ICIQ-Satisfaction: Psychometric testing of a new patient reported outcome measure for the evaluation of satisfaction after urological surgery

Authors

Alan D. Uren[1], Nikki Cotterill[2], Hashim Hashim[1], Jo Worthington[3], Dharmesh Kapoor[4], Paul Abrams[1]

1. Bristol Urological Institute

2. University of the West of England and the Bristol Urological Institute

3. University of Bristol

4. Royal Bournemouth Hospital

Abstract

Objectives: The psychometric testing of a patient reported outcome measure (PROM) to assess satisfaction after urological surgery.

Subjects/patients and methods: Following item development in 2008, the developmental ICIQ-S (dICIQ-S) was used in a randomised control trial comparing two types of surgery for prostatic obstruction. The dICIQ-S was administered at 6 weeks, 3 months and 12 months after surgery. Reliability was assessed by Cronbach's α and construct validity by the correlation of scores with concurrently administered PROMs of known validity: ICIQ-MLUTS, I-PSS, and the ICIQ-LUTSqol.

Results: A total of 410 men were included in the trial. Item-level missing data was generally low (mean 2.3%, range <1% to 9.2%). High ceiling effects were found in all items. Factor analysis identified six items which were related to surgical outcomes, which have formed the scored part of the ICIQ-S. Cronbach's α for the scored items was 0.89. As hypothesised, post-surgery dICIQ-S scores were correlated with reduced symptoms and improved quality of life as measured by the concurrent PROMs.

Conclusion: The results support the validity and reliability of the tool for evaluating satisfaction after urological surgery. The final ICIQ-S consists of six scored items for the assessment of satisfaction with 'surgical outcomes' followed by eight unscored items and is recommended for use in clinical practice and research.

Key words

Satisfaction, urology, surgery, patient reported outcome

Introduction

Although not well-defined in the literature, patient satisfaction is likely to be a multidimensional concept made up of patient experience, outcomes, expectations, and the fulfilment of these expectations (Bjertnaes et al., 2012; Sixma et al., 1998). The assessment of patient satisfaction using patient reported outcome measures (PROMs) is recognised as an important indicator of the quality of healthcare delivery (Al-Abri and Al-Balushi, 2014; Urden, 2002).

The measurement of patient satisfaction using PROMs has been used for patient-centered research into lower pelvic tract dysfunction (Akakpo et al., 2017; Mishriki et al., 2008). When investigating the success of different treatments for stress urinary incontinence (Monarch Study), patient satisfaction was high despite lower cure rates, suggesting a complete cure for incontinence may not be essential for satisfaction with treatment (Freeman et al., 2011). This is just one example of the importance of PROMs for the additional insight to be gained from the patient perspective (Robinson et al., 2003). However, in order to achieve the objective of reliably capturing patient feedback, the instrument used should be shown to be reliable and valid (Castro Diaz et al., 2017; Urden, 2002). Hence, we have identified a need for a validated multi-item PROM for the measurement of patient satisfaction after urological procedures.

The International Consultation on Incontinence Questionnaire (ICIQ) modules offer a range of patient reported outcome measures for lower pelvic dysfunction (www.ICIQ.net). They are recommended by the International Consultation on Incontinence for use in clinical trials in order to standardise outcome assessment (Abrams et al., 2006; Castro Diaz et al., 2017). The following describes the development of a new ICIQ module, the ICIQ-S, to assess patient satisfaction after urological surgery.

Subjects/patients and methods

The developmental 20-item ICIQ-S (dICIQ-S) was derived in 2008 using fourteen concept elicitation interviews. Draft items were refined by a team of multidisciplinary clinical experts alongside cognitive interviews with a further 11 patients until the items were deemed to be comprehensive and interpreted as intended.

The dICIQ-S was used within a multi-centre randomised control trial for the comparison of the clinical and cost-effectiveness of transurethral resection of the prostate (TURP) versus thulium laser transurethral vaporesection of the prostate (ThuVARP) (Worthington et al., 2017), ISRCTN registry 00788389, to evaluate its psychometric properties. As part of the trial protocol, the following questionnaires of known validity were administered at baseline pre-surgery and then alongside the dICIQ-S at 6 weeks, 3 months and 12 months post-surgery: the ICIQ-MLUTS (Donovan et al., 1996) and the I-PSS (Barry et al., 1992) for the measurement of male lower urinary tract symptoms, and the ICIQ-LUTSqol for the measurement of the impact of urinary symptoms on quality of life (Kelleher et al., 1997).

