

ICIQ: A Brief and Robust Measure for Evaluating the Symptoms and Impact of Urinary Incontinence

Kerry Avery,^{1*} Jenny Donovan,² Tim J. Peters,³ Christine Shaw,⁴
Momokazu Gotoh,⁵ and Paul Abrams¹

¹*Bristol Urological Institute, Southmead Hospital, Bristol, United Kingdom*

²*Department of Social Medicine, University of Bristol, United Kingdom*

³*Division of Primary Health Care, University of Bristol, United Kingdom*

⁴*Department of General Practice, University of Wales College of Medicine, United Kingdom*

⁵*Department of Urology, Nagoya University School of Medicine, Japan*

Aims: To develop and evaluate the International Consultation on Incontinence Questionnaire (ICIQ), a new questionnaire to assess urinary incontinence and its impact on quality of life (QoL).

Methods: A developmental version of the questionnaire was produced following systematic literature review and views of an expert committee and patients. Several studies were undertaken to evaluate the psychometric properties of the questionnaire, including content, construct and convergent validity, reliability and sensitivity to change. **Results:** The ICIQ was easily completed, with low levels of missing data (mean 1.6%). It was able to discriminate among different groups of individuals, indicating good construct validity. Convergent validity was acceptable, with most items demonstrating 'moderate' to 'strong' agreement with other questionnaires. Reliability was good, with 'moderate' to 'very good' stability in test-retest analysis and a Cronbach's alpha of 0.95. Items identified statistically significant reductions in symptoms from baseline following surgical and conservative treatment. Item reduction techniques were used to determine the final version and scoring scheme, which also demonstrated good psychometric properties. **Conclusions:** The final ICIQ comprises three scored items and an unscored self-diagnostic item. It allows the assessment of the prevalence, frequency, and perceived cause of urinary incontinence, and its impact on everyday life. The ICIQ is a brief and robust questionnaire that will be of use in outcomes and epidemiological research as well as routine clinical practice. *NeuroUrol. Urodynam.* 23:322–330, 2004. © 2004 Wiley-Liss, Inc.

Key words: outcome measure; quality of life; questionnaire; reliability; sensitivity to change; urinary incontinence; validity

INTRODUCTION

Urinary incontinence is a common and distressing condition [Donovan et al., 2002]. There has been increasing recognition that symptoms alone are poor indicators of the effect of incontinence on individuals' lives. Although not life-threatening, the symptoms of incontinence can be severely incapacitating, causing considerable impairment to various aspects of an individual's life and ultimately reducing the quality of their life (QoL). But this impact varies between individuals. Consequently, the measurement of incontinence should incorporate the impact the condition may have on QoL.

Several self-completion questionnaires have been developed to assess incontinence, but most are lengthy and have been developed for use in specific patient groups. As yet, there is no brief and simple questionnaire that would allow the assessment of the symptoms and impact of incontinence across the population that could be applied widely in clinical practice and research. In 1998, the WHO-sponsored Interna-

tional Consultation on Incontinence (ICI) initiated the development and evaluation of such a questionnaire. This study reports on the developmental and psychometric research undertaken to produce the International Consultation on Incontinence Questionnaire (ICIQ).

Abbreviations: ICIQ, International Consultation on Incontinence Questionnaire; QoL, quality of life; LUTS, lower urinary tract symptoms.

K. Avery, J. Donovan and P. Abrams designed the study. K. Avery collected the data, carried out the analyses and produced the first draft of the manuscript. T. Peters advised on statistical methods. All the authors were involved in writing the manuscript.

Ethical approval for this study was granted by the Local Research Ethics Committee at Southmead Hospital in Bristol, UK.

*Correspondence to: Kerry Avery, BSc, PhD, Bristol Urological Institute, Southmead Hospital, Westbury-on-Trym, Bristol, BS10 5NB, UK.

E-mail: kerryavery1977@hotmail.com

Received 3 December 2003; Accepted 25 March 2004

Published online 24 May 2004

in Wiley InterScience (www.interscience.wiley.com)

DOI 10.1002/nau.20041

MATERIALS AND METHODS

Validity

Detailed psychometric evaluation is necessary to determine how well a questionnaire measures the concept it aims to measure and to ensure that it is reliable and suitably responsive to symptom or QoL change [Donovan et al., 2002]. A number of studies were undertaken to develop the ICIQ and to examine its psychometric properties. Various sampling methods were employed to develop and evaluate the questionnaire in individuals who represented potential respondents, comprising samples of clinic and community-based adults of a range of ages and both sexes with differing levels of symptoms. Table I outlines the main study samples and characteristics; Table II the sub-studies and samples.

