencing FI and a reference group. The FIQL (four domains) and FISI questionnaires were validated by testing standardized measurement properties: discriminative ability, internal consistency, reproducibility, construct validity, and responsiveness.

RESULTS

A total of 55 patients and 277 reference participants were included. Patients had significant lower and higher scores at the FIQL and FISI, respectively, than references (FIQL: 2.58±0.70 and 3.92±0.36, FISI: 38.57±10.73 and 23.17±15.01; p<.001), indicating worse functioning in patients and with this the discriminative abilities of the measures. The FIQL demonstrated adequate internal consistency on all domains (0.72–0.96), except for the embarrassment domain (0.55). The reproducibility was good for both measures. A negative correlation was found between the FIQL and FISI. Furthermore, the FIQL showed a positive (0.77) and the FISI a negative correlation -0.31) with the Mental Component Summary scale of the SF-12. Responsiveness analysis showed a minimal important change of 0.40 points for the FIQL.

CONCLUSION

Validity and reliability were good in the Dutch FIQL, but inconclusive in the FISI. The Dutch FIQL can support physicians in determining the impact of FI on patient's quality of life.

INTRODUCTION

The spectrum of fecal incontinence (FI) comprises the involuntary loss of gas, liquid, and solid stool.[1] Prevalence figures differ from 1.4% to 42%, depending on the age and sex of the population and the definition of incontinence.[2-4] The incidence of FI is higher in women and higher with age. FI is known to have a major impact on the quality of life (QoL) in patients, resulting in psychological symptoms, functional loss in daily life, and even social isolation.[5] Furthermore, FI is associated with other pelvic floor disorders such as urinary incontinence and pelvic organ prolapse.[6]

To assess the severity of FI and its impact on QoL, standardized and validated instruments are needed, which can also be used to diagnose FI and to assess the effectiveness of FI therapy. It is important to establish that measures adequately evaluate the topic of interest before they can be used in clinical practice. Standardized validation processes have been developed to assess this ability.[7] Several validated questionnaires are available in English, such as the International Consultation on Incontinence Modular Questionnaire-Bowel symptoms (ICIQ-BS),[8] Manchester Bowel Questionnaire, [9] Fecal Incontinence Quality of Life scale (FIQL)[10] and the Fecal Incontinence Severity Index (FISI).[11]

We have chosen to use the FIQL and the FISI because of the grade B recommendation by the International Continence Society, prevalent use in clinical research about pelvic floor disorders, and personal preference. The FIQL was also validated in various other languages.[12–17] The FIQL is a condition-specific quality of life measure to assess FI.[10] The FISI is a scoring system to assess the severity of symptoms of FI.[11] To enable the use of proper measures in Dutch clinical practice, versions of the FISI and FIQL have yet to be translated and validated. The aim of this study is to provide validated versions of the Dutch FISI and FIQL following a standardized procedure to evaluate the measurement properties.[18]

METHODS

This study is part of a large health-related QoL study among patients visiting a tertiary pelvic floor center.[19–21] Approval was obtained by the local medical research ethics committee (MEC-2008-376) and guidelines following good clinical practice were followed.

Linguistic Validation

A standardized guideline was used for the translation of the FIQL and the FISI measures.[22] Three native Dutch each independently translated the English FIQL and FISI to

Dutch. After discussion and revision of differences, a native English speaker performed a backward translation. In a pilot study, 10 patients participated in a face-to-face validation and changes were made according to their feedback. The final versions of the FISI and FIQL were then used for this study (See "Vragenlijsten").

Study Population

Patient group

Women and men were eligible for inclusion when they experienced symptoms of FI for at least 3 months, spoke fluently Dutch, and were over 18 years old. Patients with current malignancy, dementia, mental retardation, and/or another neurological disease were excluded. Eligible patients were informed about this study by their physician during a regular outpatient visit. If interested in participation, they received an information package containing the consent form and the first two questionnaires. After written informed consent was given, these questionnaires were to be filled out during the inclusion visit, and 1 week after inclusion. The second set of questionnaires was sent out through postal mail and completed 6 months after inclusion. Treatment was not initiated during the first week of study participation. An extra question taken from the RAND 36-Item Health Survey, the RAND 36- Health Transition Item (RAND 36-HTI), was added to the last questionnaires.[23] This question compares the general health of patients to 1 year ago.

