

A Simplified Urinary Incontinence Score for the Evaluation of Treatment Outcomes

Asnat Groutz, Jerry G. Blaivas,* and Jarrod E. Rosenthal

Weill Medical College, Cornell University, New York, New York

There are no standardized definitions for anti-incontinence therapy outcomes. The present study was conducted to evaluate whether the incorporation of several non-invasive outcome measures into a new score may serve as a meaningful outcome instrument. Ninety-four consecutive sphincteric incontinent women who underwent a pubovaginal sling by a single surgeon were enrolled. All patients underwent a full clinical evaluation, including pre- and post-operative questionnaires, 24-hour voiding diary, and 24-hour pad test. Surgery outcomes were classified twice: First, by analyzing the patient questionnaire, voiding diary, and pad test separately, according to previously published criteria, and second, by combining the three outcome tools into a new response score. The new score was constructed in a simple, easy-to-remember format and divided into five categories: cure, good response, fair response, poor response, and failure. All patients were evaluated at least 1 year post-operatively. Comparison of the old and new classifications suggests that the new response score provides a more accurate evaluation of the surgical outcomes. Although 64 to 69% of the patients were originally classified as cure according to the old classification, only 44.7% were re-classified as cure by the strict criteria employed in the new score. Furthermore, the response score also differentiates between various degrees of clinical improvement (i.e., good, fair, or poor response). Twenty-five (26.6%) patients, most of whom were previously classified as cure, were re-classified as good response, whereas 20 others were re-classified as fair (13.9%), or poor (7.4%) response. Seven (7.4%) patients were re-classified as surgical failures. All were diagnosed pre-operatively as having complex sphincteric incontinence. Specific failure rates were therefore 11.3% for complex and 0% for simple cases. In conclusion, the suggested post-operative response score incorporates in a user-friendly format three popular outcome tools (i.e., 24-hour diary, 24-hour pad test, and patient questionnaire) and seems to reflect the surgical results more accurately. Further studies are needed to assess its validity and reproducibility in other treatment modalities. *Neurorol Urodynam.* 19:127–135, 2000. © 2000 Wiley-Liss, Inc.

Key words: urinary incontinence; treatment; pubovaginal sling; outcome score

INTRODUCTION

The reported success rates of anti-incontinence surgery vary considerably according to surgical techniques, patient selection, length of follow-up, study methodology, and the investigator definition of cure. Although various outcome measures

*Correspondence to: Jerry G. Blaivas, M.D., 400 East 56 Street, New York, NY 10022.

Received 8 September 1999; Accepted 21 October 1999

have been proposed for the evaluation anti-incontinence surgery, no single tool has met with widespread acceptance. To date, there are no standard definitions of cure or failure and there are no standardized, widely accepted clinical tools to assess outcomes after anti-incontinence surgery.

Minimal standards by which the efficacy of therapy for urinary incontinence should be assessed and reported were previously proposed by the Urodynamic Society [Blaivas et al., 1997]. According to these recommendations, post-treatment evaluation should consist of at least: 1) the patient's opinion of treatment outcome; 2) micturition questionnaire; 3) voiding diary; 4) pad test; 5) physical examination; 6) uroflowmetry; and 7) estimation of post-void residual urine. It was further recommended that the specific method by which data were collected should be reported. This requirement was supported by several studies in which a marked disparity between the results of chart reviews, anonymous questionnaires, independent researches interviews, and treating physician interviews was revealed [Sirls et al., 1995; Haab et al., 1997]. However, we recently found a good agreement between validated patient questionnaire, 24-hour voiding diary, 24-hour pad test, and retrospective chart review as outcome measures of pubovaginal sling surgery [Chaikin et al., 1999]. Nevertheless, we believe these outcome instruments may not always reflect the real clinical status. The present study was therefore conducted to evaluate whether the incorporation of several non-invasive outcome measures into a new response score may serve as a more meaningful outcome instrument.

METHODS

Patients

Ninety-four consecutive women who underwent a pubovaginal sling by a single surgeon (J.G.B.) were enrolled. Eighty-four of these patients represent a subset of 251 women previously reported on for pubovaginal sling [Chaikin et al., 1998] and whose outcome results were also previously used to compare four different instruments for outcome analysis [Chaikin et al., 1999]. Data are presented as mean \pm standard deviation or percentage according to the variables.

