Development and psychometric evaluation of a dutch questionnaire for the assessment of anorectal and lower urinary tract symptoms

M. Wyndaele[1], B.Y. De Winter[2], S. Van Roosbroeck[3], M. Van Outryve[4], S. De Wachter[1], G. Van Hal[3], P. Pelckmans[2], J.J. Wyndaele[1]

(1) Department of Urology, (2) Laboratory of Experimental Medicine and Pediatrics, and Laboratory of Gastroenterology, (3) Department of Epidemiology and Social Medicine, University of Antwerp, Belgium ; (4) Department of Gastroenterology, University Hospital Antwerp, Belgium.

Abstract

Background and study aims : **Epidemiological studies have shown a frequent coexistence of symptoms and diseases affecting the anorectum and lower urinary tract. To further investigate combined symptoms and pathology of both pelvic viscera we developed a self-reported questionnaire, in Dutch, which extensively evaluates habits, complaints and symptoms of both viscera. We describe the construction and the psychometric properties of this questionnaire.**

Patients and methods : **This prospective study was conducted in 56 patients with anorectal symptoms, 41 patients with lower urinary tract symptoms and in a control group of 91 people. The following psychometric properties of the questionnaire were evaluated : content validity, construct validity, criterion validity, test-retest reliability and internal consistency.**

Results : **The questionnaire covered all important domains, was well interpreted and showed good acceptability (content validity). The questionnaire clearly differentiated the patient populations (construct validity). The criterion validity of the questionnaire was excellent. The test-retest reliability of the questionnaire was acceptable in all three the study populations (overall median kappa : 0.64 ; Inter Quartile Range : 0.56-0.75 ; mean agreement : 88%). The internal consistency of both anorectal and lower urinary tract symptom questions was high (Crohnbach's alpha of 0.78 and 0.80 respectively).**

Conclusions : **This questionnaire is a valid and reliable instrument for the assessment of anorectal and lower urinary tract symptoms. It can provide further insights into the epidemiology of concomitant bowel and bladder disorders and, accordingly, can contribute to a more efficient diagnostic and therapeutic approach in patients with such disorders.** (Acta gastroenterol. belg., **2011**, 74, **295-303**).

Key words : Questionnaire, validation, bladder, bowel.

Introduction

Epidemiological studies have shown that symptoms and diseases affecting the anorectum and the lower urinary tract (LUT) frequently coexist. Worldwide, community-based studies show a high prevalence of combined fecal (FI) and urinary (UI) incontinence (1,2). Additionally, a high incidence of FI in women with LUT disorders (3-10), and an association of urinary tract problems with irritable bowel syndrome have been described (11).

Both anorectal and LUT symptoms (LUTS) have a negative impact on the patients' quality of life (QoL) (12), and, concomitant pathology in the other organ system is known to further reduce the QoL (13,14). The relation between the anorectum and the LUT has therapeutic consequences as well : treatment of constipation in children can lead to a resolution of UI and can prevent further recurrence of urinary tract infections (15). On the other hand, treatment of overactive bladder (OAB) with bladder relaxing drugs can cause constipation (16).

Many questionnaires have been designed to assess bowel and urinary symptoms independently. However, because of the interrelationships between both pelvic viscera, it is important to integrate the initial evaluation of anorectal and LUT disorders, rather than segregate them along traditional specialty boundaries (17). To serve that purpose, several questionnaires that assess pelvic floor symptoms and bother have been developed and validated. The Pelvic Floor Disorders Distress Inventory (PFDI), the Pelvic Floor Disorders Impact Questionnaire (PFIQ) and the International Consultation on Incontinence Questionnaire (ICI-Q) are the most referenced (18,19). Based on these existing questionnaires, we constructed a self-reported questionnaire, in Dutch, updated to the current standards in both gastroenterology (20,21) and urology (22). The questionnaire extensively evaluates habits, complaints and symptoms of both viscera, as well as possible risk factors for combined pathology (i.e. age, obstetric and surgical history, use of medication). No psychometrically validated versions of the PFDI, the PFIQ or the ICI-Q in Dutch could be found to date. We created our questionnaire to serve a scientific purpose : to assess the prevalence, the characteristics as well as the risk factors of concomitant functional bowel and bladder complaints. In contrast to the PFDI, the PFIQ and the ICI-Q, our questionnaire is not a condition-specific outcome measure and should therefore not be used to evaluate the state, impact or evolution of pelvic floor disorders in patients. However as it assesses the presence of complaints and symptoms in both organ systems, it can be used clinically by physicians as a lead for additional investigations.

Correspondence to : Jean-Jacques Wyndaele, Department of Urology, University Hospital Antwerp, Wilrijkstraat 10, B-2650 Edegem, Belgium.
E-mail : jeanjacques.wyndaele@ua.ac.be

Submission date : 01/09/2010
Acceptance date : 27/01/2011

In this study we describe the construction and the psychometric properties of this questionnaire in Dutch for the assessment of anorectal and lower urinary tract symptoms. According to the guidelines by the ICI, and in line with the psychometric validation of the PFDI and the PFIQ, the following properties of the questionnaire were tested : content validity, construct validity, criterion validity, test-retest reliability and internal consistency (18,19).