Developmental Work

To ensure that the questionnaire reflected the content domain of the incontinence, could be clearly and easily interpreted by respondents and make sense to those in the clinical area [Nunnally and Bernstein, 1994]; consultations were carried out with the ICI expert committee and in-depth interviews with 63 UK urology clinic attendees (46 females, 17 males) with incontinence or other lower urinary tract symptoms (LUTS). Patients were observed completing the questionnaire and interviewed to establish their comprehension of items. A developmental version of the questionnaire (*d*ICIQ) was then produced. The following psychometric properties of the *d*ICIQ were then assessed.

- Content validity: Response rates and missing data indicate the acceptability of items [Donovan et al., 2002]. Consecutive UK urology clinic attendees with incontinence or other LUTS and patients registered with two UK community general practices were invited to complete the questionnaire. Response rates and percentage levels of missing data were calculated.
- Construct validity: This refers to the relationships between the questionnaire and underlying theories [Rust and Golombok, 1999], for example, that the prevalence of incontinence is higher among females than males at all ages [Herzog and Fultz, 1990]. The ability of the *d*ICIQ to discriminate between individuals of different genders, ages and patient groups and with different types of incontinence was assessed by comparing levels of incontinence in samples of urology clinic attendees and community-based individuals of different ages and sexes, using chi square (χ^2) analyses for unpaired categorical data [Litwin, 1995]. Incontinence was defined as a minimum leakage of 'about once a week or less often' by the *d*ICIQ item assessing 'amount of leakage'.
- Convergent validity: As there is no 'gold standard' questionnaire for incontinence, the relationships between items in the *d*ICIQ, ICSmale short form (ICSmaleSF) [Donovan et al., 2000] and Bristol Female Lower Urinary Tract Symptoms (BFLUTS) [Jackson et al., 1996] questionnaires, which assess related concepts, were investigated

TABLE I. Samples and Socio-Demographic Characteristics of Studies Used in the Psychometric Testing of the ICIQ

	Invited	Participated
Total baseline sample	Total = 634: females = 455; males = 179	Total = 469: females = 324; males = 145 Mean age = 57.2 years, range, 23.4–101.3
Bristol clinic sample	Total = 374: females = 305; males = 69	Total = 223: females = 182; males = 41 Mean age = 56.4 years, range, 23.4–90.9 Response rate = 60%
Leicester community sample	Total = 230: females = 129; males = 101	Total = 221: females = 126; males = 95 Mean age = 58.7 years, range, 40.8–101.3 Response rate = 96%
Bristol community sample	Total = 30: females = 21; males = 9	Total = 25: females = 16; males = 9 Mean age = 50.1 years, range, 24.5–73.3 Response rate = 83%
Total community sample	Total = 260: females = 150; males = 10	Total = 246: females = 142; males = 104 Mean age = 57.8 years, range, 24.5–101.3
Leicester community sample		
Bristol community sample		
Total surgical intervention sample	Total = 75: females = 63; males = 12	Total = 57: females = 47; males = 10 Mean age = 56.5 years, range, 21.6–88.2 Response rate = 76%
Surgical intervention sample— <i>d</i> ICIQ	Total = 49: females = 40; males = 9	
Surgical intervention sample—ICIQ	Total = 26: females = 23; males = 3	
Conservative management sample	Total = 266: females = 11; males = 255	Total = 206: females = 7; males = 199 Mean age = 66.6 years, range, 41.3–87.4 Response rate = 77%
Second Bristol clinic sample	Total = 165: females = 154; males = 11	Total = 105: females = 98; males = 7 Mean age = 55.3 years, range, 20.1–88.5 Response rate = 64%

TABLE II. Sub-Samples Used for Particular Studies in the Psychometric Testing of the ICIQ

	Invited	Participated
Convergent validity—BFLUTS Sub-sample of Bristol clinic sample	Females = 258	Females = 118 Mean age = 57.7 years; range, 24.4–88.3 Response rate = 46%
Convergent validity—ICSmaleSF Sub-sample of Bristol clinic sample	Males = 58	Males = 27 Mean age = 58.6 years; range, 23.6–82.6 Response rate = 47%
Convergent validity—KHQ Sub-sample of Bristol clinic sample	Females = 144	Females = 91 Mean age = 55.2 years; range, 20.1–88.5 Response rate = 63%
Stability (test-retest reliability) Sub-sample of Bristol clinic sample	Total = 223: females = 182; males = 41	Total = 144: females = 121; males = 23 Mean age = (tm)58.1 years; range, 24.5–90.9 Response rate = 65%

using Spearman's rank correlation coefficient for ordered categorical data [Altman, 1991].