Reference group

An ISO-certified (ISO 26362) panel was used as a reference group. Stratification for age, educational level, and residence was performed to create a representative group of the Dutch general population. Beforehand, the presence of FI was unknown.

Questionnaire

The questionnaire consisted of three measures:

- The FIQL is a condition-specific measure that evaluates QoL in patients who experience FI. The measure consists of 29 questions, subdivided into four subscales: Lifestyle, Coping/Behaviour, Depression/Self-perception, and Embarrassment. With a few exceptions, responses are graded on a 4-point Likert-scale ranging from 1 "strongly agree" to 4 "strongly" disagree. The additional response option "not applicable" is graded with a score of 4, in concordance with Dr. Rockwood, developer of the FIQL.[10] Question 1, general health, is graded from 1 "excellent" to 5 "poor" and is reversely scored. Question 4, FI specific depression, is graded from 1 "extremely so" to 6 "not at all." The scale scores are calculated by adding the numerical values of all responses in that specific scale and then dividing by its

number of items. Scale scores are only calculated if at least half of the items have been answered. Higher scores indicate a better QoL.[10]

- The FISI is a severity rating score for FI. This score evaluates the frequency of four types of FI: gas, mucus, and liquid and solid stool. Frequency options range from "two or more times a day" to "never." Surgeon and patient specific ratings are available. In this study, the scores were based on patient specific ratings. Higher scores indicate more severe FI.[11]
- The Short Form Health Survey (SF-12) is a general health QoL questionnaire. Based on the responses two summary measures can be calculated, the physical component summary (PCS-12) and the mental component summary (MCS-12).[24] The SF-12 was distributed to the patient group only.

Measurement Properties

The following measurement properties were evaluated following standardized quality criteria[18]:

Reliability

Internal consistency

This indicates the extent to which the items in a (sub) scale assess the same construct. A Cronbach's alpha of >0.70 was considered to show adequate and >0.50 to show moderate internal consistency.[18]

Reproducibility

This indicates the degree to which scores in a stable patient remain similar on repeated measurements. The intraclass correlation coefficient (ICC) was calculated to evaluate test-retest reliability, that is, the extent to which patients can be differentiated from each other, despite measurement error. A value of at least 0.70 was required for adequate test-retest reliability.[25] Agreement concerns the similarity between scores on repeated measures. The measurement error (systematic and random error of a score not due to true changes) was quantified as limits of agreement (LOA). The LOA represent the range of a specific percentage of the differences between two sets of scores and are calculated as the mean change in score ±1.96 * standard deviation (SD) of the changes.[26] The impact of the measurement error has been calculated by relating the range of the LOA to the range of the total score of the measure.

Validity

Content validity

This is the extent to which the measure addresses the concepts of interest in the target population. Face validity is the degree to which the measures indeed presents an ad-

equate reflection of the construct to be measured, in this case clinical symptoms and quality of life dependent on FI.[18] Floor or ceiling effects are present if more than 15% of the responders reported the highest or the lowest score.[18] Floor and ceiling effects were determined for total and scale scores.

Criterion validity

This is the correlation between the measure of interest and a gold standard. No perfect gold standard exists for the FIQL or FISI. The SF-12 was chosen as a substitute, because it is a commonly used generic health-related QoL measures and consists of a physical and a mental component scale, which could correspond to the FISI and FIQL, respectively. The correlation between the FIQL and FISI was determined as well. Pearson's *r* was used to calculate the correlations between the different measures.[18]

Construct validity

This is the extent to which predefined hypotheses about the scores of the FIQL and FISI measures in relation to other measurements could be confirmed. If at least 75% of these hypothesis were correct, construct validity was considered adequate.[18] We formulated the following hypotheses:

a. FISI and FIQL scores will be negatively correlated.
b. FISI scores will be higher in patients than in the reference group.
c. FIQL scores will be lower in patients than in the reference group.
d. In patients FIQL scores will be positively correlated with MCS-12 scores.