Material and methods

Questionnaire Design

The questionnaire, constructed in Dutch in 2006, is based on the PFDI, PFIQ and the ICI-Q (18,19). To cover an as wide as possible range of anorectal symptoms and LUTS, and to ensure that the items questioned correspond with the current guidelines for both anorectal and LUT disorders, an extensive literature review has been done. The gastroenterological questions are consistent with the Rome III diagnostic criteria for functional bowel and anorectal disorders (20,21), the urological questions with the guidelines proposed by the International Continence Society (ICS) in 2005 (22). To further ensure that the questionnaire reflects the content domain of functional bowel and bladder disorders and their combination, we consulted the gastroenterologists, coloproctologists and urologists of the pelvic clinic of our hospital. We constructed the questionnaire to be easily understood, while maintaining clarity in text size and structure. The questionnaire is designed to be self-completed, which is the preferred mode of administration for both anorectal and LUT questions (23,24).

The questionnaire is available upon request to the corresponding author. A short introductory text to the ten-page questionnaire (A4 size) explains the purpose of the questionnaire and states that its content is subject to medical secrecy. The questionnaire is composed of five sections : impact on QoL (one question), anorectal symptom questions (17 main questions), LUTS questions (16 main questions), obstetric history (one main question) and medical history (three main questions). There are 27 or 28 (for men and women, respectively) yes/no questions, 4 questions with multiple answers (maximum four), 2 visual analogue scale questions and 4 open numerical questions (e.g. frequency). Two versions of the questionnaire were made, both containing the same questions with a different order of appearance of the anorectal symptom and LUTS questions, depending on the department the questionnaire was used in (Urology or Gastroenterology). The order presented in this manuscript is identical to the questionnaire for the Gastroenterology department. The sections on anorectal symptoms and on LUTS contain a lot of similar questions. Where possible the questionnaire was designed in such a way that both the questions and the response options followed the same format. To clarify the answer or to retrieve additional information on a complaint or symptom, 14 of the main questions can lead to sub-questions depending on the answer given.

Study Population

Ethical approval was granted by the ethical committee of our hospital and all participants gave written informed consent for this prospective study.

The questionnaire was administered to three pre-treatment patient populations :

1) 56 patients attending the Gastroenterology clinic with complaints of lower bowel or anorectal dysfunction (GASTRO)
2) 41 patients attending the Urology clinic with complaints of LUT dysfunction (URO)
3) 91 patients attending the Orthopaedic or General Internal Medicine clinic, with no primary complaints related to bowel or bladder (CONTROL)

For the first two groups, complaints related to the other organ system were not considered an exclusion criterion. The given numbers represent the patients completing the first questionnaire. The first questionnaire was completed without assistance. Insufficient Dutch proficiency was considered an exclusion criterion.

Validity

To assess the validity of the questionnaire we evaluated content validity, construct validity and criterion validity (18,19).

Content validity

To ensure that all relevant aspects of anorectal and LUT function were covered, the questionnaire was reviewed by experts from the pelvic clinic of our hospital. Participating patients were invited to comment on the content, design and clarity of the questionnaire. The incidence of missing data was measured as this is an indicator of the acceptability of the questions.

Construct validity

Construct validity refers to the relationships between the questionnaire and underlying theories (25). We examined the ability of the questionnaire to differentiate between patient groups and controls. We compared the prevalence of anorectal symptoms between the GASTRO and CONTROL groups, and of LUTS between the URO and CONTROL groups. The prevalence of UI was also compared between males and females as the prevalence of UI in women is higher at all ages (26). A consensus on the correlation between FI and gender has not been documented (27).

Criterion validity

For pelvic floor disorders, no clear gold standard exists against which to measure the criterion validity of questionnaires.

As no validated combined anorectal and LUT questionnaires exist in Dutch, we opted to test the self-reported patient's answers of a random sample of URO patients, against the responses to an interview-assisted completion of the questionnaire together with a physician.

Reliability

Test-retest reliability (stability)

The stability of the responses, over a period of one month in which the patients' symptom status was not expected to change, was tested in the three patient groups. After 3 weeks, all patients received a second questionnaire by mail. Patients were asked to return this second questionnaire on the next consultation (GASTRO and URO patients) planned one month after the initial visit, or, by mail (CONTROL). If no response was received, a third questionnaire was mailed to these patients after a 3-week interval.

To avoid the learning of answers, patients who first completed the questionnaire for the Urology department, received the questionnaire for the Gastroenterology department and vice versa. To be able to exclude patients with a change in their underlying condition, two questions were added to the second questionnaire : one regarding a possible progression in their complaints or pathology, and another on a possible treatment initiated during the time between answering the first and second questionnaire.

Internal consistency

Internal consistency refers to the extent to which items within the questionnaire are related to each other. The correlation between the anorectal symptom questions and between the LUTS questions were assessed by Crohnbach's alpha coefficient (25) using baseline data from the three patient groups. We also evaluated the psychometric benefit of sub-grouping both anorectal symptom and LUTS questions into categories for storage and evacuation.

Statistical Analysis

Statistical analyses were performed using SAS (version 9.1, SAS Institute Inc., Cary, NC, USA) and SPSS V 15 (SPSS Inc., IL, USA).

To test the ability of our questionnaire to differentiate between patient groups and controls (construct validity), we used a Chi square analysis for unpaired categorical data (statistical significance was set at $P < 0.05$).

Test-retest reliability of the questionnaire was determined by calculating the κ statistic and the associated 95% confidence intervals (95% CI) for each categorical question (28). In addition, the proportion of overall agreement for each question was presented. The criterion validity was measured in a similar way by comparing the physician's version, with the independently reported patient's responses. Values are presented as mean κ (and 95% CI limits) for individual questions, and median κ (and interquartile range, IQR) for a group of questions.