Reliability

A reliable questionnaire is consistent, stable and reproducible [Litwin, 1995].

- (d) **Stability:** The stability of individuals' responses to the questionnaire items over a period in which their symptom status would not be expected to change was assessed in patients from the clinic sample who completed a second *d*ICIQ within approximately 2 weeks of the first questionnaire. Agreement between test and retest responses were analysed by graphical interpretation and the weighted Kappa statistic for ordered categorical data [Altman, 1991], rather than a correlation coefficient that would not adequately represent levels of agreement [Bland and Altman, 1986; Altman, 1991].
- (e) **Internal consistency:** The correlation between the questionnaire items was assessed by Cronbach's alpha coefficient [Rust and Golombok, 1999] using baseline data from the clinic and community samples.

Sensitivity

- (f) **Sensitivity to detect change in symptoms or QoL** [Armitage and Berry, 1994] was investigated in patients in the intervention arm of a trial comparing continence nurse practitioner (CNP) with standard care, and patients undergoing surgery for incontinence. CNP care was based on primary evidence-based clinical interventions including fluid intake awareness, pelvic floor awareness, bladder re-education, antibiotics for urinary tract infection, treatment of *Candida* (thrush), the need for medication change, consideration of iatrogenic causes and advice on incomplete bladder emptying. Surgical intervention included insertion of artificial urinary sphincter, anterior vaginal repair, injection of bulking agents (including col-

lagen and synthetic polymers), bladder neck suspension, augmentation cystoplasty and sling procedures (including suburethral sling and tension-free vaginal or transvaginal tape/TVT).

The percentage change in the presence of symptoms between baseline and follow-up (approximately 8 weeks for conservative management and 12 weeks for surgical intervention) was calculated. The Wilcoxon signed ranks test for paired ordered categorical data [Altman, 1991] was used to determine whether symptom levels differed significantly. The difference in the degree of change between patients receiving conservative management and surgical intervention was investigated using univariable regression with a single binary explanatory variable [Altman, 1991].

Self-Diagnostic Item

Clinicians on the expert panel indicated the need for patients completing this short questionnaire to be able to indicate the perceived cause of the incontinence. This was kept separate from the scored part of the questionnaire, as it was not concerned with severity or impact. The question was subjected to testing of content, construct and convergent validity and sensitivity.

Devising the Final ICIQ and a Scoring Scheme

A 'principal factor' analysis was undertaken to ascertain the most suitable factor solution for the underlying structure of the questionnaire [Armitage and Berry, 1994] using all available baseline data. Redundant items were removed following consideration of the results from (a)–(f) above. The psychometric properties of the score were investigated and descriptive statistics used to determine how well it described the data [Altman, 1991]. The ability of the score to reflect theories relating to incontinence in specific patient groups was investigated using univariable and linear regression methods [Altman, 1991]. Agreement between the ICIQ score and the

ICSmaleSF 'Incontinence' and 'Voiding' subscores [Donovan et al., 2000] and the BFLUTS 'Incontinence' and 'Quality of Life' factors [Brookes et al., in press] was investigated by Pearson's product moment correlation coefficient (r) for interval data [Altman, 1991]. The relationship between the ICIQ score and scores on the King's Health Questionnaire [Kelleher et al., 1997] was similarly analysed using data from consecutive female urology clinic attendees with incontinence or LUTS who simultaneously completed the KHQ. The ICIQ score's stability was assessed by calculating the percentage agreement between the scores obtained at test and retest and the weighted Kappa statistic [Altman, 1991]. Data from intervention studies were used to investigate the score's sensitivity.

RESULTS

Validity

- (a) Content validity: Interviews and review by clinical and social science experts indicated that items in the *d*ICIQ were well interpreted and covered all important domains. Response rates were good (see Tables I and II). Eighty-seven percent of individuals fully completed all items, with most items demonstrating very low levels of missing data (mean 1.6%, range, <1% to 2%).
- (b) Construct validity: The *d*ICIQ clearly differentiated between males and females, with community-based women reporting more incontinence than men (58.9% and 25.2% respectively, $P < 0.001$). It also detected a lower prevalence of incontinence in the community sample than in urology clinic attendees (44.7% and 97.2% respectively, $P < 0.001$) and demonstrated a clear association between sex and the perceived causes of incontinence in both the clinic and the community samples ($P < 0.001$). As anticipated, stress incontinence was the most predominant in community women, in contrast to men where urge incontinence was the most commonly reported. Whilst urge and stress incontinence were commonly observed among male and female clinic attendees, respec-

tively, this was exceeded by the presence of mixed incontinence in both sexes.