Responsiveness

This is the ability of a measure to detect clinical important change over time. This was only calculated for patients who received treatment. The area under the receiver operating characteristic (ROC) curve (AUC) shows the ability of a measure to distinguish patients who have improved from those who have not. The RAND 36-HTI was used as an external criterion to determine the AUC for the FIQL and FISI scores. An AUC of at least 0.50 was considered adequate.[18]

Interpretability

This is the degree to which a qualitative meaning can be assigned to the quantitative scores of the measure. To indicate a true clinical relevant improvement we used the minimal important change (MIC), defined as the optimal ROC cut-off point; the value for which the sum of proportions for misclassifications ([1-sensitivity]+[1-specificity]) was smallest.[7] The LOA should be smaller than the MIC.[18]

Statistical Methods

To determine an adequate sample size quality criteria were followed, which stated that at least 50 patients were required.[18] We reported continuous data as mean and SD and the categorical data as numbers and percentages. Differences between patient and reference group were evaluated using the Student's t-test and Chi-square test for continuous and categorical variables, respectively. One-way analysis of variance (ANOVA) was used for the evaluation of more than two independent groups. Treatment options were divided into three subgroups; conservative, pharmaceutical, and surgical. A two-sided *p*-value of <.05 was considered significant. Statistical analysis was performed using SPSS version 21.0 (IBM Crop., Armonk, NY).

RESULTS

Out of the 91 patients interested in study participation, 55 (60%) completed the first questionnaire (Figure 4.1). Table 4.1 displays the characteristics of patient and reference group (n=277) and shows a higher percentage of women in the patient group than in the reference group. The total score and four subscale scores of the FIQL are significantly lower in the patient group, indicating a worse quality of life due to FI. The FISI score is significantly higher in the patient group, indicating more severe symptoms of FI.

Reliability

Internal consistency

For the FIQL the Cronbach's alpha ranged from 0.55 to 0.96 on the total score and on the four separate domains in the patient group. It was below 0.70 for the embarrassment domain only, indicating moderate internal consistency for that domain and adequate internal consistency for the total score and the three other domains. For the reference group the Cronbach's alpha for the FIQL were adequate as well (Table 4.2). The FISI had a low Cronbach's alpha of 0.22 in patients, while it had a moderate Cronbach's alpha of 0.66 in the reference group (Table 4.2).

Reproducibility

The test retest period had a mean duration of 7.5 days. Table 4.2 demonstrates the excellent reliability for the total (ICC 0.95) and domain scores (ICC 0.80–0.95) of the FIQL. A good reliability was shown for the FISI (ICC 0.72) as well. The impact of the measurement error is expected to be 25% when relating the range of the LOA (-0.51 to 0.47) to the range of the total FIQL score (0–4). For the FISI the impact of the measurement error is expected to be 48%, LOA (-16.26 to 13.02).