Statistical analysis

The statistical software package SPSS was used for the analysis, in order to evaluate the psychometric properties of the dICIQ-S in the trial population. Items were considered to have a floor effect or ceiling effect if the frequency of responses were found to be skewed to either the ‘least

satisfied' or 'most satisfied' options respectively. Instrument reliability was evaluated by internal consistency (Cronbach's α) where a value of ≥0.7 was considered acceptable and >0.9 to indicate probable redundancy within the item pool (Nunnally and Bernstein, 1994). Exploratory factor analysis identified items which could be grouped together by their common variance (Yong and Pearce, 2013), and informed the instrument's potential scoring. Construct validity was investigated by using Spearman's correlation coefficient to examine expected relationships according to known theory (Hays et al., 1993). Hypotheses were that the instrument should detect a higher satisfaction if patients had less urinary symptoms and reduced impact on quality of life post-surgery (Mishriki et al., 2008). Positive expectations may also influence patient satisfaction with outcomes (Marschall-Kehrel et al., 2006) so patients with more severe baseline symptoms might be expected to have higher reported satisfaction post-surgery. To investigate these hypotheses, the dICIQ-S scores were compared with the overall symptom score from the ICIQ-MLUTS (scored from 0-52), the overall quality of life item 22 from the ICIQ-LUTSqol (scored 0-10 on a Likert scale), and the overall impact on quality of life item within the I-PSS (scored 0-6 on a Likert scale). For each of these PROMs a higher score indicates a higher symptom severity or impact on quality of life. Mean scores were calculated and one-way analysis of variance used to evaluate any trends across the repeat administrations.

Results

A total of 410 men (mean age 69.5) were recruited into the trial. The number of patients that completed each of the questionnaires at each time-point is shown in table 1.

Table 1. Numbers of questionnaires completed and used for the analysis at each time-point.

Questionnaire	Baseline (n, %)	6 weeks (n, %)	3 months (n, %)	12 months (n, %)
ICIQ-S	n/a	338 (82)	343 (84)	346 (84)
IPSSqol	187*(46)	335 (82)	335 (82)	339 (83)
ICIQ-MLUTS	177*(43)	301 (73)	323 (79)	324 (79)
ICIQ-LUTSqol	186* (45)	333 (81)	338 (82)	335 (82)

*PROM data was only analysed for patients who did not have an indwelling catheter.

There were low levels of missing data for every item (mean 2.3%, range <1% to 9.2%) except for item 12 'complications' which had missing data of 6.8% and satisfaction with sex-life which was 9.2% (item 14) (table 2). Ceiling effects were present in all of the items (frequency of responses to items ranged from 41% to 91% for the most 'satisfied' response option). No floor effects were found (percentage of responses >100/X (where X was the number of response options).

Table 2. Item response distribution of the dICIQ-S at 12 months.

dICIQ-S item	**% missing data**	**% of responses most 'satisfied' response option**	**% of responses least 'satisfied' response option**
Q3 How would you rate the outcome of your surgery?	0.7	72	3.4
Q4 How prepared did you feel for the surgery?	1.0	73	1
Q5 How satisfied were you with any explanation about your	1.0	68	1