- (c) Convergent validity: Agreement between responses to *d*ICIQ and BFLUTS items measuring the 'frequency' and 'usual amount' of leakage ranged from 'moderate' to 'strong' (Table III). Agreement between responses to *d*ICIQ and BFLUTS/ICSmaleSF items assessing the perceived causes of incontinence ranged from 'weak' to 'moderate'.

Reliability

- (d) Stability: The reliability of the symptom items is presented in Table IV. Agreement was 'good' to 'very good' for all items excluding 'overall quality of life', which was 'moderate'.
- (e) Internal consistency: Cronbach's alpha coefficient was very high (0.95) indicating excellent internal consistency but also some redundancy. Item reduction was therefore undertaken. Factor analyses confirmed a single strong underlying factor. Redundant and outlying items were removed based on clinical judgement and psychometric methods. A number of items were closely related to each other, with correlations greater than 0.83. Those demonstrating the weakest psychometric properties, the lowest factor loadings and lowest intercorrelations with the other items were removed. The final questionnaire (Fig. 1) comprises three scored items (Questions 1–3) and a self-diagnostic item (Question 4, not scored), with a Cronbach's alpha coefficient of 0.92.

Sensitivity

- (f) There was an observed decrease in the percentage of patients reporting symptoms on each symptom item following conservative management, with all except 'overall quality of life' reaching statistical significance (Table V). There was also a decrease in the percentage reporting symptoms on all except one item relating to the perceived

TABLE III. Agreement Between Responses to Items in the *d*ICIQ and the BFLUTS and ICSmaleSF Questionnaires

Item	BFLUTS			ICSmaleSF		
	Spearman's r_s	P -value	95% confidence interval*	Spearman's r_s	P -value	95% confidence interval*
Frequency of leakage	0.86	<0.001	¹ 3.3–3.6, ² 3.2–3.5	—	—	—
Usual amount of leakage	0.53	<0.001	¹ 1.5–1.8, ² 1.3–1.6	—	—	—
Perceived cause of leakage						
Before reaching the toilet	0.35	0.0001	¹ 0.6–0.8, ² 0.9–1.0	0.24	0.23	¹ 0.4–0.8, ² 0.9–1.0
When coughing or sneezing	0.44	<0.001	¹ 0.7–0.9, ² 0.9–1.0	0.58	0.002	¹ 0.2–0.6, ² 0.5–0.9
When asleep	0.47	<0.001	¹ 0.1–0.3, ² 0.5–0.6	0.50	0.008	¹ 0.1–0.5, ² 0.4–0.8
When active or exercising	0.29	0.002	¹ 0.5–0.7, ² 0.9–1.0	—	—	—
After urinating and dressed	—	—	—	0.45	0.023	¹ 0.3–0.7, ² 0.7–1.0
No obvious reason	0.55	<0.001	¹ 0.4–0.6, ² 0.7–0.9	0.24	0.23	¹ 0.4–0.8, ² 0.6–1.0

*¹*d*ICIQ, ²BFLUTS/ICSmale.

TABLE IV. Agreement Between Test and Retest Responses for Nine dICIQ Symptom Items

Item		Crude agreement (%)	Kappa value	P-value	95% confidence interval*
Q3a	Frequency of leakage	92.0	0.73	<0.001	¹ 3.1–3.5, ² 3.0–3.5
Q3b	Frequency of leakage—bother	88.7	0.68	<0.001	¹ 6.1–7.1, ² 5.6–6.7
Q5a	Frequency of protection use	95.7	0.90	<0.001	¹ 1.6–2.0, ² 1.7–2.1
Q6a	Usual amount of leakage	92.3	0.71	<0.001	¹ 1.4–1.7, ² 1.4–1.6
Q6b	Worst amount of leakage	89.5	0.67	<0.001	¹ 1.7–2.0, ² 1.7–2.0
Q7	Interference with everyday life	90.2	0.74	<0.001	¹ 5.3–6.4, ² 5.2–6.2
Q8	Interference with social life	87.6	0.70	<0.001	¹ 4.4–5.6, ² 4.4–5.5
Q9	Interference with sex life	88.9	0.75	<0.001	¹ 3.0–4.7, ² 2.8–4.5
Q10	Overall quality of life	85.3	0.58	<0.001	¹ 5.0–6.1, ² 5.0–6.0

*¹test, ²retest.

causes of incontinence, with all but two reaching statistical significance. Males experienced improvement in more of their symptoms overall (ranging from approximately –2% to –15%) than females (approximately 0% to –29%). Decreases in the percentage of patients reporting symp-

toms on all symptom items were observed following surgical intervention (Table VI), with each highly statistically significantly better at follow-up. Decreases were also observed in the percentage reporting each perceived cause of incontinence, although only three of seven reached statistical significance. Again, males appeared to experience decreases in more of their symptoms overall (ranging from approximately 0% to –86%) than females (approximately 0% to –52%). Much larger improvements were detected for all symptoms in those patients who received surgical intervention (ranging from approximately 39% to 50%) in comparison with those who received conservative management (2% to 15%).

