The responsiveness of the ICS*male* questionnaire to outcome: evidence from the ICS-'BPH' study

J.L. DONOVAN[1], S.T. BROOKES, J.J.M.C.H. DE LA ROSETTE[2], T.J. PETERS[1], D. PORRU[3], A. KONDO[4], N. DABHOIWALA[5], R. MILLARD[6], R. BOSCH[7], J. NORDLING[8], A. MATOS FERREIRA[9], K. HÖFNER[10], H. MOSTAFID[11], S. WALTER[12], I. NISSENKORN[13], C. FRIMODT MOLLER[14], M. MENDES SILVA[15], C. CHAPPLE[16] and P. ABRAMS[17]

[1]Department of Social Medicine, University of Bristol, UK, [2]Department of Urology, Nijmegen, The Netherlands, [3]Urologic Clinic, University of Cagliari, Italy, [4]Department of Urology, Nagoya University School of Medicine, Japan, [5]Department of Urology, Amsterdam, The Netherlands, [6]The Prince Henry Hospital, Sydney, Australia, [7]Department of Urology, Rotterdam, The Netherlands, [8]Department of Urology, Herlev, Denmark, [9]Department of Urology, Lisbon, Portugal, [10]Department of Urology, Hannover Medical School, Germany, [11]Department of Urology, Guy's, London, UK, [12]Department of Urology, Odense, Denmark, [13]Department of Urology, Kfar Saba, Israel, [14]Department of Urology, Hellerup, Denmark, [15]Servico de Urologia, Lisbon, Portugal, [16]Department of Urology, Sheffield, UK and [17]Bristol Urological Institute, Southmead Hospital, Bristol, UK

Objective To evaluate the responsiveness of the ICS*male* questionnaire to the outcome of treatments for lower urinary tract symptoms (LUTS).

Patients and methods Consecutive men aged >45 years attending 23 urology centres in 12 countries, with symptoms suggestive of bladder outlet obstruction secondary to benign prostatic hyperplasia (BPH), were recruited to Phase I of the International Continence Society (ICS)-'BPH' study. In Phase II of the ICS-'BPH' study, 355 men in 15 centres in nine countries were followed up, having proceeded to treatment according to clinical practice. All men completed the ICS-'BPH' study questionnaire at baseline and follow-up, including the ICS*male* which concerns LUTS and related problems.

Results Patients included in Phase II were similar to those in Phase I according to age and levels of baseline symptoms. Patients received a range of treatments: 32% TURP, 29% drug therapies, 20% watchful waiting, 9% minimally invasive therapies and 10% 'others' (including open prostatectomy). For patients who underwent TURP, most LUTS, including voiding and filling symptoms, were highly statistically significantly better at follow-up than at baseline ($P<0.0001$). For drug, minimally invasive and 'other' treatments, fewer LUTS were highly statistically significantly better. For those undergoing watchful waiting, no symptoms were significantly different between baseline and follow-up.

Conclusion The ICS*male* questionnaire, in addition to being psychometrically valid and reliable, is responsive to change in outcome.

Keywords Questionnaire, lower urinary tract symptoms, quality of life, benign prostatic hyperplasia

Introduction

Increasingly, questionnaires are being used to assess symptoms and the impact on quality of life of men with LUTS related to benign prostatic disease, e.g. the AUA symptom score [1], IPSS [2], DAN-PSS-1 [3], Maine Medical Assessment Program [4], ICS*male* questionnaire [5], Bolognese Questionnaire [6], BPH-HRQOL measure [7], BPH Impact Index [8] and ICS*QoL* questionnaire [9]. The most widely used measure, the IPSS, is now used commonly in randomized trials and other evaluative studies to assess outcome, particularly symptom improvement after intervention. However, there are concerns that the IPSS only encompasses some aspects of outcome. In particular, it contains only seven symptoms,

i.e. frequency, incomplete emptying, intermittency, urgency, weak stream, hesitancy and nocturia [2], while it has been shown that men with LUTS experience a much wider range of symptoms, some of which are considerably more troublesome than those included in the IPSS [3,5,10]. Recently, data indicating that the DAN-PSS may be more sensitive to outcome have been published [11].

The ICS*male* questionnaire is a wide-ranging questionnaire including 20 symptom questions in total, most of which also have a sub-question attached concerning the degree of problem caused. Thus far, the ICS*male* questionnaire has been shown to be psychometrically valid and reliable at baseline [5]. However, it is important that questionnaires used to assess outcome are evaluated to ensure that they are appropriately responsive to changes over time or after interventions. The ICS*male*

questionnaire was developed during Phase I of the ICS-'BPH' study, a large international study whose main aim was to investigate the relationships between reported symptoms and the results of urodynamic studies [12]. Phase II of the study has involved an observational follow-up of patients recruited to Phase I in centres willing to collaborate with this activity. This article reports on the evaluation of the responsiveness of the ICS*male* questionnaire to outcome after treatment among patients in Phase II of the ICS-'BPH' study.