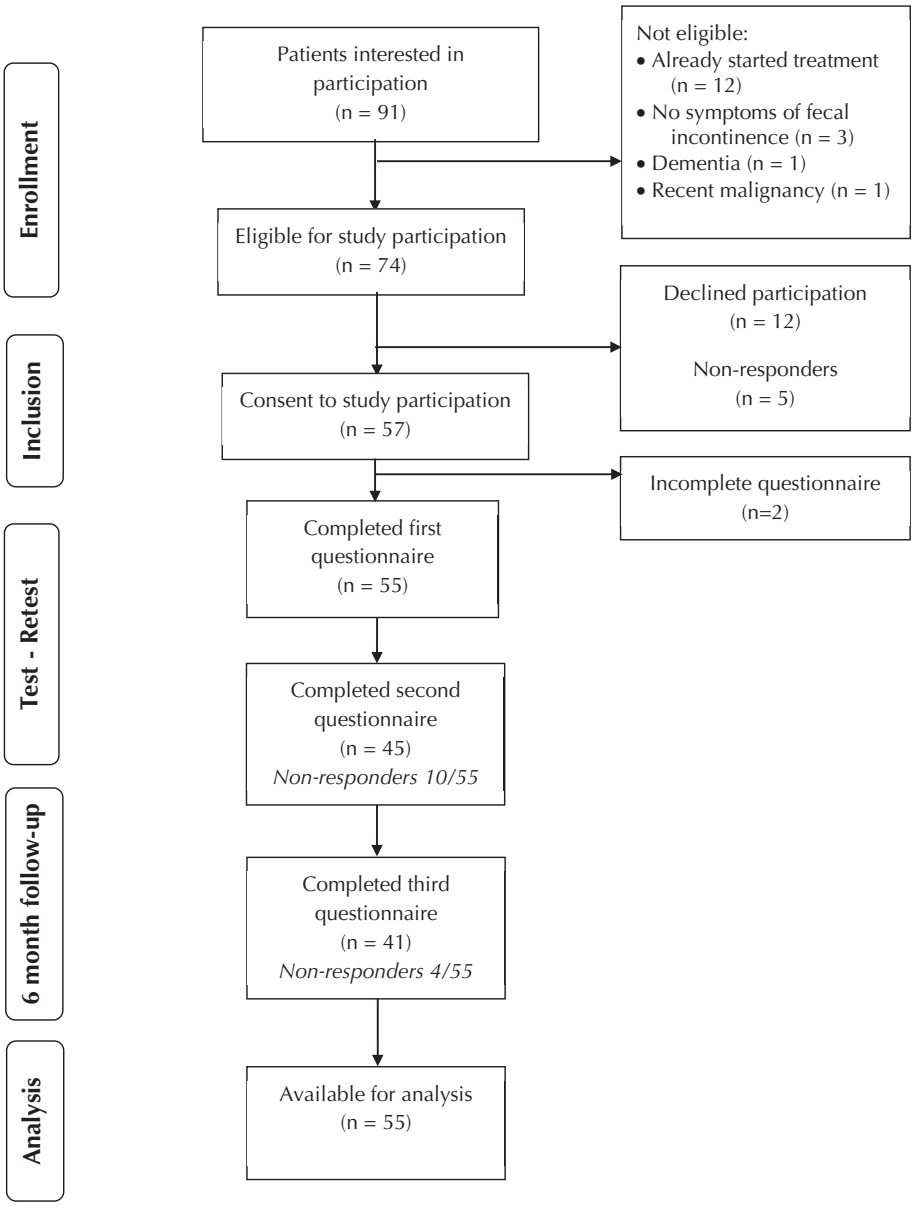

Figure 4.1 Flowchart: inclusion process of patient group

Validity

Content validity

Face validity was considered adequate for the FIQL and FISI by the researchers and 10 patients during the pilot study. For the total score of the FIQL, a ceiling effect was observed in 68% of the reference group, reflecting the low prevalence of FI in this group.

Table 4.1 Demographic and Clinical Characteristics of the Study Group (Patient vs Reference).

	Patient group (n=55)	Reference group (n=277)	*p*-value
Gender, number of women (%)	52 (95)	141 (51)	< .001
Age (years), mean±SD	59.3 ± 12.45	48.7 ± 16.2	< .001
Educational level, number(%)			
Lower	20 (36)	59 (21)	.01
Middle	25 (45)	133 (48)	
Higher	10 (18)	85 (31)	
Scores, mean ± SD			
FISI[a]	38.57 ± 10.73	23.17 ± 15.01	<.001
FIQL[b]	2.58 ± 0.70	3.92 ± 0.36	<.001
Lifestyle	2.82 ± 0.93	3.92 ± 0.33	<.001
Coping/behaviour	2.07 ± 0.73	3.85 ± 0.43	<.001
Depression/self-perception	3.20 ± 0.98	4.04 ± 0.41	<.001
Embarrassment	2.23 ± 0.80	3.86 ± 0.43	<.001
Scores at 6 month follow-up, mean±SD	(n=45)		
FISI[a]	35.91 ± 14.41		.16[d]
FIQL[b]	2.72 ± 0.66		.01[d]
Lifestyle	2.97 ± 0.76		.02[d]
Coping/behaviour	2.24 ± 0.77		.01[d]
Depression/self-perception	3.33 ± 0.79		.17[d]
Embarrassment	2.37 ± 0.89		.06[d]
Scores SF-12, mean±SD	(n = 23)		
PCS-12[c]	41.50 ± 11.66		
MCS-12[c]	46.26 ± 13.39		