For questions with a single categorical response variable, intersource agreement (patient-patient and patient-interview) was assessed by the unweighted κ statistic, which represents the proportion of agreement beyond that expected by chance alone : it is scaled to vary from -1 to 1. Kappa coefficients were interpreted using the guidelines given by Landis and Koch (29), i.e. 0.01-0.20 as slight ; 0.21-0.40 as fair ; 0.41-0.60 as moderate ; 0.61-0.80 as substantial ; and 0.81-1.00 as almost perfect agreement. A negative κ value indicates poorer agreement than chance. Because a low κ-value may reflect a low prevalence of that symptom in the cohort and not lack of agreement, the proportion of positive and negative agreement was also calculated for questions with a high agreement and low κ (30,31). For questions with multiple categorical, not ordinate responses, intersource agreement was assessed by a weighted κ statistic (32). This weighting system assigns more weight to small degrees of disagreement than to larger degrees. The 95% CI of κ provides a test of the hypothesis that the underlying value of κ differs from zero.

To permit calculation of the κ statistic, the QoL question (visual analogue scale) was divided into three clinically relevant categories : a score of 0-3 represented little to no impact of the stool or micturition problems on daily life, a score of 4-6 represented a medium impact and a score of 7-10 represented a high impact. The question about stool type (Bristol Stool Scale) was divided into three clinically relevant stool types to allow analysis of its test-retest reliability.

The Intraclass Correlation Coefficient (ICC) was used to evaluate the stability of the questions about frequency of micturition and defecation.

The internal consistency of the questionnaire was tested with the Crohnbach's α.

Results

Of 91 CONTROL patients who completed the first questionnaire, 68 returned the second questionnaire (response rate : 75%). All the GASTRO (n = 56) and URO (n = 41) patients brought the second questionnaire to their next appointment (response rate : 100%). Consequently, the final analytic sample was composed of 165 patients (62% female), of which 56 (84% female) GASTRO patients, 41 (59% female) URO patients and 68 (46% female) CONTROL patients.

The mean overall age was 55 ± 14 years (range 16-86). The mean age was 56 (± 16), 60 (± 14) and 51(± 12) years for GASTRO, URO and CONTROL patients respectively.

Validity

Content validity

Review of the questionnaire by patients and by experts from the pelvic clinic of our hospital indicated that it was well interpreted and covered all important domains.

Table 1. — **Prevalence of anorectal symptoms as reported by the GASTRO and CONTROL groups**

Anorectal symptom	Frequency (%)		Chi Square
	GASTRO	CONTROL	
Rectal bleeding/mucus	36.4%	7.6%	0.000
Postponing defecation with desire	52.7%	6.1%	0.000
Postponing defecation with strong desire	58.5%	35.9%	0.015
Fecal incontinence	49.1%	4.7%	0.000
Flatal incontinence	64.8%	59.4%	0.544
Hesitancy	45.5%	3.0%	0.000
Use of laxatives	39.3%	4.5%	0.000
Tenesmus	56.4%	19.4%	0.000
Action required to start defecation	55.6%	12.3%	0.000
Straining to defecate	50.9%	43.9%	0.444
Anal blockage	51.8%	21.5%	0.001
Pain during defecation	42.9%	9.2%	0.000
Feeling of incomplete bowel evacuation	61.5%	18.6%	0.000

Table 2. — **Prevalence of LUTS as reported by the URO and CONTROL groups**

LUTS	Frequency (%)		Chi Square
	URO	CONTROL	
Nocturia	87.2%	60.3%	0.004
Postponing micturition with desire	36.6%	10.9%	0.002
Postponing micturition with strong desire	74.4%	35.5%	0.000
Urgency	67.5%	26.6%	0.000
Urinary incontinence	70.7%	19.4%	0.000
Hesitancy	12.2%	4.5%	0.144
Action required to start micturition	14.6%	1.6%	0.009
Straining to void	22.0%	15.6%	0.411
Dysuria	19.5%	0.0%	0.000
Need to immediately re-void	43.9%	25.8%	0.052
Feeling of incomplete bladder emptying	36.8%	7.6%	0.000
Postmicturition leakage	56.1%	49.2%	0.492
Intermittency	45.5%	15.1%	0.002
Slow stream	77.8%	42.9%	0.027

Most items demonstrated low levels of missing data (mean 3.7% ± 1.7%, range 1.2%-6.7%), indicating a good acceptability of the questions, except for the feeling of incomplete bowel emptying (9.1% missing).

Construct validity

The prevalence of anorectal symptoms reported by the GASTRO and CONTROL groups are shown in Table 1. The prevalence of LUTS in the URO and CONTROL groups are listed in Table 2. As anticipated, the questionnaire easily differentiated the populations : CONTROL patients reported significantly less anorectal symptoms and LUTS than GASTRO and URO patients, respectively. A sequence of chi-squared tests showed a significant association between response incidence and population for every item, except flatal incontinence, straining to defecate, hesitancy, straining to void and postmicturition leakage.

The LUTS questions of our questionnaire clearly differentiated between males and females, with women reporting more UI than men (49.0% and 22.6%, respectively, P = 0.001).

Criterion validity

The study sample contained 33 patients (tables 3 and 4). The overall median κ, including the QoL, obstetric history and medical history questions, was 1.00 (IQR : 1.00-1.00). There were no questions for which the 95% CI included a value of zero. The κ for criterion validity could not be calculated in 15 questions as one of the variables was a constant. The overall agreement for all sections was close to perfect, resulting in a mean agreement of 100% for most sections and of 98% for the questions regarding medical history.