dICIQ-S item	% missing data	% of responses most 'satisfied' response option	% of responses least 'satisfied' response option
surgery?			
Q6 What do you think about the decision you made to have the surgery?	0.7	91	0.3
Q7 Overall, has the result of your surgery been...?	1.4	56	9
Q8 With regards to the surgery, how much pain did you suffer..?	1.7	46	2
Q9 With regards to the pain, was it..?	32***	41	11
Q10 With regards to the treatment you had for the pain, were you..?	35***	50	0.3
Q11 Compared to before your treatment, are your everyday activities..?	3.1	69	1
Q12 Were there any complications or side effects with your surgery?	6.8	69	7
Q13 Would you say you have been able to return to a 'normal life' after surgery?	1.7	63	1.7
Q14 Compared to before you had the surgery, is your sex life..?	9.2 (28)**	5.1	15
Q15 Would you recommend this surgery to friends or relatives with similar problems?	1.7	78	1.8
Q16 If you were in the same situation again, would you still have the same surgery?	1.4	80	1.4
Q17 Compared to how you felt before the surgery, how is your condition now?	1.7	78	1.7
Q18 Do you have any new symptoms that you did not have before the surgery?	1.7	80*	18*
Q19 If you had to spend the rest of your life with your symptoms as they are now, how would you feel?	1.0	64	1
Q20 Overall, how satisfied were you with your surgery?	1.7	63	0.7
Q21 Have you had any previous treatment or procedures for this problem?	3.1	83*	14*
Q22 If you had previous treatment or procedures for this problem, how satisfied were you?	26**	3.4	3.1

Notes: Item 1 was 'gender'. Item 2 was 'date of birth'.

* Yes/no response option

**% responded 'not applicable'
***Skip instruction if not applicable

dICIQ-S Scoring

Eight items had ordinal response options and were scored from 0-3, 0-4 or 0-5, depending on the number of response options. Items 1 and 2 were gender and age, and the other items were not applicable to all patients, or were not ordinal, so were not included in the scoring.

Reliability

As shown in table 3, overall internal consistency was high for the scored items (Cronbach's α 0.86). Exploratory factor analyses with parallel principal component analysis showed two underlying factors relating the items. Factor 1 contained six items (Cronbach's α of 0.89) which could be conceptually related by satisfaction with 'surgical outcomes'. Factor 2 contained items 4 and 5 which were conceptually related to the 'service experience'. The overall Cronbach's α was improved if items 4 or 5 were deleted. Item-total correlations for these items were relatively low at 0.24 and 0.46.

Table 3. Item reliability statistics and factor loadings of the dICIQ-S scored items (12 months).

Item (dICIQ-S)	**Item-total correlation**	**Cronbach's α if item deleted**	**Factor 1 'surgical outcomes'**	**Factor 2 'service experience'**
Q3 How would you rate the outcome of your surgery?	0.77	0.82	0.87	
Q4 How prepared did you feel for the surgery?	0.24	0.88		0.92
Q5 How satisfied were you with any explanation about your surgery?	0.46	0.86		0.52
Q13 Would you say you have been able to return to a 'normal life' after your surgery?	0.60	0.85	0.67	
Q15 Would you recommend this surgery to friends or relatives with similar problems?	0.79	0.83	0.84	
Q16 If you were in the same situation again, would you still have the surgery?	0.74	0.83	0.81	
Q17 Compared to how you felt before your surgery, how is your condition now?	0.75	0.83	0.85	
Q19 If you had to spend the rest of your life with your symptoms as they are now, how would you feel?	0.79	0.83	0.87	

Notes: Overall scored items Cronbach's α 0.86. Factor 1: Cronbach's α 0.89.

Factor extraction method: Principal component analysis. Rotation method: Varimax with Kaiser normalisation.

Construct validity

As hypothesised, higher levels of satisfaction post-surgery were associated with greater reduction in both symptom severity and improvement in quality of life after surgery. The dICIQ-S surgical outcome factor score and standalone item Q20 (overall satisfaction) were negatively correlated with ICIQ-MLUTS, ICIQ-LUTSqol and IPSSqol scores at all time points (Table 4). These correlations were all

statistically significant at the 0.01 level (Spearman's correlations). Alongside the dICIQ-S surgical outcome factor score, this also provides evidence of validity of the overall satisfaction item (Q20) as a valid standalone item score when administered alongside the other items. Baseline MLUTS, IPSS and Qol scores were not correlated with the dICIQ-S surgical outcome score at 6W, 3M and 12M (Table 5). Thus, no relationship was found between levels of satisfaction with outcomes after surgery, and the pre-surgery severity of symptoms or their impact on quality of life.

Table 4. Correlations of concurrent PROMs at 6 weeks, 3 month and 12 month administrations.