[a] Higher FISI scores indicate more severe FI.
[b] Higher FIQL scores indicate a better FI quality of life.
[c] Scores higher than 50 indicates better quality of life, scores lower than 50 indicates poorer quality of life.
[d] Difference between baseline and 6 month follow-up.

For the FISI, a floor effect was observed in the reference group (16%), again reflecting the low prevalence of FI.

Criterion validity

In the patient group no correlation was found between FIQL (or any of its domain scores) and FISI, while in the reference group significant negative correlations were found (Table 4.3). Correlations of the FIQL and the FISI with the PCS-12 were insignificant (Table 4.3). Significant positive correlations were found for all four domains and total score of the FIQL with the MCS-12, and a significant negative correlation was found for the FISI and MCS-12.

Table 4.2 The Internal Consistency is presented with the Cronbach's alpha for total and domain scores

	Cronbach's alpha Patients	Referents	Test	Retest	Intraclass correlation coefficient (95%CI)	Mean change ± SD	Limits of agreement[a]
FIQL Total (n=40)	0.95	0.96	2.61 ± 0.69	2.64 ± 0.71	0.95 (0.90 – 0.97)	-0.02 ± 0.25	-0.51 – 0.47
Lifestyle	0.93	0.94	2.82 ± 0.93	2.78 ± 0.97	0.94 (0.89 – 0.97)	0.04 ± 0.33	-0.61 – 0.68
Coping/behaviour	0.86	0.93	2.08 ± 0.76	2.17 ± 0.82	0.93 (0.88 – 0.96)	-0.09 ± 0.29	-0.66 – 0.48
Depression/self-perception	0.89	0.79	3.27 ± 0.92	3.26 ± 0.89	0.95 (0.90 – 0.97)	0.01 ± 0.30	-0.58 – 0.60
Embarrassment	0.55	0.72	2.30 ± 0.80	2.36 ± 0.88	0.80 (0.67 – 0.89)	-0.07 ± 0.53	-1.11 – 0.97
FISI Total (n=42)	0.22	0.66	37.83 ± 10.18	39.45 ± 9.57	0.71 (0.53 – 0.84)	-1.62 ± 7.47	-16.26 – 13.02

The ICC and LOA scores reflect the reproducibility of both questionnaires. [a] Limits of agreement is calculated as: mean *change* ± 1.96 * SD*change*

Table 4.3 The correlation between FIQL total, FIQL domains, and FISI in the patient group (n=55) and the reference group (n=277), *the numbers in italic* presents the clinical correlation of the FIQL.

	FIQL Total	Lifestyle	Coping / behavior	Depression/self-perception	Embarrassment	FISI
FIQL Total	1.00					
	1.00					
Lifestyle	0.87[a]	1.00				
	0.57[a]	*1.00*				
Coping/behavior	0.82[a]	0.81[a]	1.00			
	0.66[a]	*0.71[a]*	*1.00*			
Depression/self-perception	0.89[a]	0.69[a]	0.67[a]	1.00		
	0.86[a]	*0.37[a]*	*0.37[a]*	*1.00*		
Embarrassment	0.64[a]	0.32[b]	0.30[b]	0.48[a]	1.00	
	0.57[a]	*0.56[a]*	*0.59[a]*	*0.32[a]*	*1.00*	
FISI	-0.17[c]	-0.16[c]	-0.22[c]	-0.22[c]	-0.02[c]	1.00
	-0.25[a]	*0.28[a]*	*-0.32[a]*	*-0.30[a]*	*-0.25[a]*	*1.00*
PCS-12	0.15[c]	0.18[c]	0.19[c]	0.28[c]	-0.36[c]	-0.36[c]
MCS-12	0.77[a]	0.58[a]	0.68[a]	0.70[a]	0.46[b]	-0.31[c]