Reliability

Test-retest reliability (stability)

The overall median kappa (κ) for all questions was 0.64 [Inter Quartile Range (IQR) : 0.56-0.75], indicating

Table 3. — **Test-retest reliability and criterion validity for the anorectal symptom questions in all patients (n = 165)**

Question	*Reproducibility*			*Criterion validity*		
	Observed Agreement (%)	**κ or weighted κ***	**95% CI**	**Observed Agreement (%)**	**κ or weighted κ***	**95% CI**
Change in frequency in past 3 months	82	0.53*	0.37 - 0.68	100	1.00	1.00 - 1.00
Stool Type (Bristol Stool Scale)				100	1.00	1.00 - 1.00
Type 1 or 2	93	0.66	0.48 - 0.85	100	1.00	1.00 - 1.00
Type 3 or 4	82	0.54	0.39 - 0.69	100	1.00	1.00 - 1.00
Type 5, 6 or 7	82	0.54	0.39 - 0.69	100	1.00	1.00 - 1.00
Rectal bleeding/mucus	93	0.74	0.59 - 0.88	100	1.00	1.00 - 1.00
Reason for defecation						
Clock	93	0.52	0.28 - 0.77	100	1.00	1.00 - 1.00
Precaution	**97**	**0.27**	(-0.17) - 0.71	100	/	/
Desire	78	0.42	0.26 - 0.57	100	1.00	1.00 - 1.00
Strong Desire	78	0.41	0.25 - 0.57	100	1.00	1.00 - 1.00
Postponing defecation with desire	89	0.68	0.55 - 0.82	100	1.00	1.00 - 1.00
Need to run	**83**	**-0.07**	(-0.18) - 0.04	100	1.00	1.00 - 1.00
Loss of urine	86	0.59	0.20 - 0.98	100	1.00	1.00 - 1.00
Postponing defecation with strong desire	82	0.63	0.50 - 0.75	100	1.00	1.00 - 1.00
Need to run	89	0.71	0.48 - 0.94	100	1.00	1.00 - 1.00
Loss of urine	85	0.69	0.47 - 0.92	100	1.00	1.00 - 1.00
Fecal Incontinence	94	0.81	0.69 - 0.93	97	0.78	0.38 - 1.00
Stress Incontinence	89	0.69	0.30 - 1.00	100	/	/
Urgency Incontinence	89	0.75	0.44 - 1.00	100	/	/
Other Incontinence	83	0.56	0.12 - 1.00	100	/	/
Amount of leakage	83	0.64	0.33 - 0.94	100	/	/
Liquid or solid or both	**58**	**0.28***	(-0.07) - 0.63	100	1.00	1.00 - 1.00
Use of pad	92	0.83	0.60 - 1.00	100	/	/
Effect of position change	100	1.00	1.00 - 1.00	100	/	/
Influence on daily activities	91	0.74	0.41 - 1.00	100	/	/
Flatal incontinence	85	0.69	0.58 - 0.81	100	1.00	1.00 - 1.00
Hesitancy	87	0.59	0.44 - 0.75	100	1.00	1.00 - 1.00
Use of laxatives	94	0.80	0.67 - 0.92	100	1.00	1.00 - 1.00
Tenesmus	84	0.63	0.50 - 0.76	100	1.00	1.00 - 1.00
Action required to start defecation	90	0.76	0.65 - 0.87	100	1.00	1.00 - 1.00
Position change	85	0.68	0.44 - 0.92	100	1.00	1.00 - 1.00
Straining	83	0.55	0.26 - 0.84	100	1.00	1.00 - 1.00
Reduction of coeles	90	0.65	0.34 - 0.97	100	1.00	1.00 - 1.00
Anal Digitation	95	0.86	0.66 - 1.00	100	1.00	1.00 - 1.00
Other	**83**	**0.14**	(-0.24) - 0.51	100	1.00	1.00 - 1.00
Straining to defecate	81	0.62	0.50 - 0.75	100	1.00	1.00 - 1.00
Anal blockage	90	0.80	0.70 - 0.89	100	1.00	1.00 - 1.00
Pain during defecation	88	0.63	0.48 - 0.78	100	1.00	1.00 - 1.00
Anal	86	0.59	0.20 - 0.98	100	/	/
Lower Abdomen	81	0.62	0.29 - 0.95	100	/	/
Between legs	95	0.64	0.01 - 1.00	100	/	/
Other	95	/	/	100	/	/
Feeling of incomplete bowel evacuation	84	0.61	0.47 - 0.75	100	1.00	1.00 - 1.00

κ = Mean Kappa ; 95% CI = 95% Confidence Interval ; **Bold** = κ values < 0.40 and their respective agreement.

a substantial agreement. The 95% Confidence Interval (CI) included a value of zero in only eight questions. Overall agreement was high : the questionnaire had a mean agreement of 88%.

Fifteen percent of questions had a $\kappa > 0.80$, indicating perfect agreement ; 47% between 0.61 and 0.80, indicating substantial agreement ; 30% between 0.41 and 0.60, indicating moderate agreement. Only 8% had a $\kappa < 0.40$ indicating slight or fair agreement. The κ-value could not be calculated in four questions as one of the variables was a constant (one anorectal symptom sub-question and three LUTS questions).