Preoperative Evaluation

Pre-operatively, all patients underwent meticulous clinical evaluation, which included a complete history and physical examination, standardized urinary questionnaire, 24-hour voiding diary, 24-hour pad test, urine culture, non-invasive uroflowmetry, post-void residual urine volume, video urodynamics, and urethrocystoscopy. Multi-channel video urodynamics were performed according to the recommendations of the International Continence Society [Abrams et al., 1988] except for cystometry. Contrary to these recommendations, the patient was not instructed to try to inhibit micturition during the filling phase, but rather to report sensations to the examiner. The cystometrogram was performed using radiographic contrast and a 7-F double lumen catheter via constant infusion at a medium filling rate, with rectal pressure monitoring. Sphincteric incontinence was defined as visible urinary leakage during increased abdominal pressure in the absence of detrusor overactivity. Sphincteric incontinence was further classified as simple or complex according to previously published criteria [Chaikin et al., 1998]. Complex cases included urge incontinence,

“pipe stem urethra” (a fixed scarred urethra), urethral or vesicovaginal fistula, urethral diverticulum, grade 3 or 4 cystocele, or neurogenic bladder. Simple cases comprised women that did not meet the criteria for complex incontinence, including cases of detrusor instability without urge incontinence and cases of prior surgical failure.

Postoperative Evaluation

Post-operatively the women were scheduled to be evaluated at 1, 3, 6, and 12 months and thereafter annually. At each visit, history, focused examination with a full bladder, 24-hour voiding diary, 24-hour pad test, uroflowmetry, and post-void residual urine were obtained. All patients were contacted in person at follow-up or by telephone and asked to complete a validated post-operative voiding questionnaire [Haab et al., 1997] administered by a blinded investigator (J.E.R.). The 24-hour pad test was performed at the same time of the voiding diary, before the follow-up visit. Patients were allowed to wear as many pads as they believed necessary. The used pads were enclosed in a self-sealing plastic bag. The number and weight of the pads were recorded before being given to the patients and on return to the follow-up visit. The voiding diary included the following data: time of micturition, voided volumes (milliliters), incontinence episodes and type of incontinence (urge, stress, or unaware incontinence), although for the purpose of this analysis, no distinction was made between post-operative stress and urge incontinence.

Surgery outcomes were classified twice: First, by analyzing the patient questionnaire, voiding diary, and pad test separately, according to previously published criteria [Chaikin et al., 1999], and second, by combining the three outcome tools into a new response score. We then examined each case in which there was a disparity between the two analyses to find out whether the new response score reflects better the clinical outcome than each of the three tools separately.

The Old Classification

In the first analysis, we used the following criteria to define cure, improvement, and failure by each of the outcome tools separately:

1. 24-hour voiding diary
 - a. Cure: no urinary incontinence (urge or stress) episodes.
 - b. Improvement: post-operatively decrease of $\geq 50\%$ in the total number of incontinence episodes.
 - c. Failure: post-operatively decrease of $< 50\%$ in the total number of incontinence episodes.
2. 24-hour pad test
 - a. Cure: post-operatively total weight gain of the pads ≤ 2 g.
 - b. Improvement: post-operatively decrease of $\geq 50\%$ in the total weight gain.
 - c. Failure: post-operatively decrease of $< 50\%$ in the total weight gain.
3. Patient questionnaire
 - a. Cure: the patient considers herself cured by the operation and satisfied.
 - b. Improvement: the patient considers herself improved and will recommend it to a friend.
 - c. Failure: the patient considers the operation to have failed.

The New Response Score (Appendix 1)

In the second analysis, we rated the post-operative data according the following criteria:

1. 24- hour voiding diary
 - a. No urinary incontinence (urge or stress) episodes: 0 points.
 - b. 1–2 incontinence episodes: 1 point.
 - c. ≥ 3 incontinence episodes: 2 points.
2. 24-hour pad test
 - a. Total weight gain of the pads ≤ 8 g: 0 points.
 - b. Total weight gain of the pads 9–20 g: 1 point.
 - c. Total weight gain of the pads >20 g: 2 points.
3. Patient questionnaire
 - a. The patient considers herself as cured: 0 points.
 - b. The patient considers herself as improved: 1 point.
 - c. The patient considers the operation to have failed: 2 points.

The response score is the sum of the above sub-scores and is classified as:

Cure: total score 0
 Good response: total score 1–2
 Fair response: total score 3–4
 Poor response: total score 5
 Failure: total score 6.

RESULTS

Ninety-four consecutive patients were prospectively enrolled. The mean age of the patients was 58 ± 13 years. There were 32 simple (34%) and 62 complex (66%) cases of sphincteric incontinence. The mean pre-operative number of incontinence episodes per 24 hours was 7.9 ± 4.9 . The mean pre-operative 24-hour weight gain of the pads was 170.6 ± 144.5 g. All patients were evaluated at least 1 year postoperatively (mean, 3.4 ± 2.5 years).