The correlation between the FIQL and FISI, and PCS-12 and MCS-12 in the patient group only (n=23) was calculated to reflect the criterion validity.
[a] $p < .01$ [b] $p < .05$ [c] $p > .05$

Construct validity

Three of four predefined hypotheses were confirmed, indicating good construct validity.

a. An insignificant correlation was found between FIQL and FISI scores, and the hypothesis could therefore not be confirmed.
b. FISI scores of patients were higher than those of the reference group.
c. FIQL scores of patients were lower scores than those of the reference group.
d. The MCS-12 was positively correlated with the FIQL on all domains.

Responsiveness

At 6 month follow-up 45 patients completed the FIQL and FISI, of whom 32 patients had received treatment. Treatment was conservative in 17 patients, pharmaceutical in nine patients and surgical in six patients. The change in scores is shown in Table 4.1. The AUC for the FIQL in this group was 0.69, with a *p*-value of .08 (Table 4.4). For the FISI an AUC of 0.45 was found, with a *p*-value of .62.

Interpretability

Table 4.4 displays a MIC for the FIQL of 0.40 with 63% correctly identified as improved and 75% as not improved. The MIC is just within the range of the LOA, indicating that a change of 0.40 points could possibly be attributed to measurement error. For the FISI a MIC of 11.50 with 100% correctly identified as improved and 16% as not improved is found. This MIC is within the LOA range, indicating that with a change of 11.50 points a true improvement cannot be de distinguished from a measurement error.

Table 4.4 The FIQL and FISI scores in patients who have received treatment and their corresponding RAND-36 response reflect the responsiveness and interpretability.

	Number (%) (n=31)[a]	FIQL scores (n=31)[b]	FISI scores (n=31)[c]
RAND-36 health transition item			
Much worse / a little worse	5 (16)	0.23 ± 0.88	2.80 ± 9.42
Same	15 (48)	0.28 ± 0.34	-8.21 ± 15.85
A little better	5 (16)	0.40 ± 0.39	1.60 ± 6.43
Much better	6 (19)	0.74 ± 0.45	-11.67 ± 15.27
Area under the ROC curve		0.69	0.45
p-value		.08	.62
Minimal important change		0.40	11.50
Sensitivity; specificity		0.63 ; 0.75	1.00 ; 0.16

The RAND-36 functions as an anchor.

Data presented are in number (%) or $mean_{change}$ ±SD_{change} between baseline and follow-up at 6 months.

[a]Responsiveness has only been reported for the 31 patients who received treatment.

[b]Positive scores indicate an improvement in quality of life.

[c]Negative scores indicate an improvement in symptoms.

DISCUSSION

With this study, we have addressed the need for validated measures in the Dutch language for the assessment of the impact and severity of FI. We found an adequate reliability and validity for the FIQL, for the FISI these remain inconclusive. The results differed significantly between patient and reference groups for both measures, indicating worse quality of life and more severe symptoms of FI in the patient group, thereby demonstrating their discriminative power.

The Cronbach's alpha ranged between 0.55 and 0.95 for the total score and the four separate domains, which is similar to other validation studies.[12,13,15,17] The internal consistency was, therefore, adequate for all but one domain. The Cronbach's alpha for the embarrassment scale of 0.55 was below the minimum of 0.70 that is considered adequate. Such a low Cronbach's alpha has previously also been reported in other validation studies.[12,15,17] This embarrassment scale consists of (only) three questions. It addresses two unrelated aspects of shame, which may have led to a diverse pattern of answers and thus a relatively low Cronbach's alpha. In our opinion, the lower internal consistency for this scale is not problematic.