The proportion of overall agreement, the κ-values and their 95% CI for the categorical anorectal symptom and

Table 4. — **Test-retest reliability and criterion validity for the LUTS questions in all patients (n = 165)**

Question	*Reproducibility*			*Criterion validity*		
	Observed Agreement (%)	**κ or weighted κ***	**95% CI**	**Observed Agreement (%)**	**κ or weighted κ***	**95% CI**
Nocturia	83	0.73*	0.63 - 0.84	100	1.00	1.00 - 1.00
Reason for micturition						
Clock	**96**	**0.38**	(-0.01) - 0.77	100	1.00	1.00 - 1.00
Precaution	88	0.54	0.35 - 0.73	100	1.00	1.00 - 1.00
Desire	78	0.41	0.24 - 0.57	100	1.00	1.00 - 1.00
Strong Desire	81	0.46	0.29 - 0.63	100	1.00	1.00 - 1.00
Postponing micturition with desire	89	0.72	0.59 - 0.84	97	0.93	0.80 - 1.00
Need to run	79	0.48	0.11 - 0.84	100	1.00	1.00 - 1.00
Loss of urine	78	0.43	0.05 - 0.81	100	1.00	1.00 - 1.00
Postponing micturition with strong desire	82	0.64	0.51 - 0.76	100	1.00	1.00 - 1.00
Need to run	89	0.51	0.17 - 0.85	96	0.88	0.65 - 1.00
Loss of urine	86	0.68	0.45 - 0.90	100	1.00	1.00 - 1.00
Urgency	80	0.60	0.47 - 0.72	96	0.92	0.78 - 1.00
Urinary Incontinence	94	0.88	0.81 - 0.96	100	1.00	1.00 - 1.00
Stress Incontinence	93	0.85	0.71 - 0.99	96	0.92	0.75 - 1.00
Urgency Incontinence	85	0.70	0.52 - 0.89	100	1.00	1.00 - 1.00
Other Incontinence	**78**	**0.27**	(-0.04) - 0.58	100	1.00	1.00 - 1.00
Amount of leakage	79	0.67*	0.51 - 0.84	100	1.00	1.00 - 1.00
Use of pad	96	0.89	0.73 - 1.00	100	1.00	1.00 - 1.00
Effect of position change	84	0.59	0.33 - 0.84	100	1.00	1.00 - 1.00
Influence on daily activities	80	0.59	0.37 - 0.80	100	1.00	1.00 - 1.00
Hesitancy	93	0.49	0.23 - 0.74	100	1.00	1.00 - 1.00
Action required to start micturition	94	0.63	0.42 - 0.85	100	1.00	1.00 - 1.00
Position change	100	1.00	1.00 - 1.00	100	/	/
Straining	86	0.70	0.17 - 1.00	100	1.00	1.00 - 1.00
Reduction of coeles	86	/	/	100	/	/
Other	86	0.70	0.17 - 1.00	100	1.00	1.00 - 1.00
Straining to void	88	0.68	0.55 - 0.82	100	1.00	1.00 - 1.00
Dysuria	94	0.58	0.35 - 0.81	100	1.00	1.00 - 1.00
Urethra	100	/	/	100	1.00	1.00 - 1.00
Lower Abdomen	100	1.00	1.00 - 1.00	100	1.00	1.00 - 1.00
Between legs	**75**	**0.38**	(-0.2) - 0.97	100	/	/
Other	75	/	/	100	/	/
Intermittency	86	0.56	0.39 - 0.73	100	1.00	1.00 - 1.00
Slow stream	86	0.63	0.49 - 0.78	100	1.00	1.00 - 1.00
Change in flow	88	0.75	0.64 - 0.86	100	1.00	1.00 - 1.00
Need to immediately re-void	86	0.69	0.57 - 0.81	100	1.00	1.00 - 1.00
Feeling of incomplete bladder emptying	91	0.74	0.61 - 0.88	100	1.00	1.00 - 1.00
Postmicturition leakage	78	0.57	0.44 - 0.70	100	1.00	1.00 - 1.00

κ = Mean Kappa ; 95% CI = 95% Confidence Interval ; **Bold** = κ values < 0.40 and their respective agreement.

LUTS questions are shown in tables 3 and 4, respectively.

The overall median κ for the anorectal symptom questions was 0.63 (IQR : 0.55-0.71), with a high mean agreement of 87%. The mean agreement was comparable in all three patient groups (82%-91%). The overall median κ was 0.57 (IQR : 0.45-0.67) in the CONTROL patients, 0.60 (IQR : 0.45-0.69) in the GASTRO and 0.72 (IQR : 0.61-0.88) in the URO patients. The ICC of defecation frequency was 0.93 (95% CI : 0.89-0.96).

The overall median κ for the LUTS questions was 0.63 (IQR : 0.53-0.70), with a high overall agreement of 87%. The mean agreement was equal in all three patient groups (86%). The overall median κ was 0.53 (IQR : 0.32-0.62) in the CONTROL patients, 0.64 (IQR : 0.51-0.72) in the GASTRO and 0.60 (IQR : 0.53-0.73) in the URO patients. The ICC of daytime voiding frequency was 0.84 (95% CI : 0.78-0.88).

The overall median κ for the questions about obstetric history was 0.63 (IQR : 0.58-0.91) and for the questions about medical history 0.84 (IQR : 0.81-0.85). The mean agreement of both was high : 92% and 95% respectively. The κ-value for the QoL question was 0.65 (95% CI : 0.45-0.85) and the agreement 78%.