Outcome results according to the old and new classifications are presented in Tables I and II. These results suggest that the new response score enables a more accurate evaluation of the surgical outcomes. Although 64–69% of the patients were originally classified as cure according to the old classification, only 44.7% were re-classified as cure by the strict criteria employed in the new score. Furthermore, the response score also differentiates between various degrees of clinical improvement (i.e., good, fair, or poor response). Twenty-five (26.6%) patients, most of whom were

TABLE I. Outcome Results According to Different Outcome Tools (Old Classification)

	Cure	Improvement	Failure
24-hour pad test	64%	25%	11%
24-hour diary	69%	12%	19%
Patient questionnaire	64%	28%	8%

TABLE II. Outcome Results According to the Suggested Response Score

	Total score	Patients	
		N = 94	%
Cure	0	42	44.7
Good response	1-2	25	26.6
Fair response	3-4	13	13.9
Poor response	5	7	7.4
Failure	6	7	7.4

previously classified as cure, were re-classified as good response, whereas 20 others were re-classified as fair (13.9%), or poor (7.4%) response.

Surgical outcomes were defined as failure when objective results were poor and the patient considered her surgery as failure. Seven (7.4%) patients fulfilled these criteria and were therefore classified as surgical failures. All were diagnosed pre-operatively as having complex sphincteric incontinence. Specific failure rates were therefore 11.3% for complex and 0% for simple cases.

In seven cases, there was a significant disparity between the patient questionnaire and the new classification. Seven patients who were re-classified as poor response considered themselves improved by the patient questionnaire. Another patient who was classified as fair response considered her surgery as failure by the patient questionnaire. Examination of these specific patients revealed the following data.

Case 1

A 78-year-old woman who presented with mixed stress/urge incontinence and a grade 4 cystocele. Pre-operative urodynamic evaluation revealed sphincteric incontinence and detrusor instability. Pre-operative 24-hour voiding diary documented seven episodes of stress and urge incontinence. The concomitant pad test showed 250 g of urine loss. Post-operatively, the patient no longer had stress incontinence, but had persistent urge incontinence. A post-operative diary showed seven episodes of urge incontinence and a pad test showed 35 g of urine loss. She considered herself improved by questionnaire, but was classified as poor response by the new outcome score.

Case 2

A 62-year-old woman, who previously underwent anterior colporrhaphy and Raz urethropexy, presented with severe stress incontinence. Pre-operative 24-hour voiding diary and pad test documented 14 episodes of stress incontinence and 275 g of urine loss. Post-operatively, the patient had persistent stress incontinence. Post-operative diary and pad test showed 13 episodes of stress incontinence and 67 g of urine loss. She considered herself improved by questionnaire, but was classified as poor response by the new outcome score.

Case 3

A 50-year-old woman, who previously underwent several anti-incontinence procedures, presented with mixed stress/urge incontinence. Pre-operative 24-hour

voiding diary and pad test documented 12 episodes of urinary incontinence and 201 g of urine loss. Further evaluation revealed vesicovaginal fistula as well as sphincteric incontinence. She underwent fistula repair and pubovaginal sling. Post-operatively, the patient had persistent stress and urge incontinence. Post-operative diary and pad test showed 10 episodes of stress incontinence and 166 g of urine loss. She considered herself improved by questionnaire, but was classified as poor response by the new outcome score.

Case 4

A 38-year-old woman, who previously underwent modified Pereyra, presented with mixed stress/urge incontinence. Pre-operative 24-hour voiding diary and pad test documented 11 episodes of urinary incontinence and 203 g of urine loss. Post-operatively, the patient had persistent urge incontinence. Post-operative diary and pad test showed nine episodes of urge incontinence and 98 g of urine loss. She considered herself improved by questionnaire, but was classified as poor response by the new outcome score.

Case 5

A 53-year-old woman, who previously underwent collagen injections, presented with mixed stress/urge incontinence. Pre-operative 24-hour voiding diary and pad test documented 17 episodes of urinary incontinence and 552 g of urine loss. Post-operatively, the patient had persistent stress and urge incontinence. Post-operative diary and pad test showed 11 episodes of stress incontinence and 444 g of urine loss. She considered herself improved by questionnaire, but was classified as poor response by the new outcome score.