Devising the Scoring Scheme

As the factor analyses and internal consistency indicated a robust measure, it was possible to combine items into a single summed score (range, 0–21). Scores could be calculated for 458 (98%) of the 469 questionnaires returned (mean 7.2, standard deviation 6.6, median 5, inter-quartile range, 0–13; observed range, 0–21). As anticipated, patients attending urology clinics reported significantly higher mean scores than individuals in the community (mean 12.7 and 2.4 respectively: $P = 0.001$, 95% confidence interval 9.6–11.1), with community-dwelling females reporting greater levels of symptoms and impact of incontinence than males (mean 2.9 and 1.6 respectively: $P = 0.001$, 95% confidence interval –2.2 to –0.5) and an increase in the scores obtained by community-based individuals with increasing age ($P = 0.018$, 95% confidence interval 0.06–0.6 per decade). A statistically significant difference in scores was also observed between individuals with different types of incontinence in the clinic ($P = 0.001$) and the community sample ($P = 0.001$), with higher scores among individuals with urge incontinence in comparison to those with stress incontinence.

The ICIQ score correlated well with the ICSmaleSF ‘Incontinence’ subscore but demonstrated only a weak correlation

Many people leak urine some of the time. We are trying to find out how many people leak urine, and how much this bothers them. We would be grateful if you could answer the following questions, thinking about how you have been, on average, over the PAST FOUR WEEKS.

1. How often do you leak urine? (Tick one box)

never 0

about once a week or less often 1

two or three times a week 2

about once a day 3

several times a day 4

all the time 5

We would like to know how much you think leaks.

2. How much urine do you usually leak (whether you wear protection or not)? (Tick one box)

none 0

a small amount 2

a moderate amount 4

a large amount 6

3. Overall, how much does leaking urine interfere with your everyday life?

Please ring a number between 0 (not at all) and 10 (a great deal)

0 1 2 3 4 5 6 7 8 9 10

not at all a great deal

ICIQ score: sum scores 1+2+3

4. When does urine leak? (Please tick all that apply to you)

never – urine does not leak

leaks before you can get to the toilet

leaks when you cough or sneeze

leaks when you are asleep

leaks when you are physically active/exercising

leaks when you have finished urinating and are dressed

leaks for no obvious reason

leaks all the time

Fig. 1. Items in the ICIQ.

TABLE V. Change (%) in Respondents Reporting 'Any Level of Symptom' Before and After Conservative Management

Symptom/item	Percentage change (%)			P-value		
	Males and females	Males	Females	Males and females	Males	Females
Frequency of leakage	-10.9	-11.3	0.0	<0.001	<0.001	0.16
Frequency of leakage—both	-15.4	-14.9	-28.6	<0.001	<0.001	0.017
Perceived cause/type of leakage						
Before reaching a toilet	-5.9	-5.1	-28.6	0.77	0.99	0.16
When cough/sneeze	-3.5	-3.1	-14.3	0.002	0.004	0.32
Whilst asleep	0.0	0.0	n/a*	0.054	0.054	n/a*
During physical activity/exercise	-2.0	-1.5	-14.3	0.001	0.002	0.32
After urination and are dressed	-13.8	-13.3	-28.6	0.027	0.046	0.16
For no obvious reason	-3.0	-3.1	0.0	0.006	0.009	0.32
All the time	-0.5	-0.5	n/a*	0.014	0.014	n/a*
Frequency of protection use	-2.0	-2.1	0.0	0.034	0.011	0.32
Usual amount of leakage	-12.2	-12.7	0.0	<0.001	<0.001	0.32
Worst amount of leakage	-7.3	-7.6	0.0	<0.001	<0.001	0.028
Interference with everyday life	-12.3	-12.8	0.0	<0.001	<0.001	0.041
Interference with social life	-9.2	-9.6	0.0	<0.001	<0.001	0.10
Interference with sex life	-1.8	-1.8	n/a*	0.001	0.001	n/a*
Overall quality of life	37.5 better	36.8 better	57.1 better	0.15	0.24	0.12
	34.4 same	34.6 same	28.6 same			
	28.1 worse	28.7 worse	14.3 worse			

*No patient reported this symptom pre- or post-intervention.

