ORIGINAL ARTICLE

A Randomized Trial of Urodynamic Testing before Stress-Incontinence Surgery

Charles W. Nager, M.D., Linda Brubaker, M.D., Heather J. Litman, Ph.D., Halina M. Zyczynski, M.D., R. Edward Varner, M.D., Cindy Amundsen, M.D., Larry T. Sirls, M.D., Peggy A. Norton, M.D., Amy M. Arisco, M.D., Toby C. Chai, M.D., Philippe Zimmern, M.D., Matthew D. Barber, M.D., Kimberly J. Dandreo, M.Sc., Shawn A. Menefee, M.D., Kimberly Kenton, M.D., Jerry Lowder, M.D., Holly E. Richter, Ph.D., M.D., Salil Khandwala, M.D., Ingrid Nygaard, M.D., Stephen R. Kraus, M.D., Harry W. Johnson, M.D., Gary E. Lemack, M.D., Marina Mihova, B.S., Michael E. Albo, M.D., Elizabeth Mueller, M.D., Gary Sutkin, M.D., Tracey S. Wilson, M.D., Yvonne Hsu, M.D., Thomas A. Rozanski, M.D., Leslie M. Rickey, M.D., David Rahn, M.D., Sharon Tennstedt, Ph.D., John W. Kusek, Ph.D. and E. Ann Gormley, M.D., for the Urinary Incontinence Treatment Network*

ABSTRACT

BACKGROUND

Urodynamic studies are commonly performed in women before surgery for stress urinary incontinence, but there is no good evidence that they improve outcomes.

METHODS

We performed a multicenter, randomized, noninferiority trial involving women with uncomplicated, demonstrable stress urinary incontinence to compare outcomes after preoperative office evaluation and urodynamic tests or evaluation only. The primary outcome was treatment success at 12 months, defined as a reduction in the score on the Urogenital Distress Inventory of 70% or more and a response of "much better" or "very much better" on the Patient Global Impression of Improvement. The predetermined noninferiority margin was 11 percentage points.

RESULTS

A total of 630 women were randomly assigned to undergo office evaluation with urodynamic tests or evaluation only (315 per group); the proportion in whom treatment was successful was 76.9% in the urodynamic-testing group versus 77.2% in the evaluation-only group (difference, −0.3 percentage points; 95% confidence interval, −7.5 to 6.9), which was consistent with noninferiority. There were no significant between-group differences in secondary measures of incontinence severity, quality of life, patient satisfaction, rates of positive provocative stress tests, voiding dysfunction, or adverse events. Women who underwent urodynamic tests were significantly less likely to receive a diagnosis of overactive bladder and more likely to receive a diagnosis of voiding-phase dysfunction, but these changes did not lead to significant between-group differences in treatment selection or outcomes.

CONCLUSIONS

For women with uncomplicated, demonstrable stress urinary incontinence, preoperative office evaluation alone was not inferior to evaluation with urodynamic testing for outcomes at 1 year. (Funded by the National Institute of Diabetes and Digestive and Kidney Diseases and the Eunice Kennedy Shriver National Institute of Child Health and Human Development; ClinicalTrials.gov number, NCT00803959.)

From the University of California at San Diego (C.W.N., M.E.A.) and Kaiser Permanente Southern California (S.A.M.) — both in San Diego; Loyola University, Chicago (L.B., K.K., E.M.); New England Research Institutes, Watertown, MA (H.J.L., K.J.D., M.M., S.T.); the University of Pittsburgh, Magee–Womens Hospital, Pittsburgh (H.M.Z., J.L., G.S.); the University of Alabama at Birmingham, Birmingham (R.E.V., H.E.R., T.S.W.); Duke University, Durham, NC (C.A.); William Beaumont Hospital, Royal Oak, and Oakwood Hospital, Dearborn (L.T.S, S.K.) — both in Michigan; the University of Utah, Salt Lake City (P.A.N., I.N., Y.H.); the University of Texas, San Antonio (A.M.A., S.R.K., T.A.R.); the University of Maryland, Baltimore (T.C.C., H.W.J., L.M.R.); the University of Texas Southwestern Medical Center, Dallas (P.Z., G.E.L., D.R.); Cleveland Clinic, Cleveland (M.D.B.); the National Institutes of Health, Bethesda, MD (J.W.K.); and Dartmouth–Hitchcock Medical Center, Lebanon, NH (E.A.G.). Address reprint requests to Dr. Nager at the University of California, San Diego, Perlman Suite 2A, Mail Code 0974, 9350 Campus Pt. Dr., La Jolla, CA 92037-1300, or at cnager@ucsd.edu.