The low internal consistency we found for the FISI could (partially) be explained by the fact that the FISI is a scoring system for symptoms of FI. The construct measured by the FISI is observable and it is therefore clear that each item of the FISI contributes to the construct of FI. Guidelines state that low internal consistency does not have to be problematic if the construct to be measured is evident.[7]

For both measures the high ICC indicates an excellent reproducibility. The ICC for the FIQL total score and its four different domain scales range from 0.80 to 0.95, similar to previous validation studies.[12–14,17] The embarrassment scale showed a good reproducibility (ICC 0.80), even though it showed a sub adequate internal consistency (Cronbach's alpha 0.55). This strengthens the earlier assumption that the low internal consistency is acceptable. Future research could focus on this embarrassment scale by assessing the impact of redistribution of the scale items, deletion of the scale or patient interviews.

Similar to other validation studies, a scoring system was used to assess the correlation between clinical symptoms and the FIQL questionnaire. The FISI measure was chosen for this study, other studies have used both the FISI and Wexner scores.[13,15,17] Negative correlations between the FIQL and FISI were seen in both patient and reference groups, however, these were only significant in the reference group. The insignificant correlation in the patient group might be explained by the small sample size. This is the only predefined hypothesis that was not confirmed for the construct validity.

We found a good correlation for the FIQL and the FISI with the MCS-12. This was expected, given that the four domains of the FIQL mainly address the experience of

FI and not so much the (physical) symptoms. In other validation studies, which used the SF-36 to determine criterion validity,[13,15] positive correlations were found between the FIQL and some of the physical domain scores of the SF-36, however, we could not repeat this since we used the PCS-12. Therefore, an adequate comparison with the other validation studies cannot be made.

Responsiveness analysis of the FIQL showed a borderline significant AUC of 0.69 (p=.08), indicating a good correlation of the FIQL with clinical change determined with the RAND 36-HTI. The small number of patients available for this analysis might have contributed to the borderline significant p-value for the FIQL, and also have caused the AUC of 0.45 for the FISI. In addition, the weak correlation between the FISI and the RAND 36-HTI could be caused by the more divergent FISI scores. However, we could not elucidate the cause of this weak correlation. The difference between the constructs to be measured, that is, the experience versus the actual severity of the FI symptoms, could explain the difference found during the responsiveness and interpretability analysis with the RAND 36-HTI for the FIQL and FISI.

The strengths of this study are the use of standardized measurement properties to evaluate the quality of the Dutch versions of the FIQL and FISI and the use of a reference group to assess the ability of both measures to detect clinically relevant differences. Limitations of this study include the small number of patients who received treatment during the 6 months follow-up. This complicated the adequate evaluation of responsiveness and interpretability for the FISI.

CONCLUSION

The Dutch version of the FIQL showed good reliability and validity for use in patients with FI. The measurement properties of the FISI were found to be inconclusive. It is recommended to confirm the responsiveness of the FIQL and measurement properties of the FISI in a bigger cohort in future research.

Acknowledgments
The authors thank Dr. P. Dewint for the recruitment of patients.