Internal consistency

Grouping all the anorectal symptom questions gave a high Crohnbach's α of 0.78. Similarly, the LUTS questions had a Crohnbach's α of 0.80. There was no clear psychometric benefit from subgrouping the questions into categories of storage and evacuation for both anorectal symptoms and LUTS : the Crohnbach's α for these groups were 0.52 and 0.83 for anorectal symptom questions and 0.73 and 0.77 for LUTS questions. With the exception of anorectal storage symptoms, these values remained high.

Discussion

The prevalence and impact of combined fecal and urinary disorders in the general population is high. More knowledge about the whole spectrum of interrelated bowel and bladder symptoms could benefit diagnosis and treatment in patients presenting with lower urinary tract and/or lower bowel symptoms. We report the results of a validation study of a comprehensive questionnaire designed to assess the full spectrum of anorectal and urinary symptoms together. The findings support the validity and reliability of the questionnaire by demonstrating good psychometric properties.

The questionnaire was designed to be self-completed and was readily understood and easily completed by the target population. Although it appears lengthy, most patients completed the questionnaire within 10-15 minutes. This relatively short time requirement may be explained by the structuring in main and sub-questions (which only need answering if a certain symptom is present), and, by the consistency in format between similar questions on bowel and bladder. The low levels of missing data indicate the absence of inappropriate questions. Experts from the pelvic clinic reviewing the questionnaire concluded that our questionnaire covers all important domains, while still being useful in clinical practice due to its simplicity and limited time requirement.

The questionnaire shows good construct validity as it is clearly able to differentiate between the clinical and CONTROL populations : anorectal symptoms and LUTS were reported more frequently in the GASTRO and URO populations, respectively, compared to the CONTROL population. Furthermore, the UI question in our questionnaire allows clear differentiation between males and females.

Criterion validity describes how well the questionnaire correlates with a "gold standard" measure that already exists. As no combined anorectal and LUT questionnaires exist in Dutch, we compared the self-reported patients' answers to their interview-assisted responses. However, it is important to note that the patient's experience of disease is similar to the physician's rating of disease for FI (23) but not for LUTS (24). Patients may also choose to give the "desirable" answer to the physician, rather than the real answer. We tried to counter this by stressing the importance of truthful responses for the future diagnostic approach and therapeutic success.

The agreement and the κ statistic for criterion validity were very high for all the questions. This may be explained by the fact that an earlier version of the questionnaire, containing only LUTS questions, had been used for years in the Urology Department. This questionnaire was used as a method to facilitate history taking and was adapted regularly, based on patient and doctor feedback. This early version of the questionnaire was never tested for validity and reliability.

A limitation of the study design and analysis is the lack of an objective clinical assessment with a gold standard measure (e.g. anorectal manometry, transit times, videoproctography or urodynamic studies). The availability of such data would permit stronger conclusions regarding the criterion validity of the questionnaire. However, questionnaires are primarily designed to measure the patient's perspective of their condition, and hence the diagnosis of a condition may be less important than the way in which it is perceived by the patient.

Test-retest reliability of the questionnaire is satisfactory, with high values of observed agreement, of the kappa (κ) statistic and of the 95% confidence interval (CI). The results for our questionnaire are comparable to previously validated instruments (33,34). Several points need attention in evaluating the stability of a questionnaire : the time interval and the mode of administration.

If the interval between test and retest is too short, the initial response may be remembered upon retesting ; if too long, the clinical condition may have altered. We tried to counter this issue by choosing a time interval of four weeks and by adding questions to the second questionnaire regarding a possible change in the patient's condition and regarding a possible initiation of treatment. A positive answer to either of these questions would have resulted in an exclusion from our study, but did not apply to any included patient.

True retesting also requires the mode of administration to be identical on the two occasions. Although the questionnaire was self-completed on both occasions, we could not avoid the questionnaire being completed in two different locations. Nevertheless, we found good agreement between answers to the questionnaires completed in the waiting room and those completed at home.

In four of the anorectal symptom questions and in three of the urological questions, the κ statistic for reproducibility was < 0.4 (Tables 3 and 4). In six of these seven questions, the concurrent agreement was $\geq 75\%$, suggesting that the low κ was attributable to a high prevalence of either a positive or a negative response and not to poor agreement (31). As suggested by Cichetti and Feinstein (30) we evaluated the positive and negative agreement (Ppos and Pneg) for the questions which had a high agreement but low κ. It was clear that the lower κ for these questions could be attributed to the

reproducibility of the less prevalent answers. One question with mainly positive answers ("The need to run when postponing defecation when desire was felt") showed a Ppos of 0.9. The other questions with high agreement and low κ had mainly negative answers and showed a Pneg of 0.91.

One item in the questionnaire had both low agreement and a low κ : "whether the stool loss in faecal incontinence was mainly solid, liquid or both". A possible explanation for the low reproducibility of this question is that stool consistency may vary, even over a shorter period of time, depending on the patient's diet and condition. The question left a third possible answer which combined both previous answer possibilities. Almost half of the patients (11 out of 24) choose this third answer on at least one occasion. This question should be interpreted taking this into account.

The good overall internal consistency was not improved by dividing the symptoms into arbitrary categories of storage and evacuation. On the other hand, the Crohnbach's α was not altered drastically for the anorectal evacuation symptoms or for the storage and evacuation LUTS, despite having fewer items in each category. This indicates a high correlation among items of these three categories. Internal consistency correlates highly with the number of items in the sub-group (35). Therefore, the low number of anorectal storage symptom questions may explain the low Crohnbach's α for this subgroup.