TABLE VI. Change (%) in Respondents Reporting 'Any Level of Symptom' Before and After Surgical Intervention

Symptom/item	Percentage change (%)			P-value		
	Males and females	Males	Females	Males and females	Males	Females
Frequency of leakage	-39.5	-28.6	-41.9	<0.001	<0.022	<0.001
Frequency of leakage—both	-47.4	-57.1	-44.9	<0.001	0.018	0.0001
Perceived cause/type of leakage						
Before reaching a toilet	-34.2	-57.1	-29.0	0.020	0.71	0.004
When cough/sneeze	-50.0	-42.9	-51.6	0.016	1.00	0.011
Whilst asleep	-2.7	-14.3	0.0	0.68	0.65	1.00
During physical activity/exercise	-47.3	-42.9	-48.4	0.040	0.78	0.023
After urination and are dressed	-15.8	0.0	-19.4	0.18	0.78	0.18
For no obvious reason	-34.2	-42.9	-32.3	0.15	0.083	0.039
All the time	-31.5	-57.1	-25.8	0.27	0.58	0.36
Frequency of protection use	-42.1	-42.9	-41.9	<0.001	0.049	0.0002
Usual amount of leakage	-39.5	-42.9	-45.2	<0.001	0.024	0.0001
Worst amount of leakage	-38.9	-28.6	-41.4	<0.001	0.016	0.0001
Interference with everyday life	-46.1	-42.9	-46.8	<0.001	0.018	0.0001
Interference with social life	-50.0	-85.7	-41.9	<0.001	0.018	0.0004
Interference with sex life	-40.9	-66.7	-37.6	0.003	0.17	0.010
Overall quality of life	61.1 better	71.4 better	58.6 better	0.010	<0.001	<0.001
	13.9 same	14.3 same	13.8 same			
	25.0 worse	14.3 worse	27.6 worse			

TABLE VII. Agreement Between ICIQ Score and ICSmaleSF and KHQ Scores

Score	Pearson's <i>r</i>	<i>P</i> -value	95% confidence interval*
ICSmaleSF incontinence score	0.74	<0.001	¹ 13.0–16.9, ² 9.4–13.7
ICSmaleSF voiding score	0.26	0.20	¹ 13.0–16.9, ² 4.7–8.7
BFLUTS incontinence factor	0.80	<0.001	¹ 12.5–14.3, ² 9.8–11.3
BFLUTS quality of life factor	0.60	<0.001	¹ 12.7–14.5, ² 7.0–8.5
KHQ	0.72	<0.001	¹ 10.7–12.8, ² 36.4–45.0

*¹ICIQ, ²ICSmaleSF/BFLUTS/KHQ.

with the ICSmaleSF 'Voiding' subscore (Table VII) [Altman, 1991]. This was anticipated, given the incontinence-specific nature of the ICIQ. The score also correlated strongly with the BFLUTS 'Incontinence Factor' and moderately with the BFLUTS 'Quality of Life Factor'. A reasonable agreement was also found between the ICIQ and KHQ scores. The test-retest reliability of the ICIQ score was 'good' (92.0%), with a high Kappa value of 0.74 ($P < 0.001$). Finally, highly significant improvements in scores were found following both conservative management (mean score 5.6 and 4.0 before and after treatment respectively: $P < 0.001$) and surgical intervention (mean score 15.2 and 5.8 before and after treatment respectively: $P < 0.001$), with greater decreases for patients who underwent surgical intervention than conservative management (mean decrease -9.5 and -1.7 respectively: $P < 0.001$).

DISCUSSION

The variable relationship between the level and impact of symptoms of incontinence has been widely acknowledged [Wyman et al., 1987; Hunskaar and Vinsnes, 1991; Peters et al., 1997]. Whilst incontinence is prevalent among the general population, it is not always bothersome, even to those reporting severe symptoms [Hunskaar and Vinsnes, 1991]. To assess the impact of incontinence comprehensively, it is therefore necessary to measure both the level of an individual's symptoms and the extent to which they impair their life.

The ICIQ has been developed to provide a simple, brief and robust questionnaire to assess the symptoms and impact of incontinence that could be used universally in clinical practice and research. The ICIQ has been shown to have high levels of validity, reliability and sensitivity, evaluated according to standard psychometric methods. The ICIQ is easily completed with low levels of missing data and studies have confirmed that it accurately measures incontinence (content validity) [Nunnally and Bernstein, 1994], adequately reflects known theories relating to incontinence (construct validity) [Rust and Golombok, 1999] and exhibits expected relationships with other measures of related concepts (convergent validity) [Litwin, 1995]. Assessments of internal consistency and stability have demonstrated that the ICIQ is highly reliable, providing consistent, stable and reproducible data [Litwin, 1995]. This is particularly important if it is to be used to monitor the status of patients' symptoms over time [Fayers and Machin, 2000]. The ability of the ICIQ to detect changes in

the level and impact of patients' symptoms following various treatments for incontinence has also been demonstrated. Again, the sensitivity of a questionnaire is vital if it is to be used as an outcome measure to monitor a patient's status, to assess clinically important changes following treatment or to assess the efficacy of an intervention [Fayers and Machin, 2000]. Finally, a simple scoring system has been developed that is suitable, scientifically justified and demonstrates adequate psychometric properties.