*The Urinary Incontinence Treatment Network investigators are listed in the Supplementary Appendix, available at NEJM.org.

This article (10.1056/NEJMoa1113595) was published on May 2, 2012, at NEJM.org.

N Engl J Med 2012.

In the United States in 2010, approximately 260,000 women underwent surgical treatment of stress urinary incontinence.[1] Urodynamic studies, which assess physiological variables during bladder storage and emptying, are often performed preoperatively to confirm and characterize the clinical features of stress urinary incontinence or to guide decisions about modifications in treatment.[2-4] However, these studies have not been shown to improve surgical outcomes, they are uncomfortable and costly (payments allowed by Medicare are greater than $500 for the three-part study),[5] and they increase the risk of urinary tract infection.[6] A Cochrane review[7] and the National Institute for Health and Clinical Excellence in the United Kingdom[8] have recommended that randomized, controlled trials be performed to address the question of whether performing preoperative urodynamic studies improves outcomes.

We conducted a randomized trial involving women with uncomplicated, stress-predominant urinary incontinence who were planning to undergo surgery, in order to determine whether outcomes at 1 year among women who underwent only an office evaluation were inferior to those among women who also underwent preoperative urodynamic studies.

METHODS

STUDY DESIGN AND PROCEDURES

The Value of Urodynamic Evaluation (VALUE) study, an 11-center, randomized, noninferiority trial, compared the results among women who underwent an office evaluation without urodynamic testing (evaluation-only group) with those among women who underwent urodynamic testing in addition to the office evaluation (urodynamic-testing group) before their planned surgery. Details of the study design and methods have been published previously,[9] and the protocol is available with the full text of this article at NEJM.org.

Women presenting with urinary incontinence underwent a standardized basic office evaluation and were eligible for the study if they were 21 years of age or older, had a history of symptoms of stress urinary incontinence for at least 3 months, and had a score on the Medical, Epidemiological, and Social Aspects of Aging (MESA) questionnaire for stress urinary incontinence that was greater than the score on this questionnaire for urgency incontinence,[10] a postvoiding residual urine volume of less than 150 ml, a negative urinalysis or urine culture, a clinical assessment of urethral mobility, a desire for surgery for stress urinary incontinence, and a positive provocative stress test (defined as an observed transurethral loss of urine that was simultaneous with a cough or Valsalva maneuver at any bladder volume). Exclusion criteria were previous surgery for incontinence, a history of pelvic irradiation, pelvic surgery within the previous 3 months, and anterior or apical pelvic-organ prolapse of 1 cm or more distal to the hymen. Eligible patients were invited to participate in the study and asked to provide consent before any urodynamic testing was performed. After written informed consent had been obtained, study surgeons recorded their diagnoses on a comprehensive checklist of clinical diagnoses.

Patients were randomly assigned to a study group with the use of an automated randomization system stratified according to surgeon; more than 90% of the surgeons were fellowship-trained. Women in the urodynamic-testing group underwent noninstrumented uroflowmetry with a comfortably full bladder, filling cystometry with Valsalva leak-point pressures, and a pressure-flow study. Urethral pressure profilometry or urodynamic testing with the use of video was permitted if it was routinely performed as part of the preoperative investigation at the study site. Testing followed the Good Urodynamic Practice guidelines of the International Continence Society,[11] and interpretation conformed to International Continence Society nomenclature.[12] After interpretation of the urodynamic tests, study physicians again completed the same comprehensive checklist of clinical diagnoses without viewing their previous entries. At office visits 3 and 12 months after treatment, outcome data were obtained by study personnel who were unaware of the group assignments.

The protocol was approved by the institutional review board at each site, and an independent data and safety monitoring board reviewed the progress and safety of the study. The third author, the senior statistician for the study, vouches for the accuracy of the reported data and for the fidelity of the study to the protocol.