Concurrent instrument	dICIQ-S factor 1 score			ICIQ-S Q20 Overall satisfaction		
	6W	3M	12M	6W	3M	12M
ICIQ-MLUTS	-0.66	-0.65	-0.52	-0.55	-0.49	-0.54
ICIQ-LUTSqol	-0.70	-0.67	-0.60	-0.56	-0.51	-0.61
IPSSqol	-0.78	-0.72	-0.72	-0.59	-0.57	-0.62

Notes: Spearman's correlations used. All sig. at 0.01 level.

Table 5. Correlation of concurrent PROMs baseline scores with dICIQ-S surgical outcome scores at 6 weeks, 3 month and 12 month administrations.

Baseline concurrent PROM scores	dICIQ-S factor 1 score		
	6W	3M	12M
ICIQ-MLUTS	-0.27	-0.13	-0.04
ICIQ-LUTSqol	-0.15	-0.13	-0.13
IPSSqol	-0.00	-0.10	-0.09

Notes: Spearman's correlations used. All non-sig. at 0.05 level.

As shown in table 6, mean scores as measured by all three concurrent PROMs decreased over the study period (reported symptom severity reduced and impact on quality life improved). The dICIQ-S surgical outcome score, and 'overall satisfaction' item Q20 score correspondingly increased.

Table 6. Mean PROM scores at each administration.

PROM	Mean score			
	Baseline	6M	3M	12M
dICIQ-S factor 1 score (0 to 26)	n/a	22.1	22.6	23.3
dICIQ-S Q20 Overall satisfaction (0 to 10)	n/a	8.35	8.63	8.78
IPSSqol (0 to 6)	4.95	2.02	1.59	1.15
ICIQ-MLUTS (0 to 56)	22.1	10.2	8.62	7.30
ICIQ-LUTSqol (0 to 10)	6.3	2.86	2.06	1.38

Notes: One-way repeated measure ANOVA with Bonferroni correction used for pairwise comparisons. All were significant at the 0.05 level.

Final ICIQ-S

Decisions to modify, remove items and their response options were made by the development team using the available evidence. The final ICIQ-S consists of 6 scored and 8 unscored items (Table 7). Factor 1 is the scored section of the questionnaire and these six items may be summed in order to provide a score for the evaluation of satisfaction with 'surgical outcomes' (ICIQ-S outcome score, range 0-24). Factor 2 items were not scored separately as two items alone do not have sufficient

reliability. The unscored eight items are standalone items which can be analysed separately when required.

Six items were removed in the final version of the ICIQ-S that assessed the patient's satisfaction with their decision to undertake surgical treatment, return to physical activity, sex life, whether they had any new symptoms post-surgery, whether they had previous surgery, and satisfaction with any previous surgery. The response options for these items were not ordinal, had high missing data or were considered of limited utility by the development team. Missing data highlighted that the pain items were only relevant for analysis if there was a timeframe associated. Thus, the item for whether the severity of pain was more (or less) than expected, was replaced by an item 'how much pain are you experiencing now?' A minor change to the wording of the middle response option for three items was changed to clarify the midpoint of the ordinal scale.

Discussion

The current study has shown evidence to support the validity and reliability of the ICIQ-S, a PROM designed for the evaluation of patient satisfaction after urological surgery. The original items were derived from patient interviews and the psychometric performance of the PROM has now been tested in a large population of men after surgery for prostatic obstruction.

All of the items exhibited ceiling effects, but this is not unusual for satisfaction questionnaires (Yellen et al., 2002). Hence we are not concerned that the ceiling effects compromise the instrument's ability to detect differences between subjects (FDA, 2009), as these may simply reflect the 'true' levels of satisfaction due to the known efficacy of the intervention within the sample. Construct validity was evidenced by the expected association of a higher satisfaction score with a lower postoperative urinary symptom burden (ICIQ-MLUTS) and reduced impact on quality of life (ICIQ-LUTSqol, IPSSqol). This is encouraging evidence of validity and gives confidence in the measurements made. It also suggests that the ICIQ-S will be applicable to other interventions, not quite so successful as prostate resection. There was no association between the severity of baseline symptoms and the satisfaction with outcome score after surgery. Nevertheless, the apparent failure of an instrument to detect a relationship may not always be due to the instrument, but to the true absence of the relationship (Bowling, 2004). As found elsewhere, the post-operative symptom burden is likely to be the predominant factor, rather than the extent of improvement, when reporting outcome satisfaction (Kane et al., 1997). The mean scores of the dICIQ-S and overall satisfaction item increased over the time period and the corresponding trend was also evident for the other questionnaires. Following TURP, the continued improvement in satisfaction and outcomes over an extended follow-up period, has been found in other studies (Mishriki et al., 2008), and the detection of this trend is further evidence of the reliability of the new 'outcome score' and overall satisfaction item.