The scored items shall be of particular use in research. Several existing questionnaires to assess incontinence cover a wide range of symptoms and QoL issues, and most are rather lengthy. The trade-off between producing a questionnaire that is both efficient *and* comprehensive enough to serve the needs of patients, clinicians and researchers alike is an issue that developers are often faced with [Donovan et al., 2002]. Simplicity and brevity of interpretation are offset by crudeness and the loss of qualitative description to some degree. Whilst the ICIQ was intended to be comprehensive, the objective to also keep it brief meant that all known symptoms of incontinence would not be included. Instead, items were selected, using the methods described above, to produce a questionnaire that would be comprehensive for the assessment of the frequency, severity and impact on QoL of urinary incontinence in the widest range of patients possible. Item reduction techniques were then used to produce a shorter questionnaire that can be more efficiently administered and places less burden on respondents. Short questionnaires are ideal for routine administration in clinical settings, particularly when frequent assessment is required or where time is limited. The simplicity and brevity of the ICIQ will make it of use to general practitioners and clinicians in primary and secondary care institutions to screen for incontinence, to obtain a brief yet comprehensive summary of the level, impact and perceived cause of symptoms of incontinence and to facilitate patient-clinician discussions. Studies requiring a more detailed investigation of these aspects should use, in addition, other longer questionnaires. Whilst the items in the ICIQ may be scored to give an overall indication of the level and impact of incontinence, the responses to individual items may also be interpreted individually. In clinical practice, for example, it may be important to monitor changes in a specific symptom over time or following treatment. Levels of missing data for ICIQ items are rare and therefore no recommendations are made for imputing missing values. Where scored items have

missing data, users are advised to consider each remaining item separately.

The Cronbach's alpha for the three scored items indicated a high level of internal consistency. The fact that the decrease in the alpha value after discarding redundant items was only slight was expected, given that the remaining items still measured related concepts. However, the decision to remove further items was rejected given the intention of the ICIQ to provide a brief summary view of the frequency, amount and impact of incontinence. The Incontinence Severity Index [Sandvik et al., 1993] is a simple measure to assess the severity of incontinence among females, comprising just two items assessing the frequency and the amount of leakage. A major difference between that and this questionnaire is that the ICIQ allows patients to express their perception of the impact of leakage and whether they consider it to be a problem. The ICIQ simultaneously assesses the self-perceived impact of incontinence alongside symptom severity. This is particularly important when making decisions regarding whether an individual is likely to require or benefit from treatment and in evaluating the effectiveness of treatments. The non-scored self-diagnostic item was included by the expert committee because it was thought to be useful in clinical practice, to understand the patient's perception of the cause of their incontinence.

The Second ICI recommended that all randomised trials evaluating treatments for incontinence should employ standardised, validated questionnaires to assess their impact on patient outcome [Donovan et al., 2002]. Much larger improvements were detected for all symptoms in those patients who received surgical intervention in comparison with those who received conservative management. This suggests that the ICIQ is not only responsive to relevant and clinically important changes following intervention but is also sensitive enough to detect *varying* levels of change from different interventions that are expected to induce varying levels of improvement. This should ensure it is a particularly useful measure of outcome, as well as in epidemiological surveys. The use of the brief ICIQ should facilitate the comparison of data across studies. The present study investigated responsiveness to overall treatment modalities in relatively small samples. Further assessment of the questionnaire using larger numbers and a larger surgical series is required. It will also be important to explicitly investigate the responsiveness of the ICIQ to individual treatments.