OUTCOMES

The primary outcome, treatment success, was measured by means of two validated instruments, the Urogenital Distress Inventory[13] and

10.1056/NEJMoa1113595 NEJM.org

the Patient Global Impression of Improvement.[14] We defined treatment success as a reduction in the Urogenital Distress Inventory score from baseline to 12 months of 70% or more and a Patient Global Impression of Improvement response of "very much better" or "much better" at 12 months. The Urogenital Distress Inventory is a 20-item patient-reported measure that assesses the presence of urinary incontinence, urgency, frequency, and voiding dysfunction and the extent to which the patient is bothered by these symptoms. Scores range from 0 to 300, with higher scores indicating greater distress. The 70% cutoff value was selected on the basis of the previous experience of the study investigators and receiver-operating-characteristic curve analyses from a previous surgical trial,[15] which showed that the 70% cutoff had high sensitivity (94%) and acceptable specificity (43%) for overall success in that trial. The Patient Global Impression of Improvement is a patient-reported measure of perceived improvement that is obtained by asking study participants, "How is your urinary tract condition now, as compared with how it was before you received treatment for your urinary leakage?" Responses are on a 7-point scale from "very much better" to "very much worse." This instrument correlates with the frequency of incontinence episodes, pad tests, and quality of life as it relates to incontinence.[14]

Secondary measures of incontinence outcomes were assessed by means of the Incontinence Severity Index (with scores ranging from 1 to 12 and higher scores indicating greater severity),[16] the MESA questionnaire (with scores ranging from 0 to 200 and higher scores indicating greater severity),[10] the Incontinence Impact Questionnaire (with scores ranging from 0 to 400 and higher scores indicating a more negative effect on quality of life),[13] the Medical Outcomes Study 12-Item Short Form Health Survey (with scores ranging from 0 to 200 and higher scores indicating better health),[17] the Patient Global Impression of Severity (with scores ranging from 1 [normal] to 4 [severe]),[14] and a summary score for patient satisfaction that was based on responses to questions developed for this study (with scores ranging from 0 to 100 and higher scores indicating better satisfaction).[9] At 12 months after treatment, a provocative stress test at a bladder volume of 300 ml was performed by an outcome assessor who was unaware of the study assignments.[9] Adverse events were assessed after surgery, at discharge, and at 3 and 12 months postoperatively.

PLANNED SUBGROUP ANALYSIS

We considered that the use of urodynamic studies might result in a change from surgical to nonsurgical therapy and that the patients who chose nonsurgical therapy might be less likely to meet the definition of successful treatment used for the primary outcome. To determine whether urodynamic studies might improve outcomes only among women who underwent surgery, a planned subgroup analysis was performed to compare surgical outcomes only among women in the study who underwent surgical treatment. For this analysis, the primary outcome was defined as successful surgical treatment (i.e., a score reduction of 70% or more on the Urogenital Distress Inventory and a response of "much better" or "very much better" on the Patient Global Impression of Improvement) and a negative standard-volume stress test at 12 months.[9]

STATISTICAL ANALYSIS

Assuming a significance level of 5% and a true success rate in each group of 70% with a noninferiority margin of 11 percentage points, we calculated that we needed to enroll 270 women in each study group to have 80% power for determining whether the results in the evaluation-only group were noninferior to those in the urodynamic-testing group. The investigators selected the 11% noninferiority margin on the basis of clinical judgment that this was a reasonable threshold for a trade-off between a decrease in the rate of successful treatment and the potential benefits of eliminating urodynamic studies from preoperative assessment. Assuming a 10% dropout rate, a sample of 300 women per group was required. Noninferiority was declared if the upper boundary of the 95% confidence interval for the between-group difference in the success rate was less than 11%. To minimize bias toward noninferiority, only women who were treated per protocol (i.e., who underwent the randomly assigned evaluation) were considered in the primary outcome analysis.

We also performed an intention-to-treat analysis that included all women who underwent randomization, but this was considered a secondary analysis. Other secondary outcomes were assessed in the intention-to-treat population.