One psychometric property was not assessed in this study : responsiveness of the questionnaire to change. However, our questionnaire is not a condition-specific outcome measure and is therefore not designed to reliably detect the overall effect of treatment and to detect clinically meaningful change. Therefore, we choose not to test its responsiveness.

The questionnaire assessed in this paper serves the role of a symptom inventory and does not assess the extent to which the symptoms impair the patients' overall QoL. Such questionnaires are available for LUTS (19) and for fecal incontinence (36). We accept that the lack of a measure of the degree of bother or of a more extensive measure of the impact on QoL, may represent a shortcoming. However, the inclusion of a specific QoL assessment for each symptom, or group of symptoms, would considerably lengthen the questionnaire and reduce its user friendliness. The application of separate QoL questionnaires specific for anorectal symptoms and LUTS may be useful in conjunction with the current questionnaire for clinical research.

This questionnaire combines an extensive list of both anorectal and urological questions. We decided to use it as a means to evaluate the concomitance of habits, complaints and symptoms in patients with both bowel and bladder disorders, as the current knowledge about this remains limited. Further studies with this instrument intend to identify discriminating questions that can help in diagnostic and therapeutic planning for patients with concomitant bowel and bladder disorders. This may permit to select a shorter list of questions that can be used in everyday clinical practice. This questionnaire was evaluated for reproducibility and validity and as the results were satisfactory, the questionnaire can now be used for further surveys.

Conclusions

We designed and psychometrically tested a new questionnaire in Dutch which evaluates habits, complaints and symptoms in patients with bowel and bladder disorders. This instrument is considered to be a valid and reliable instrument for the assessment of anorectal and lower urinary tract symptoms. This questionnaire can provide further insights into the epidemiology of concomitant bowel and bladder disorders and, accordingly, can contribute to a more efficient diagnostic and therapeutic approach in patients with such disorders.

Acknowledgements

We want to thank Prof J. Somville (Orthopaedic department) and Prof G. Moorkens (Department General Internal Medicine) for their help in composing the CONTROL group with patients contacting their departments.

Funding

Onderzoeksmandaat UA : Research fellowship University Antwerp, Faculty of Medicine.

FWO : Fonds Wetenschappelijk Onderzoek Vlaanderen/ Research Foundation – Flanders.

Research Fund Urology, Faculty Medicine, University Antwerp.

References

1. ROBERTS R.O., JACOBSEN S.J., REILLY W.T., PEMBERTON J.H., LIEBER M.M., TALLEY N.J. Prevalence of combined fecal and urinary incontinence : a community-based study. *J. Am. Geriatr. Soc.*, 1999, **47** (7) : 837-841.
2. UUSTAL F.E., WINGREN G., KJOLHEDE P. Factors associated with pelvic floor dysfunction with emphasis on urinary and fecal incontinence and genital prolapse : an epidemiological study. *Acta Obstet. Gynecol. Scand.*, 2004, **83** (4) : 383-389.
3. KHULLAR V., DAMIANO R., TOOZS-HOBSON P., CARDOZO L. Prevalence of faecal incontinence among women with urinary incontinence. *Br. J. Obstet. Gynaecol.*, 1998, **105** (11) : 1211-1213.
4. GORDON D., GROUTZ A., GOLDMAN G. *et al.* Anal incontinence : prevalence among female patients attending a urogynecologic clinic. *Neurourol. Urodyn.*, 1999, **18** (3) : 199-204.
5. LEROI A.M., WEBER J., MENARD J.F., TOUCHAIS J.Y., DENIS P. Prevalence of anal incontinence in 409 patients investigated for stress urinary incontinence. *Neurourol. Urodyn.*, 1999, **18** (6) : 579-590.
6. MANNING J., EYERS A.A., KORDA A., BENNESS C., SOLOMON M.J. Is there an association between fecal incontinence and lower urinary dysfunction ? *Dis. Colon Rectum*, 2001, **44** (6) : 790-798.
7. NG S.C., CHEN Y.C., LIN L.Y., CHEN G.D. Anorectal dysfunction in women with urinary incontinence or lower urinary tract symptoms. *Int. J. Gynaecol. Obstet.*, 2002, **77** (2) : 139-145.