Patients and methods

In all, 1271 consecutive men aged >45 years, and attending 23 urology departments in 12 countries, with symptoms suggestive of BOO secondary to BPO were recruited to Phase I of the ICS-'BPH' study [13]. Men were excluded if they had significant urological disease (e.g. prostate cancer), neurological disease, had undergone previous prostatic surgery, or were taking medication active on the lower urinary tract. Each patient underwent a physical examination, uroflowmetry and pressure-flow studies, and completed the ICS-'BPH' study questionnaire which includes the ICS*male* questionnaire [10]. Results concerning the psychometric validity and reliability of the ICS*male* questionnaire at baseline have been reported elsewhere [5].

Fifteen centres in nine countries agreed to participate in Phase II of the study. Patients proceeded to treatment according to normal clinical practice and clinicians were asked to follow patients up ≈12 months after the baseline evaluation; 422 patients in total were followed-up in Phase II of the ICS-'BPH' study, 355 of whom completed the ICS*male* questionnaire in full (84%).

Evaluating the responsiveness of the ICSmale questionnaire

There has been considerable controversy about the most appropriate methods of measuring change using questionnaires. There are three main aspects to the measurement of change; differentiating between those who change a great deal and those who change little, the identification of factors which are associated with a good outcome, and inferring treatment effects from group differences, commonly in clinical trials [14]. Where a questionnaire results in a simple score, treatment effects can be assessed by examining differences before and after treatment between the intervention and control group, using unpaired *t*-tests, analysis of covariance or repeated-measures ANOVA [14]. However, the ICS*male* questionnaire does not produce a simple score and so responsiveness of each question was examined by comparing baseline and follow-up symptom levels for each of the major treatment groups.

Statistical methods

Frequency distributions were obtained for symptoms and related problems caused; ANOVA and Student-Newman-Keuls tests were used for comparisons between treatment groups with respect to age and length of follow-up [15]. An extension of Fisher's exact test was also used for comparing proportions. The Wilcoxon signed-ranks test for paired data was used to determine whether baseline and follow-up questionnaire items differed significantly within each treatment group. To adjust for multiple significance testing, a Bonferroni correction which maintains a 5% overall significance level would indicate that test-wise *P* values should only be considered significant if they are <0.0025 for these comparisons [15].

Results

ICS*male* questionnaire data were available for 355 patients followed after their initial recruitment from the following countries; Table 1 shows the numbers and proportions of patients followed. As only a proportion of patients was followed, it was important to examine whether they were different from the baseline group according to age and symptom levels. The proportions in each of the 10-year age bands were very similar for the follow-up group compared with the original ICS-'BPH' study baseline sample: 24% in those <60 years at follow-up, compared with 21% at baseline, 41% compared with 43% for the 60–69 years group, and 35% compared with 36% for those >70 years. The followed patients were also very similar to the full baseline sample in terms of their levels of 'any mention of a symptom'; eight symptoms were different by either

Table 1 Patients included in Phase II of the ICS-'BPH' study, by country

| Country | No. of patients (%) at | | |
	Baseline	Follow-up ICSmale	Died
Netherlands	391	144 (37)	2
Italy	58	48 (83)	
Denmark	121	36 (30)	
Japan	66	31 (47)	1
Australia	47	21 (46)	1
Portugal	49	18 (37)	
Germany	75	14 (19)	
UK*	66	33 (50)	
Israel	10	10 (100)	

*Includes 13 patients too late to be included in the original ICS-'BPH' study baseline dataset.

Table 2 Type of treatment by country, by age in 10-year age bands and the mean follow-up

Country	Type of treatment, n (%)				
	TURP	MIT	WW	Drugs	Other
Netherlands	16 (11)	20 (14)	52 (36)	52 (36)	4 (3)
Italy	36 (75)	0	1 (2)	7 (15)	4 (8)
Denmark	21 (58)	5 (14)	5 (14)	2 (6)	3 (8)
Japan	11 (35)	6 (19)	5 (16)	7 (23)	2 (7)
Australia	11	1	2	1	6
Portugal	3	0	4	7	4
Germany	0	0	0	14	0
UK	15 (45)	1 (3)	2 (6)	4 (12)	11 (33)
Israel	2	0	1	7	0
Age distribution, years					
<60	22 (20)	7 (23)	15 (24)	24 (26)	6 (27)
60–69	39 (36)	18 (58)	29 (46)	36 (41)	7 (32)
>70	47 (44)	6 (19)	19 (30)	29 (33)	9 (41)
Mean follow-up, months	22	15	10	11	22

MIT, Minimally invasive treatment.