Descriptive statistics were computed; nonpara-

metric statistics were presented for nonnormally distributed variables. Wilcoxon rank-sum tests and t-tests were used for the comparison of continuous variables; chi-square tests and Fisher's exact tests were used to compare categorical variables, as appropriate. Linear regression and logistic-regression models were fit to assess whether outcomes differed by treatment group with adjustment for unbalanced baseline variables. For measures collected at two time points, paired t-tests and McNemar's tests were used, as appropriate. Sensitivity analyses were performed by classifying missing data for primary outcome measures as all treatment successes and as all treatment failures in order to examine the consistency of our findings. Analyses were performed with the use of SAS statistical software, version 9.2 (SAS Institute).

RESULTS

STUDY POPULATION AND GROUP ASSIGNMENTS

Between November 2008 and June 2010, a total of 630 women underwent randomization (315 in each group) at 11 participating sites (see the Supplementary Appendix, available at NEJM.org). Of the 53 participating surgeons, 38 were urogynecologists and 15 were urologists; more than 90% were fellowship-trained. In the per-protocol analysis, primary outcome data were available for 264 women in the urodynamic-testing group and 259 in the evaluation-only group (Fig. 1).

The provocative stress test, which had to be positive for inclusion in the study, was performed at a median volume of 200 ml (25th percentile, 68 ml; 75th percentile, 300 ml). Demographic and clinical characteristics at baseline for women with primary outcome data available were generally similar between the two groups (Table 1, and Table 1 in the Supplementary Appendix), although the data differed modestly with respect to six of the measures assessed; specifically, women assigned to the urodynamic-testing group had a longer duration of incontinence and were more likely to smoke, to not be taking estrogen-replacement therapy, to have received nonsurgical treatment for urinary incontinence, to have urethral mobility, and to have a higher score (indicating greater severity) on the Incontinence Severity Index. Table 2 in the Supplementary Appendix describes the baseline characteristics of all patients who underwent randomization. Table 3 in the Supplementary Appendix shows the urodynamic characteristics for the urodynamic-testing group; 97% of women in this group had stress incontinence confirmed by urodynamic testing.

Figure 1 (facing page). Study Enrollment.
The major reasons for not meeting the inclusion criteria were not meeting the definition of stress-predominant urinary incontinence (1032 women), having prolapse (639), or having a history of surgery for incontinence or other conditions (528). Reasons for declining to participate were related to study procedures (e.g., too invasive or requiring too much time; 164 women), randomization (210), or insurance coverage (5). Administrative reasons included a delay in starting the study (e.g., because of delayed approval by the institutional review board; 114 women), no staff members at the study site who were certified in the patient's native language (43), and other reasons (156). Of the 40 women who provided informed consent but were deemed ineligible before randomization occurred, 10 withdrew consent (3 were not available to start treatment within 6 weeks after randomization, 1 did not have data for the provocative stress test, and 6 had other reasons), 11 had data from urodynamic studies reviewed in the previous 12 months, 7 did not have data for the provocative stress test, 5 were not available to start treatment within 6 weeks after randomization, and 7 had other reasons. Of the 11 women who withdrew consent after random assignment to urodynamic testing, 5 did not complete the testing and were not included in the per-protocol analysis. Of the 27 women lost to follow-up, 1 did not have stress urinary incontinence (according to the MESA questionnaire) and 2 did not complete urodynamic testing; these women were not included in the per-protocol analysis.

PRIMARY OUTCOME

The rate of treatment success was 76.9% (203 of 264 women) in the urodynamic-testing group as compared with 77.2% (200 of 259) in the evaluation-only group (Fig. 2). The between-group difference of −0.3 percentage points (95% confidence interval, −7.5 to 6.9) met our predetermined criterion for the noninferiority of office evaluation alone. Results were similar in the intention-to-treat population (Fig. 2). In a sensitivity analysis, with missing data classified as either all successes or all failures, the results were similar (Table 4 in the Supplementary Appendix). A post hoc analysis with adjustment for baseline differ-