The intention is for the tool to be widely applicable for use in clinical practice and research after urological procedures. We also have a large amount of data from the MASTER study, a randomised controlled trial of the male sling and the artificial urinary sphincter for post-prostatectomy incontinence, which will be published elsewhere. This shows that even when surgery may not result in cure, the patient is satisfied and would recommend his operation to others: this finding is similar to the Monarch study in women (Freeman et al., 2011). In difficult to cure conditions, satisfaction becomes a concept that is more pragmatic and meaningful to patients and therefore relevant in the

assessment of health care. The intention is to develop the questionnaire so that it can be used after a variety of investigations and therapeutic interventions, including urodynamic studies, sacral nerve stimulation and botulinum toxin injections for overactive bladder as well as gynaecological procedures. Cognitive interviews with females post-pelvic organ prolapse surgery is part of the ongoing validation of the ICIQ-satisfaction in gynaecological surgery, that will be reported later. For different populations or procedures, we would encourage the *unscored* items to be adapted, and new items added as appropriate using standard methods of development, in order to tailor assessment for the specific population. Researchers and clinicians are encouraged to use the ICIQ-S to investigate its further utility in other contexts, including routine clinical practice and other patient populations.

Conclusion

The ICIQ-S is a 14-item instrument that covers aspects of experience, expectations, and outcomes to evaluate satisfaction after urological surgery. A statistically coherent scoring system has been derived for the evaluation of satisfaction with surgical outcomes for six items. The remaining eight items are unscored individual items and may be used according to clinical circumstances. The ICIQ-S may be downloaded along with the other ICIQ modules with permission at www.iciq.net.

Table 7. Final ICIQ-S including the scored and unscored items.

SCORED ITEMS: Outcome score sum score 1-6 (Range, 0-24)
1. How would you rate the outcome of your surgery?
2. Compared to how you felt before your surgery, how is your condition now? (0-4)
3. Would you say you have been able to return to a 'normal life' after your surgery? (0-3)
4. If you were in the same situation again, would you still have the surgery? (0-4)
5. Would you recommend this surgery to friends or relatives with similar problems? (0-4)
6. If you had to spend the rest of your life with your symptoms as they are now, how would you feel? (0-5)
UNSCORED ITEMS
7. How prepared did you feel for the surgery?
8. How satisfied were you with any explanation about your surgery?
9. After the surgery, how much pain did you experience?
10. How satisfied were you with any treatment you had for pain?
11. How much pain are you experiencing now?
12. Were there any complications or side effects with your surgery?
13. Overall, has the result of your surgery been…?
14. Overall, how satisfied were you with your surgery?

References

Abrams, P., Avery, K., Gardener, N., Donovan, J., and ICIQ Advisory Board (2006). The International Consultation on Incontinence Modular Questionnaire: www.iciq.net. J. Urol. *175*, 1063–1066; discussion 1066.

Akakpo, W., Pineda, M.A., and Burnett, A.L. (2017). Critical Analysis of Satisfaction Assessment After Penile Prosthesis Surgery. Sex. Med. Rev. *5*, 244–251.

Al-Abri, R., and Al-Balushi, A. (2014). Patient Satisfaction Survey as a Tool Towards Quality Improvement. Oman Med. J. *29*, 3–7.

Barry, M.J., Fowler, F.J., O'Leary, M.P., Bruskewitz, R.C., Holtgrewe, H.L., Mebust, W.K., and Cockett, A.T. (1992). The American Urological Association symptom index for benign prostatic hyperplasia. The Measurement Committee of the American Urological Association. J. Urol. *148*, 1549–1557; discussion 1564.

Bjertnaes, O.A., Sjetne, I.S., and Iversen, H.H. (2012). Overall patient satisfaction with hospitals: effects of patient-reported experiences and fulfilment of expectations. BMJ Qual. Saf. *21*, 39–46.