0 or 1%, and nine by 2–4%. Only incomplete emptying varied by more than this, i.e. 81% in the baseline sample compared with 75% in the follow-up group.

Patients were followed for a mean of 16 months after baseline (range 2–44). Of the 355 men followed, 115 (32%) underwent TURP, 101 (29%) drug treatment, 72 (20%) watchful waiting, 33 (9%) a minimally invasive therapy (including 11 microwave, nine laser, 13 transurethral incision) and 34 (10%) some other procedure, including 10 an open prostatectomy. Follow-up times were statistically significantly different between the treatment groups (Table 2; $P < 0.001$). A Student-Newman-Keuls test, maintaining a 5% experiment-wise significance level, indicated that the differences were statistically significant only between surgical (particularly TURP and other treatments) and more conservative treatments (drug and minimally invasive therapies, and watchful waiting).

Table 2 also shows the differences in treatments by countries in Phase II of the study. The treatments received were statistically significantly different between countries, with patients in Italy, Denmark, Australia, Japan and the UK tending to undergo surgical treatments, and patients from the Netherlands, Portugal, Israel and Germany having more conservative treatments (an extension of the Fisher's exact test, $P < 0.0001$). Table 2 also shows the age distributions within treatment groups; although older men were more likely to undergo TURP, differences between groups were not statistically significant (ANOVA, $P = 0.14$).

For patients who underwent TURP, most LUTS (15 in all) were highly statistically significantly better at follow-up than baseline ($P < 0.0001$), i.e. frequency, nocturia, urgency, urge incontinence, bladder pain, hesitancy, straining to start, straining to continue, reduced stream, intermittency, dysuria, incomplete emptying, terminal dribble, post-void dribble, and repeated urination. Figure 1 shows the proportions of men with any mention of a symptom before and after each treatment. For patients who received 'other' treatments, seven symptoms were highly statistically significantly better at follow-up than baseline ($P < 0.0001$), i.e. nocturia, hesitancy, strain to start, strain to continue, reduced stream, intermittency, terminal dribble; and five more, i.e. urgency, urge incontinence, dysuria, repeated urination and frequency, were significantly better ($P < 0.0025$; Fig. 1). For patients who received minimally invasive treatments, there were fewer LUTS (four) that were highly statistically significantly better at follow-up than at baseline ($P < 0.0001$); nocturia, hesitancy, reduced stream and intermittency; with urgency and frequency significantly better ($P < 0.0025$; Fig. 1). For those receiving drug treatment, five LUTS reached the highest level of statistical significance ($P < 0.0001$), i.e. nocturia, urgency, hesitancy, reduced stream and incomplete emptying, with five more (bladder pain, strain to continue, intermittency, post-void dribble and repeated urination) significantly better ($P < 0.0025$). For those undergoing watchful waiting, no LUTS were significantly different between baseline and follow-up (Fig. 1).

All symptoms reaching statistical significance showed an improvement between baseline and follow-up, with the proportion of men improving being considerably higher for those undergoing TURP than other treatments. Table 3 shows the percentage change in 'any mention of a symptom' between baseline and follow-up. The different treatments had clearly different impacts on symptoms, with the surgical treatments (TURP, 'other' and minimally invasive therapies) having the greatest impact, and drugs and watchful waiting the least.

The picture was similar for the problems caused; problems relating to 10 LUTS (nocturia, urgency, hesitancy, strain to start, strain to continue, reduced stream, intermittency, incomplete emptying, terminal dribble and frequency) were statistically significantly better at follow-up than baseline for those undergoing TURP ($P < 0.0025$; Fig. 1). For 'other treatments', four LUTS were statistically significantly better (frequency, nocturia, urgency and terminal dribble, $P < 0.0025$). With minimally invasive therapies, problems related only to intermittency were statistically significantly better ($P < 0.0025$). With drugs, only problems associated with frequency and nocturia reached this level of significance, and again the watchful waiting group remained stable (Table 3).

Fig. 1. The proportion of patients reporting 'any mention of symptom' for symptoms and problems before (green) and after (red) (a) TURP, (b) 'other' treatments, (c) minimally invasive therapies, (d) drugs and (e) watchful waiting.