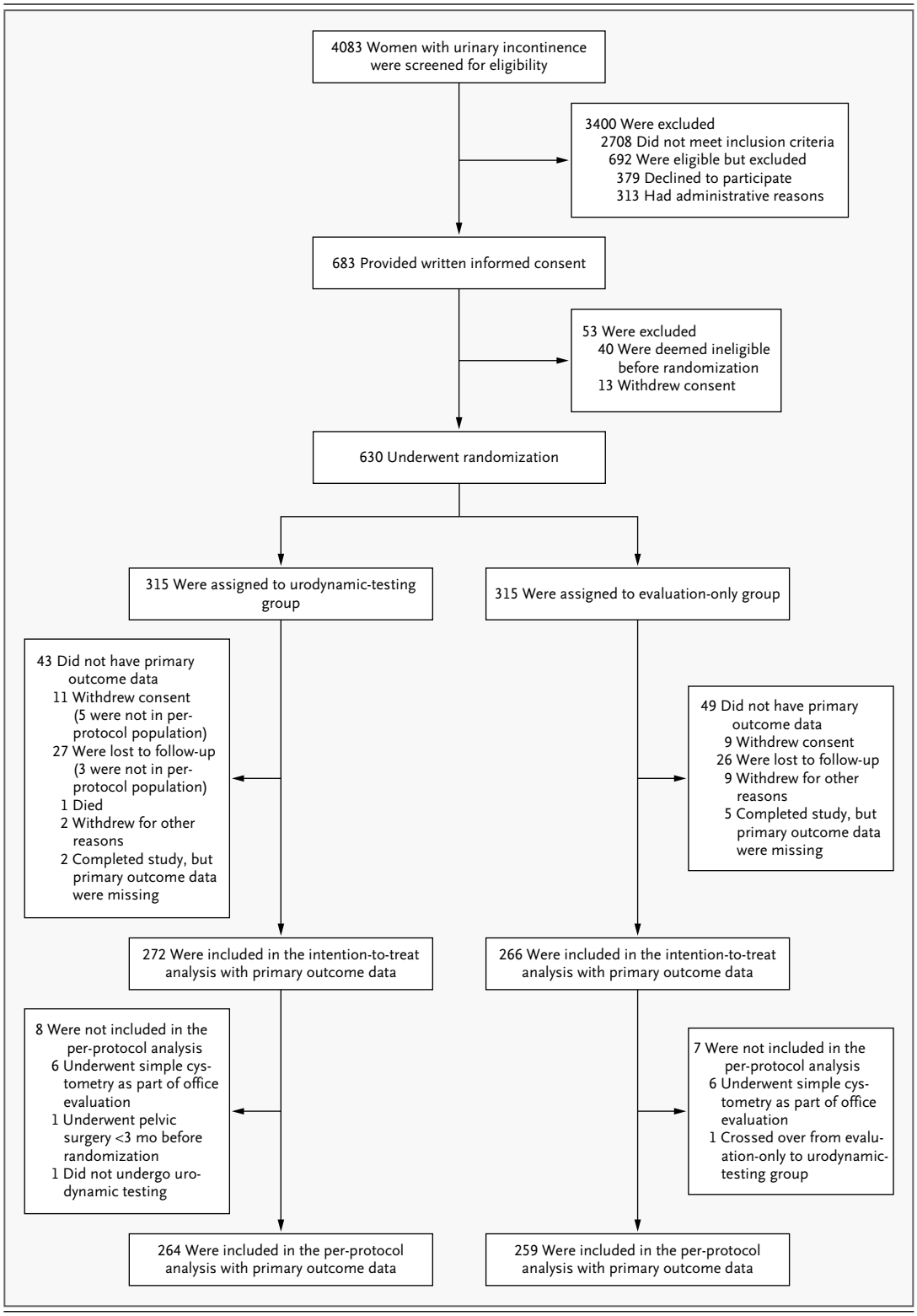

Table 1. Baseline Demographic and Clinical Characteristics of the Patients with Primary Outcome Data in the Per-Protocol Analysis.*