Bowling, A. (2004). Measuring health: A Review of Quality of Life Measurement Scales (Maidenhead, Berkshire, England; New York, NY: Open University Press).

Castro Diaz, D., Robinson, D., Bosch, R., Constantini, E., Cotterill, N., Espuna-Pons, M., Kocjancic, E., Lemos, N., Tarcan, T., and Yoshida, M. (2017). Patient-Reported Outcome Assessment. In Incontinence, 6th Edition 2017: International Consultation on Incontinence, Tokyo, September 2016, P. Abrams, L. Cardozo, A. Wagg, and A. Wein, eds. (International Continence Society), pp. 541–599.

Donovan, J.L., Abrams, P., Peters, T.J., Kay, H.E., Reynard, J., Chapple, C., De La Rosette, J., Kondo, A., and others (1996). The ICS-'BPH'Study: the psychometric validity and reliability of the ICSmale questionnaire. Br. J. Urol. *77*, 554–562.

FDA (2009). Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims.

Freeman, R., Holmes, D., Hillard, T., Smith, P., James, M., Sultan, A., Morley, R., Yang, Q., and Abrams, P. (2011). What patients think: patient-reported outcomes of retropubic versus trans-obturator mid-urethral slings for urodynamic stress incontinence--a multi-centre randomised controlled trial. Int. Urogynecology J. *22*, 279–286.

Hays, R.D., Anderson, R., and Revicki, D. (1993). Psychometric considerations in evaluating health-related quality of life measures. Qual. Life Res. Int. J. Qual. Life Asp. Treat. Care Rehabil. *2*, 441–449.

Kane, R.L., MacIejewski, M., and Finch, M. (1997). The Relationship of Patient Satisfaction with Care and Clinical Outcomes. Med. Care *35*, 714–730.

Kelleher, C.J., Cardozo, L.D., Khullar, V., and Salvatore, S. (1997). A new questionnaire to assess the quality of life of urinary incontinent women. BJOG Int. J. Obstet. Gynaecol. *104*, 1374–1379.

Marschall-Kehrel, D., Roberts, R.G., and Brubaker, L. (2006). Patient-reported outcomes in overactive bladder: The influence of perception of condition and expectation for treatment benefit. Urology *68*, 29–37.

Mishriki, S.F., Grimsley, S.J.S., Nabi, G., Martindale, A., and Cohen, N.P. (2008). Improved quality of life and enhanced satisfaction after TURP: prospective 12-year follow-up study. Urology *72*, 322–326; discussion 326-328.

Nunnally, J., and Bernstein, I. (1994). The assessment of reliability. In Psychometric Theory, (McGraw-Hill Inc), pp. 248–292.

Robinson, D., Anders, K., Cardozo, L., Bidmead, J., Dixon, A., Balmforth, J., and Rufford, J. (2003). What Do Women Want?: Interpretation of the Concept of Cure. Female Pelvic Med. Reconstr. Surg. *9*, 273.

Sixma, H.J., Kerssens, J.J., Campen, C.V., and Peters, L. (1998). Quality of care from the patients' perspective: from theoretical concept to a new measuring instrument. Health Expect. Int. J. Public Particip. Health Care Health Policy *1*, 82–95.

Urden, L.D. (2002). Patient satisfaction measurement: current issues and implications. Outcomes Manag. *6*, 125–131.

Worthington, J., Taylor, H., Abrams, P., Brookes, S.T., Cotterill, N., Noble, S.M., Page, T., Swami, K.S., Lane, J.A., and Hashim, H. (2017). A randomised controlled trial to determine the clinical and cost effectiveness of thulium laser transurethral vaporesection of the prostate (ThuVARP) versus transurethral resection of the prostate (TURP) in the National Health Service (NHS) - the UNBLOCS trial: a study protocol for a randomised controlled trial. Trials *18*, 179.

Yellen, E., Davis, G.C., and Ricard, R. (2002). The measurement of patient satisfaction. J. Nurs. Care Qual. *16*, 23–29.

Yong, A.G., and Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. Tutor. Quant. Methods Psychol. *9*, 79–94.