Table 3 The change in symptoms (%) from before to after each treatment

Symptom	TURP	MIT	WW	Drugs	Other
Nocturia	−48	−28	16	14	−41
Urgency	−27	−19	0	−9	−23
Urge incontinence	−29	9	1	−7	−32
Bladder pain	−23	−3	−1	14	−31
Stress incontinence	6	9	−5	1	−18
Misc incontinence	−4	−4	6	2	−9
Hesitancy	−44	−47	4	−10	−40
Strain to start	−35	−43	−28	−2	−63
Strain to continue	−42	−28	−1	−6	−59
Reduced stream	−37	−44	−4	−4	−33
Intermittency	−34	−10	0	−3	−35
Dysuria	−18	−19	0	−11	−39
Incomplete emptying	−24	−28	3	−3	−27
Terminal dribbling	−19	−16	2	1	−18
Post-void dribble	−20	−28	−10	−5	−32
Nocturnal incontin.	−2	−4	−3	−4	−17
Repeated urination	−37	−34	−6	−19	−43
Acute retention	−11	0	5	−1	−6
Frequency	−36	−37	0	−7	−38

MIT, Minimally invasive treatment.

Discussion

These results suggest that the ICS*male* questionnaire is sensitive to expected changes after interventions. The changes were most apparent and significant for patients who underwent TURP, where most LUTS (15 in all) were very significantly better at follow-up than at baseline ($P<0.0001$). Those who received drug therapies, five LUTS were highly significantly better, with five more significantly so; those receiving 'other' therapies also had improvements at follow-up. For minimally invasive therapies, four LUTS were very significantly better at follow-up than baseline and two were significantly better. No LUTS were statistically significantly different between baseline and follow-up for those undergoing watchful waiting.

The patterns of symptom improvement are intriguing; after TURP, the greatest number and widest range of symptoms were significantly improved, including both filling and voiding symptoms. There were no symptoms that were significantly improved by other treatments but not by TURP. There was some consistency in the symptoms that were significantly improved by all interventions (i.e. anything other than watchful waiting), with nocturia, hesitancy, urgency, intermittency and reduced stream significantly improved by all active interventions.

However, there were several symptoms which were improved only after TURP and one or two other treatments; e.g. frequency was significantly improved after any of the surgical treatments (TURP, minimally invasive and 'other' treatments), but not after drugs. Straining to continue and repeated urination were significantly improved after drugs and 'other' treatments. Urge incontinence, straining to start, dysuria and terminal dribble were significantly improved after 'other' treatments and TURP. Similarly, bladder pain, incomplete emptying and post-void dribble were significantly improved after drugs and TURP. These patterns suggest that treatments could be targeted towards patients with particular symptom profiles, although these findings will need to be confirmed by further research.

The picture was similar for the degree of problem caused by symptoms as for the symptoms themselves, with TURP having the greatest impact, and drugs, minimally invasive and 'other' treatments having smaller effects, with the watchful-waiting group remaining stable.

Although these results are very significant statistically and likely to be robust, they are based on observational data, and so it is important to consider the existence and strength of biases that might be present. For example, it is clear that there are similarities as well as differences between the treatment groups in terms of age and symptom levels. There were no statistically significant differences in age among the treatment groups. Symptom levels were similar for those undergoing TURP or minimally invasive therapies, and were also similar for those receiving drug therapies and watchful waiting. However, as would be expected, men who underwent surgical, 'other' and minimally invasive treatments had, on the whole, higher levels of symptoms than those who received more conservative therapies (drugs and watchful waiting).

It was possible only to follow a proportion of patients recruited to Phase I of the ICS-'BPH' study in Phase II. Centres were requested to follow-up as many patients as possible, but it is evident from Table 1 that the completeness of the follow-up varied between centres. The patients followed were apparently similar to those who were not in terms of age distribution and levels of symptoms, but it remains a possibility that the ICS*male* questionnaire may not have been responsive to the outcomes of patients who were not followed-up.

Although there are several differences between countries and treatment groups at both baseline [16] and follow-up, the strength of the significant improvements in symptoms between baseline and follow-up suggests that the ICS*male* questionnaire is sensitive to changes after interventions. However, these findings will need to be confirmed by truly comparable patient groups in randomized controlled trials.

In conclusion, the results show that, among patients who were followed, symptom and problem questions in the ICS*male* questionnaire are sensitive to changes that might be expected after various treatments. Most LUTS

(and their resultant problems caused) improved significantly after TURP, including both storage and voiding symptoms. Fewer symptoms exhibited this pattern for men receiving 'other' treatments, drugs and minimally invasive therapies such as laser and microwave treatment. No significant changes were apparent for those who had undergone watchful waiting. This evidence suggests that the ICS*male* questionnaire, in addition to being psychometrically valid and reliable [5] is responsive to changes in outcome.