8. LACIMA G., ESPUNA M., PERA M., PUIG-CLOTA M., QUINTO L., GARCIA-VALDECASAS J.C. Clinical, urodynamic, and manometric findings in women with combined fecal and urinary incontinence. *Neurourol. Urodyn.*, 2002, **21** (5) : 464-469.
9. SOLIGO M., SALVATORE S., MILANI R. *et al.* Double incontinence in urogynecologic practice : a new insight. *Am. J. Obstet. Gynecol.*, 2003, **189** (2) : 438-443.
10. WYNDAELE M., DE WINTER B.Y., PELCKMANS P., WYNDAELE J.J. Lower bowel function in urinary incontinent women, urinary continent women and in controls. *Neurourol. Urodyn.*, 2010.
11. MONGA A.K., MARRERO J.M., STANTON S.L., LEMIEUX M.C., MAXWELL J.D. Is there an irritable bladder in the irritable bowel syndrome ? *Br. J. Obstet. Gynaecol.*, 1997, **104** (12) : 1409-1412.
12. BHARUCHA A.E., ZINSMEISTER A.R., LOCKE G.R. *et al.* Prevalence and burden of fecal incontinence : a population-based study in women. *Gastroenterology*, 2005, **129** (1) : 42-49.
13. FIALKOW M.F., MELVILLE J.L., LENTZ G.M., MILLER E.A., MILLER J., FENNER D.E. The functional and psychosocial impact of fecal incontinence on women with urinary incontinence. *Am. J. Obstet. Gynecol.*, 2003, **189** (1) : 127-129.
14. MARKLAND A.D., RICHTER H.E., KENTON K.S. *et al.* Associated factors and the impact of fecal incontinence in women with urge urinary incontinence : from the Urinary Incontinence Treatment Network's Behavior Enhances Drug Reduction of Incontinence study. *Am. J. Obstet. Gynecol.*, 2009, **200** (4) : 424-428.
15. LOENING-BAUCKE V. Urinary incontinence and urinary tract infection and their resolution with treatment of chronic constipation of childhood. *Pediatrics*, 1997, **100** (2 Pt 1) : 228-232.
16. WYNDAELE J.J., GOLDFISCHER E.R., MORROW J.D. *et al.* Effects of flexible-dose fesoterodine on overactive bladder symptoms and treatment satisfaction : an open-label study. *Int J Clin Pract*, 2009, **63** (4) : 560-567.
17. WALL L.L., DELANCEY J.O. The politics of prolapse : a revisionist approach to disorders of the pelvic floor in women. *Perspect. Biol. Med.*, 1991, **34** (4) : 486-496.
18. BARBER M.D., KUCHIBHATLA M.N., PIEPER C.F., BUMP R.C. Psychometric evaluation of 2 comprehensive condition-specific quality of life instruments for women with pelvic floor disorders. *Am J Obstet Gynecol*, 2001, **185** (6) : 1388-1395.
19. AVERY K., DONOVAN J., PETERS T.J., SHAW C., GOTOH M., ABRAMS P. ICIQ : a brief and robust measure for evaluating the symptoms and impact of urinary incontinence. *Neurourol Urodyn*, 2004, **23** (4) : 322-330.
20. LONGSTRETH G.F., THOMPSON W.G., CHEY W.D., HOUGHTON L.A., MEARIN F., SPILLER R.C. Functional bowel disorders. *Gastroenterology*, 2006, **130** (5) : 1480-1491.
21. BHARUCHA A.E., WALD A., ENCK P., RAO S. Functional anorectal disorders. *Gastroenterology*, 2006, **130** (5) : 1510-1518.
22. DONOVAN J., BOSCH R., GOTOH M. *et al.* Symptom and Quality of Life Assessment. *In* : ABRAMS P., CARDOZO L., KHOURY S., WEIN A. (eds). 3rd ed. Paris : Editions 21, 2009. 519-584.
23. ROCKWOOD T.H., CHURCH J.M., FLESHMAN J.W. *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** (12) : 1525-1532.
24. KHAN M.S., CHALIHA C., LESKOVA L., KHULLAR V. The relationship between urinary symptom questionnaires and urodynamic diagnoses : an analysis of two methods of questionnaire administration. *BJOG*, 2004, **111** (5) : 468-474.
25. RUST J., GOLOMBOK S. *Modern psychometrics : The science of psychological assessment.* 2nd ed. London : Routledge, 1999.
26. HERZOG A.R., FULTZ N.H. Prevalence and incidence of urinary incontinence in community-dwelling populations. *J. Am. Geriatr. Soc.*, 1990, **38** (3) : 273-281.
27. PRETLOVE S.J., RADLEY S., TOOZS-HOBSON P.M., THOMPSON P.J., COOMARASAMY A., KHAN K.S. Prevalence of anal incontinence according to age and gender : a systematic review and meta-regression analysis. *Int. Urogynecol. J. Pelvic Floor Dysfunct.*, 2006, **17** (4) : 407-417.
28. SZKLO M., NIETO F.J. *Epidemiology. Beyond the basics.* Sudbury : Jones and Bartlett Publishers, 2004.
29. LANDIS J.R., KOCH G.G. The measurement of observer agreement for categorical data. *Biometrics*, 1977, **33** (1) : 159-174.
30. CICCHETTI D.V., FEINSTEIN A.R. High agreement but low kappa : II. Resolving the paradoxes. *J. Clin. Epidemiol.*, 1990, **43** (6) : 551-558.
31. FEINSTEIN A.R., CICCHETTI D.V. High agreement but low kappa : I. The problems of two paradoxes. *J. Clin. Epidemiol.*, 1990, **43** (6) : 543-549.
32. FLEISS J.L. *Statistical Methods for Rates and Proportions.* New York : Wiley, 1981.
33. BHARUCHA A.E., LOCKE G.R. III, SEIDE B.M., ZINSMEISTER A.R. A new questionnaire for constipation and faecal incontinence. *Aliment. Pharmacol. Ther.*, 2004, **20** (3) : 355-364.
34. ADELSTEIN B.A., IRWIG L., MACASKILL P., KATELARIS P.H., JONES D.B., BOKEY L. A self administered reliable questionnaire to assess lower bowel symptoms. *BMC Gastroenterol.*, 2008, **8** : 8.
35. STREINER D.L., NORMAN G.R. *Health measurement scales : a practical guide to their development and use.* Oxford : Oxford University Press, 1989.
36. REILLY W.T., TALLEY N.J., PEMBERTON J.H., ZINSMEISTER A.R. Validation of a questionnaire to assess fecal incontinence and associated risk factors : Fecal Incontinence Questionnaire. *Dis. Colon Rectum*, 2000, **43** (2) : 146-153.