Existing incontinence questionnaires have generally been developed for use in specific patient groups, such as men or women or with urge or stress incontinence. This precludes their universal application and has been a barrier for comparative research. The present study has confirmed that the ICIQ is applicable and performs well in individuals of varying ages, genders, patient groups, settings and diagnoses. Its universal nature makes the ICIQ a suitable criterion measure for the routine evaluation of patients, as a standard epidemiological tool, or as a benchmark outcome measure in clinical practice

and research. On a global scale, the questionnaire has been adapted for use in 35 languages and other English-speaking populations in accordance with standard methods [Guillemin et al., 1993; Armitage and Berry, 1994] (contact authors for further information). Data from Japan have confirmed that the psychometric properties of the questionnaire have been retained throughout this process [Subcommittee of Symptom and Quality of Life Assessment, 2001]. The ICIQ is being increasingly used and further evaluated in both clinical practice and research in the UK and internationally [Subcommittee of Symptom and Quality of Life Assessment, 2001; Karantanis et al., 2002; Peters et al., 2004]. Further research is required, however, to investigate its usefulness in routine clinical practice, for specific groups, such as the frail elderly, and in the developing world. Publication here places the questionnaire in the public domain (Fig. 1), and researchers and clinicians are encouraged to use it freely.

CONCLUSIONS

The ICIQ is a brief and robust measure of the symptoms and impact of incontinence that is suitable for use in clinical practice and research. Its use will facilitate the comparison of findings from different settings and studies, and thereby enable a more consistent and unified approach to the assessment of urinary incontinence and its impact on people's lives.

REFERENCES

- Altman D. 1991. Practical statistics for medical research. London: Chapman and Hall.
- Armitage P, Berry G. 1994. Statistical methods in medical research. 3rd edition. Oxford: Blackwell Science Ltd.
- Bland M, Altman D. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1:307–10.
- Brookes S, Donovan J, Wright M, et al. The responsiveness and scoring of the Bristol Female Lower Urinary Tract Symptoms Questionnaire. *Am J Obstet Gynecol* (in press).
- Donovan J, Peters T, Abrams P, et al. 2000. Scoring the Short Form ICSmaleSF questionnaire. *J Urol* 164:1948–55.
- Donovan J, Badia X, Corcos J, et al. 2002. Symptom and quality of life assessment. In: Cardozo L, Khoury S, Wein A, editors. *Incontinence: Proceedings of the Second International Consultation on Incontinence*, July 1–3, 2001. 2nd edition. Plymouth: Health Publication Ltd. pp 267–316.
- Fayers P, Machin D. 2000. *Quality of life: Assessment, analysis and interpretation*. New York: John Wiley & Sons, Inc.
- Guillemin F, Bombardier C, Beaton D. 1993. Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. *J Clin Epidemiol* 46:1417–32.
- Herzog A, Fultz N. 1990. Prevalence and incidence of urinary incontinence in community-dwelling populations. *J Am Geriatr Soc* 38:273–81.
- Hunskar S, Vinsnes A. 1991. The quality of life in women with urinary incontinence as measured by the Sickness Impact Profile. *J Am Geriatr Soc* 39:378–82.
- Jackson S, Donovan J, Brookes S, et al. 1996. The Bristol Female Lower Urinary Tract Symptoms questionnaire: Development and psychometric testing. *Br J Urol* 77:805–12.
- Karantanis E, Fynes M, Stanton S. 2002. Comparison of the ICI-Q SF questionnaire and the 24 hour pad test with other measures for evaluating the severity of stress incontinence—A pilot study [abstract]. *Neurourol Urodyn* 21:315–6.

- Kelleher C, Cardozo L, Khullar V, et al. 1997. A new questionnaire to assess the quality of life of urinary incontinent women. *Br J Obstet Gynaecol* 104:1374–9.
- Litwin M. 1995. How to measure survey reliability and validity. London: Sage Publications Ltd.
- Nunnally J, Bernstein I. 1994. Psychometric theory. 3rd edition. New York: McGraw-Hill, Inc.
- Peters T, Donovan J, Kay H, et al. 1997. The International Continence Society “benign prostatic hyperplasia” study: The bothersomeness of urinary symptoms. *J Urol* 157:885–9.
- Peters T, Horrocks S, Stoddart H, et al. 2004. Factors associated with variations in older people’s use of community-based continence services. *Health Soc Care Community* 12:53–62.
- Rust J, Golombok S. 1999. Modern psychometrics: The science of psychological assessment. 2nd edition. London: Routledge.
- Sandvik H, Hunskaar S, Seim A, et al. 1993. Validation of a severity index in female urinary incontinence and its implementation in an epidemiological survey. *J Epidemiol Community Health* 47:497–9.
- Subcommittee of Symptom and Quality of Life Assessment in International Consultation on Incontinence, Gotto M, Donovan J, Corcos J, et al. 2001. Scored ICIQ-SF (International Consultation on Incontinence Questionnaire—Short Form) for symptoms and QOL assessment in patients with urinary incontinence. *J Jpn Neurog Bladder Soc* 12: 227–31.
- Wyman J, Harkins S, Choi S, et al. 1987. Psychosocial impact of urinary incontinence in women. *Obstet Gynecol* 70:379–81.