Acknowledgements

References

1 Barry MJ, Fowler FJ, O'Leary MP *et al.* The American Urological Association Symptom Index for benign prostatic hyperplasia. *J Urol* 1992; **148**: 1549–57

2 Cockett ATK, Khoury S, Aso Y *et al.* eds, *Proceedings of the Second International Consultation on Benign Prostatic Hyperplasia.* Jersey: Scientific Communication International Ltd 1993

3 Hansen BJ, Flyger H, Brasso K *et al.* Validation of the self-administered Danish Prostatic Symptom Score (DAN-PSS-1) system for use in benign prostatic hyperplasia. *Br J Urol* 1995; **76**: 451–8

4 Fowler FJ Jr, Wennberg JE, Timothy RP, Barry MJ, Mulley AG Jr, Hanley D. Symptom status and quality of life following prostatectomy. *JAMA* 1988; **259**: 3018–22

5 Donovan JL, Abrams P, Peters TJ *et al.* The ICS-'BPH' Study: the psychometric validity and reliability of the ICSmale questionnaire. *Br J Urol* 1996; **77**: 554–62

6 Bolognese JA, Kozloff RC, Kunitz SC, Grino PB, Patrick DL, Stoner E. Validation of a symptoms questionnaire for benign prostatic hyperplasia. *Prostate* 1992; **21**: 247–54

7 Lukacs B, Leplege A, MacCarthy C, Comet D. Construction and validation of a BPH specific health related quality of life scale including evaluation of sexuality. Proceedings of the AUA. *Urology* 1997; **153** (suppl)

8 Barry M, Fowler F, O'Leary M, Bruskewitz R, Holtgrewe H, Mebust W. Measuring disease-specific health status in men with benign prostatic hyperplasia. *Med Care* 1995; **33**: AS145–AS155

9 Donovan JL, Kay HE, Peters TJ *et al.* Using ICSQoL to measure the impact of lower urinary tract symptoms on quality of life: evidence from the ICS-'BPH' study. *Br J Urol* 1997; **80**: 712–21

10 Peters TJ, Donovan JL, Kay HE *et al.* The International Continence Society 'Benign Prostatic Hyperplasia' Study: The bothersomeness of urinary symptoms. *J Urol* 1997; **157**: 885–9

11 Hansen BJ, Mortensen S, Mensink HJA *et al.* Comparison of the Danish prostate symptom score with the IPSS, Madsen-Iversen and Boyarsky symptom indexes. *Br J Urol* 1998; **81**: 36–41

12 de la Rosette JJMCH, Witjes WPJ, Schäfer W *et al.* Relationships between lower urinary tract symptoms and bladder outlet obstruction: results from the ICS-BPH study. *Neurourol Urodynam* 1998; **17**: 99–108

13 Abrams P, Donovan JL, de la Rosette JJMCH, Schafer W, and the ICS-'BPH' Study Group. International continence society 'benign prostatic hyperplasia' study: background, aims and methodology. *Neurourol Urodynam* 1997; **16**: 79–91

14 Streiner DL, Norman GR. *Health Measurement Scales.* Oxford: OUP 1989

15 Armitage P, Berry G. *Statistical Methods in Medical Research.* Oxford: Blackwell Science, 1987

16 Witjes WPJ, de la Rosette JJMCH, Donovan JL *et al.* The ICS-'BPH' Study: International differences in lower urinary tract symptoms and related bother. *J Urol* 1997; **157**: 1295–300

Authors

J.L. Donovan, PhD, Senior Lecturer.
S.T. Brookes, MSc, Research Associate.
J.J.M.C.H. de la Rosette, Urologist.
T.J. Peters, PhD, Reader in Medical Statistics.
D. Porru, Urologist.
A. Kondo, Urologist.
N. Dabhoiwala, Urologist.
R. Millard, Urologist.
R. Bosch, Urologist.
J. Nordling, Urologist.
A. Matos Ferreira, Urologist.
K. Höfner, Urologist.
H. Mostafid, Urologist.
S. Walter, Urologist.
I. Nissenkorn, Urologist.
C. Frimodt Moller, Urologist.
M. Mendes Silva, Urologist.
C. Chapple, Urologist.
P. Abrams, MD FRCS, Consultant Urologist.
Correspondence: Dr J.L. Donovan, Department of Social Medicine, University of Bristol, Canynge Hall, Whiteladies Road, Bristol BS8 2PR, UK.