Case 6

A 70-year-old woman, who previously underwent collagen injections, presented with mixed stress/urge incontinence. Pre-operative 24-hour voiding diary and pad test documented 10 episodes of urinary incontinence and 139 g of urine loss. Post-operatively, the patient had persistent urge incontinence. Post-operative diary and pad test showed four episodes of urge incontinence and 42 g of urine loss. She considered herself improved by questionnaire, but was classified as poor response by the new outcome score.

Case 7

A 49-year-old woman, who previously underwent several anti-incontinence procedures, presented with stress incontinence. Pre-operative 24-hour voiding diary and pad test documented three episodes of urinary incontinence and 110 g of urine loss. Further evaluation revealed a suture granuloma at the bladder neck as well as sphincteric incontinence. She underwent excision of the suture granuloma and pubovaginal sling. Post-operatively, the patient had *de novo* urge incontinence. Post-operative diary and pad test showed 3 episodes of incontinence and 50 g of urine loss. She considered herself improved by questionnaire, but was classified as poor response by the new outcome score.

Case 8

A 63-year-old woman, who previously underwent total abdominal hysterectomy, presented with stress incontinence. Pre-operative 24-hour voiding diary and pad test documented eight episodes of urinary incontinence and 90 g of urine loss. She underwent pubovaginal sling surgery. Post-operatively, the patient had *de novo* urge incontinence. Post-operative diary and pad test showed three episodes of urge incontinence and 2 g of urine loss. She considered her surgery as a failure by the questionnaire, but was classified as fair response by the new outcome score.

DISCUSSION

Outcome results of anti-incontinence procedures are usually classified and reported as cure, improvement, and failure. Theoretically, cure should imply reestablishment of normal voiding patterns. However, in most studies of surgical treatment of stress urinary incontinence, cure means that the patient no longer has stress incontinence. Yet she may have persistent, or *de novo*, urge incontinence, urinary urgency, or voiding difficulties. If urgency and detrusor instability are present pre-operatively, there is significant risk of these symptoms persisting post-operatively. If either is absent, the risk of post-operative urgency is reduced [Leach et al., 1997]. Overall, *de novo* detrusor instability was reported in a mean of 16.6% (range, 4–29%) of patients [Jarvis, 1994], whereas *de novo* urgency was reported in a median of 7% (range, 3–11%) after sling operations [Haab et al., 1996; Leach et al., 1997]. In our previously published series, 23% of the patients had post-operative persistent urge incontinence, whereas only 3% developed *de novo* urge incontinence [Chaikin et al., 1998]. Similarly, Fulford et al. [1999] noted that despite symptomatic control of stress urinary incontinence in 97% of their patients, only 80% were satisfied with the surgical results, mainly because of persistent urge syndrome. We therefore believe that no distinction should be made between different types of urinary incontinence when evaluating outcomes of any anti-incontinence procedure. Furthermore, gross classification into cure, improvement, and failure may fail to accurately reflect the real clinical status. Strict criteria for cure and failure and a detailed differentiation between various degrees of improvement, as are suggested by the new response score, may therefore provide a more meaningful tool to assess outcome results.

The response score presented in this study incorporates three popular outcome tools: 24-hour diary, 24-hour pad test, and patient questionnaire. It is quite possible that incorporation of more sophisticated tools (e.g., quality of life questionnaires, urodynamic studies) would have yielded a more accurate assessment. However, we chose the voiding diary, pad test, and patient questionnaire because they are non-invasive, inexpensive, simple tools associated with high patient compliance. Furthermore, the standardized patient questionnaire, 24-hour diary, and 24-hour pad test were all previously proven to be valid and reliable measures of urinary incontinence [Haab et al., 1997; Siltberg et al., 1997]. The suggested response score was constructed in a simple, easy-to-remember format. For the purpose of this analysis, no distinction was made between stress or urge incontinence. In addition, based on previous studies, a weight gain of up to 8 g over the 24-hour pad test was considered as normal [Lose et al., 1989; Mouritson et al., 1989; Versi et al., 1996]. Therefore, we used this cutoff point as the upper limit of a normal 24-hour pad test.

The suggested score, as was also demonstrated by the specific case studies, seems to reflect the actual surgical results more accurately. The primary disparity was found mainly between the cure rates of the old and new classifications. Approximately 64 to 69% of the patients were classified as cure by the old classification. However, objective data obtained by voiding diary and pad test demonstrated some of them still had urinary incontinence. Using the strict criteria of the new response score (i.e., no incontinence episodes, negative pad test, and subjective cure), only 44.7% of the patients were classified as cure, whereas 40.5% were classified as either good (26.6%) or fair (13.9%) response. Interestingly, previous studies demonstrated an opposite trend when patient questionnaires were compared with chart reviews or treating physician interviews. Haab et al. [1997] reported the overall (stress and urge incontinence) cure rate following pubovaginal sling, obtained by a patient questionnaire, was only 46%. Sirls et al. [1995] compared outcome results of modified Pereyra bladder neck suspension obtained by retrospective chart review with those obtained by patient questionnaire. They found a marked disparity between cure rates obtained by retrospective chart review and patient questionnaire (72 versus 47%, respectively).