REFERENCES

1. Counihan TC MR. Fecal incontinence. Current therapy in Colon and Rectal Surgery. 2nd ed. Philadelphia: Elsevier Mosby; 2005:105-12.
2. Perry S, Shaw C, McGrother C, et al. Prevalence of faecal incontinence in adults aged 40 years or more living in the community. Gut 2002;50:480-4.
3. Bartlett L, Nowak M, Ho YH. Impact of fecal incontinence on quality of life. World J Gastroenterol 2009;15:3276-82.
4. Saga S, Vinsnes AG, Morkved S, Norton C, Seim A. Prevalence and correlates of fecal incontinence among nursing home residents: a population-based cross-sectional study. BMC Geriatr 2013;13:87.
5. Brown HW, Wexner SD, Segall MM, Brezoczky KL, Lukacz ES. Quality of life impact in women with accidental bowel leakage. Int J Clin Pract 2012;66:1109-16.
6. Bezerra LR, Vasconcelos Neto JA, Vasconcelos CT, et al. Prevalence of unreported bowel symptoms in women with pelvic floor dysfunction and the impact on their quality of life. Int Urogynecol J 2014;25:927-33.
7. de Vet HC, et al. Measurement in medicine. Cambridge: Cambridge University Press; 2011.
8. Abrams P, Avery K, Gardener N, et al. The International consultation on incontinence modular questionnaire: qqq.iciq.net. J Urol 2006; 175:1063-6.
9. Kwon S, Visco AG, Fitzgerald MP, et al. Validity and reliability of the modified Manchester health questionnaire in assessing patients with fecal incontinence. Dis Colon Rectum 2005; 48:323-31.
10. Rockwood TH, Church JM, Fleshman JW, et al. Fecal Incontinence Quality of Life Scale: quality of life instrument for patients with fecal incontinence. Dis Colon Rectum 2000;43:9-16; discussion -7.
11. Rockwood TH, Church JM, Fleshman JW, et al. Patient and surgeon ranking of the severity of symptoms associated with fecal incontinence: the fecal incontinence severity index. Dis Colon Rectum 1999;42:1525-32.
12. Rullier E, Zerbib F, Marrel A, Amouretti M, Lehur PA. Validation of the French version of the Fecal Incontinence Quality-of-Life (FIQL) scale. Gastroenterol Clin Biol 2004;28:562-8.
13. Minguez M, Garrigues V, Soria MJ, Andreu M, Mearin F, Clave P. Adaptation to Spanish language and validation of the fecal incontinence quality of life scale. Dis Colon Rectum 2006;49:490-9.
14. Yusuf SA, Jorge JM, Habr-Gama A, Kiss DR, Gama Rodrigues J. [Evaluation of quality of life in anal incontinence: validation of the questionnaire FIQL (Fecal Incontinence Quality of Life)] Avaliacao da qualidade de vida na incontinencia anal: validacao do questionario FIQL (Fecal Incontinence Quality of Life). Arq Gastroenterol 2004;41:202-8.
15. Ahnis A, Holzhausen M, Rockwood TH, Rosemeier HP. FLQAI - A Questionnaire on Quality of Life in Fecal Incontinence: German translation and validation of Rockwood et al. (2000) Fecal Incontinence Quality of Life Scale (FIQLS). Z Gastroenterol 2012;50:661-9.
16. Dehli T, Martinussen M, Mevik K, et al. Translation and validation of the Norwegian version of the fecal incontinence quality-of-life scale. Scand J Surg 2011;100:190-5.
17. Ogata H, Mimura T, Hanazaki K. Validation study of the Japanese version of the Faecal Incontinence Quality of Life Scale. Colorectal Dis 2012;14:194-9.
18. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34-42.

19. Utomo E, Korfage IJ, Wildhagen MF, Steensma AB, Bangma CH, Blok BF. Validation of the Urogenital Distress Inventory (UDI-6) and Incontinence Impact Questionnaire (IIQ-7) in a Dutch population. Neurourol Urodyn 2015;34:24-31.
20. Utomo E, Blok BF, Steensma AB, Korfage IJ. Validation of the Pelvic Floor Distress Inventory (PFDI-20) and Pelvic Floor Impact Questionnaire (PFIQ-7) in a Dutch population. Int Urogynecol J 2014;25:531-44.
21. 't Hoen LA, Utomo E, Steensma AB, Blok BF, Korfage IJ. The Pelvic Organ Prolapse/Urinary Incontinence Sexual Questionnaire (PISQ-12): validation of the Dutch version. Int Urogynecol J 2015;26:1293-303.
22. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993;46:1417-32.
23. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-Item Health Survey 1.0. Health Econ 1993;2:217-27.
24. Ware J, Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996;34:220-33.
25. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods 1996;1:30-46.
26. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-10.