Characteristic	Urodynamic Testing (N=264)	Office Evaluation Only (N=259)	P Value
Age — yr	51.9±10.4	51.6±10.0	0.74
Body-mass index†	29.1±5.7	28.9±6.1	0.70
Race or ethnic group — no. (%)‡			0.43
Non-Hispanic white	209 (79.2)	191 (73.7)	
Other	55 (20.8)	68 (26.3)	
Parous — no. (%)	252 (95.5)	246 (95.0)	0.96
Duration of incontinence — mo	107.4±100.3	90.7±79.9	0.04
Postmenopausal — no./total no. (%)§	119/264 (45.1)	120/257 (46.7)	0.22
Current estrogen-replacement therapy — no./total no. (%)¶	44/166 (26.5)	57/152 (37.5)	0.04
Current smoking — no. (%)	35 (13.3)	18 (6.9)	0.04
History of nonsurgical treatment for urinary incontinence — no. (%)	174 (65.9)	148 (57.1)	0.04
History of pelvic surgery — no. (%)‖	179 (67.8)	192 (74.1)	0.11
Urethral mobility — no. (%)**	248 (93.9)	228 (88.0)	0.02
Postvoiding residual urine volume — ml			
Median	10	18	0.15
Interquartile range	5–30	5–35	
Urogenital Distress Inventory score††	125.8±44.3	121.6±43.5	0.27
Incontinence Severity Index score‡‡	7.9±2.6	7.4±2.8	0.04
MESA score§§			
Stress incontinence	73.6±16.3	71.4±19.0	0.17
Urgency incontinence	31.7±21.0	32.4±22.3	0.72
Incontinence Impact Questionnaire score¶¶	42.1±22.6	42.2±22.4	0.94
SF-12 score‖‖	98.0±14.5	96.4±13.7	0.19
Score of moderate or severe on the Patient Global Impression of Severity — no./total no. (%)***	225/262 (85.9)	227/259 (87.6)	0.44
Score of 0 on the Charlson comorbidity index — no. (%)†††	183 (69.3)	189 (73.0)	0.22

* Plus–minus values are means ±SD.
† Data on body-mass index (the weight in kilograms divided by the square of the height in meters) were available for 261 women in the urodynamic-testing group and 251 in the evaluation-only group.
‡ Race or ethnic group was self-reported.
§ Data on postmenopausal status were available for all women in the urodynamic-testing group and 257 in the evaluation-only group.
¶ Current use of estrogen-replacement therapy was recorded for 166 perimenopausal and postmenopausal women in the urodynamic-testing group and for 152 in the evaluation-only group.
‖ History of pelvic surgery included cesarean section and hysterectomy.
** Urethral mobility was defined by the study physician and could be assessed by Q-tip test,[18] point Aa on a Pelvic Organ Prolapse Quantification system[19] examination, visualization, palpation, or lateral cystogram.
†† Scores on the Urogenital Distress Inventory[13] range from 0 to 300, with higher scores indicating greater distress.
‡‡ Scores on the Incontinence Severity Index[16] range from 1 to 12, with higher scores indicating greater severity. Scores were available for 263 women in the urodynamic-testing group and 257 in the evaluation-only group.
§§ Scores on the stress and urgency indexes of the Medical, Epidemiological, and Social Aspects of Aging[10] (MESA) questionnaire range from 0 to 100, with higher scores indicating greater severity.
¶¶ Scores on the Incontinence Impact Questionnaire[13] range from 0 to 400, with higher scores indicating a more negative effect on quality of life.
‖‖ Scores on the Medical Outcomes Study 12-Item Short-Form Health Survey (SF-12)[17] range from 0 to 200, with higher scores indicating better health.
*** Scores on the Patient Global Impression of Severity[14] range from 1 (normal) to 4 (severe). Scores were available for 262 women in the urodynamic-testing group and all in the evaluation-only group.
††† Scores on the Charlson comorbidity index range from 0 to 30, with 0 indicating no coexisting conditions.[20]