Although we are pleased with the performance of each of the individual outcome tools, we believe that their incorporation into a detailed response score may more accurately reflect the actual surgical results. The suggested response score presented in this study may address this issue. Further studies are needed to assess its validity and reproducibility in other treatment modalities.

APPENDIX 1: ANTI-INCONTINENCE SURGERY RESPONSE SCORE

Postoperative 24-Hour Voiding Diary

- 0 = No urinary incontinence (urge or stress) episodes.
- 1 = 1–2 incontinence episodes.
- 2 = ≥ 3 incontinence episodes.

Postoperative 24-Hour Pad Test

- 0 = Total weight gain of the pads ≤ 8 g.
- 1 = total weight gain of the pads 9–20 g.
- 2 = total weight gain of the pads >20 g.

Patient Questionnaire

- 0 = The patient considers herself cured.
- 1 = The patient considers herself improved.
- 2 = The patient considers the operation to have failed.

Outcomes	Total score
Cure	0
Good response	1–2
Fair response	3–4
Poor response	5
Failure	6

REFERENCES

- Abrams P, Blaivas JG, Stanton SL, Anderson JT. 1988. The standardization of terminology of lower urinary tract function recommended by the International Continence Society. *Neurourol Urodynam* 7:403–26.
- Blaivas JG, Appell RA, Fantl JA, Leach G, McGuire EJ, Resnick NM, Raz S, Wein AJ. 1997. Standards of efficacy for the evaluation of treatment outcomes in urinary incontinence: Recommendations of the Urodynamics Society. *Neurourol Urodynam* 16:145–7.
- Chaikin DC, Blaivas JG, Rosenthal JE, Weiss JP. 1999. Results of pubovaginal sling for stress incontinence: A prospective comparison between four different instruments for outcome analysis. *J Urol* 162:1670–3.
- Chaikin DC, Rosenthal J, Blaivas JG. 1998. Pubovaginal fascial sling for all types of stress urinary incontinence: long-term analysis. *J Urol* 160:1312–6.
- Fulford SCV, Flynn R, Barrington J, Appanna T, Stephenson TP. 1999. An assessment of the surgical outcome and urodynamic effects of the pubovaginal sling for stress incontinence and the associated urge syndrome. *J Urol* 162:135–7.
- Haab F, Zimmern PE, Leach GE. 1996. Female stress urinary incontinence due to intrinsic sphincteric deficiency: recognition and management. *J Urol* 156:3–17.
- Haab F, Trockman BA, Zimmern PE, Leach GE. 1997. Results of pubovaginal sling for the treatment of intrinsic sphincteric deficiency determined by questionnaire analysis. *J Urol* 158:1738–41.
- Jarvis GJ. 1994. Surgery for genuine stress incontinence. *Br J Obstet Gynaecol* 101:371–4.
- Leach GE, Dmochowski RR, Appell RA, Blaivas JG, Hadley HR, Lzuber KM, Mostwin JL, O'Donnell PD, Roehrborn CG. 1997. Female stress urinary incontinence clinical guidelines panel summary report on surgical management of female stress urinary incontinence. *J Urol* 158:875–80.
- Lose G, Jorgensen L, Thunedborg P. 1989. 24-hour home pad weighing test versus 10hour ward test in the assessment of mild stress incontinence. *Acta Obstet Gynecol Scand* 68:211–5.
- Mouritsen L, Berild G, Hertz J. 1989. Comparison of different methods for quantification of urinary leakage in incontinent women. *Neurourol Urodynam* 8:579–87.
- Sirls LT, Keoleian CM, Korman HJ, Kirkemo AK. 1995. The effect of study methodology on reported success rates of the modified Pereyra bladder neck suspension. *J Urol* 154:1732–5.
- Siltberg H, Victor A, Larsson G. 1997. Pad weighing tests: the best way to quantify urine loss in patients with incontinence. *Acta Obstet Gynecol Scand Suppl* 76:28–32.
- Versi E, Orrego G, Hardy E, Seddon G, Smith P, Anand D. 1996. Evaluation of the home pad test in the investigation of female urinary incontinence. *Br J Obstet Gynaecol* 103:162–7.