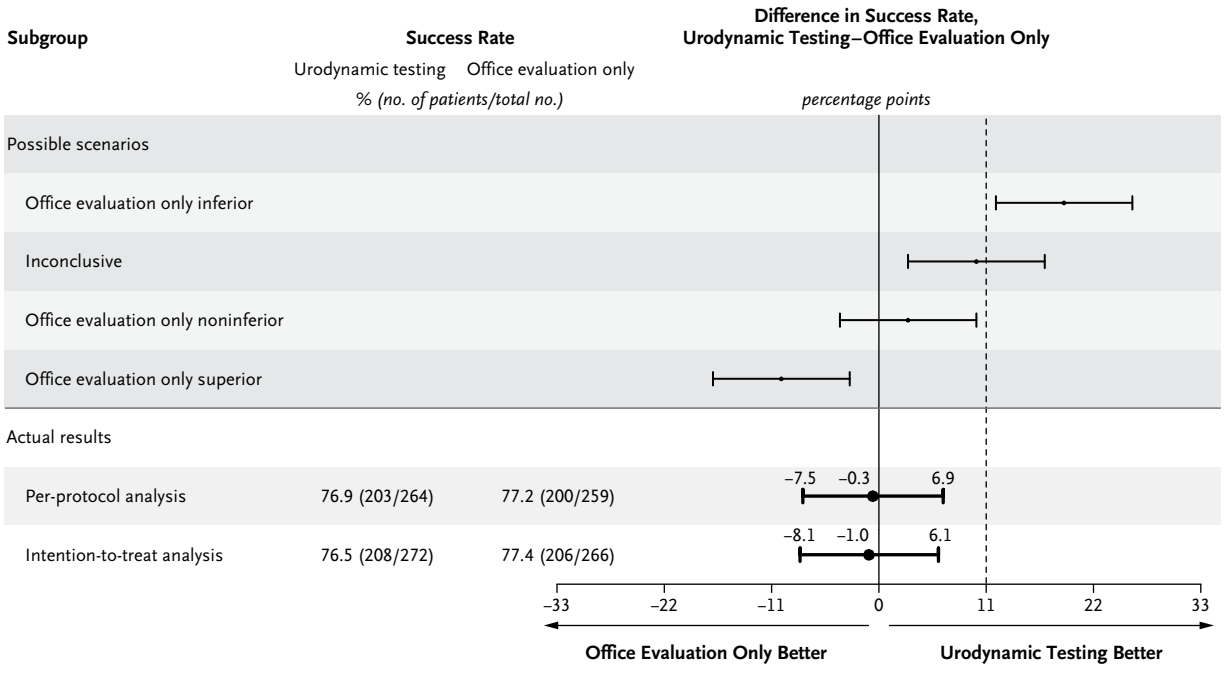

Figure 2. Primary Outcome Results.

Success was defined as a reduction of at least 70% in the Urogenital Distress Inventory score from baseline to 12 months and a response of "very much better" or "much better" on the Patient Global Impression of Improvement measure at 12 months. The horizontal I bars indicate 95% confidence intervals. The dashed vertical line denotes the predetermined noninferiority margin of 11 percentage points.

ences in duration of incontinence, Incontinence Severity Index score, and status with respect to smoking, history of nonsurgical treatment for urinary incontinence, current use of hormone-replacement therapy, and urethral mobility did not materially alter the findings (Table 5 in the Supplementary Appendix).

No significant differences were found between the urodynamic-testing and evaluation-only groups for several secondary outcome measures (Table 2). Specifically, the groups had similar changes in scores on the Incontinence Severity Index, the Patient Global Impression of Severity, and the condition-specific and global quality-of-life measures, as well as similar rates of positive provocative stress tests and similar levels of patient satisfaction at 12 months. There were no significant differences in the percentages of patients with any adverse event between those assigned to the urodynamic-testing group and those assigned to the evaluation-only group (21.3% and 19.4%, respectively; P=0.55) (Table 7 in the Supplementary Appendix)

SUCCESS OF SURGICAL TREATMENT

In the urodynamic-testing group, 17 women (5.4%) did not undergo surgery (10 canceled the surgery, 4 changed to nonsurgical treatment, and 3 had medical contraindications), and in the evaluation-only group, 27 women (8.6%) did not undergo surgery (21 canceled, 4 had medical contraindications, 1 was lost to follow-up, and 1 underwent surgery after the 12-month visit) (P=0.12). Of the 586 women who underwent surgery, 443 had follow-up stress-test data available (70.5% in the urodynamic-testing group and 70.2% in the evaluation-only group). Surgical treatment was successful (with the definition of success expanded to include a negative stress test at a bladder volume of 300 ml at 12 months) in 154 of 222 women (69.4%) in the urodynamic-testing group and in 161 of 221 (72.9%) in the evaluation-only group (P=0.42).

CLINICAL DIAGNOSIS AND TREATMENT

After the office evaluation, there were no significant differences in clinical diagnoses between the

Table 2. Outcomes.*

Outcome	Urodynamic Testing (N=272)	Office Evaluation Only (N=266)	P Value
Primary			
70% reduction in Urogenital Distress Inventory score — no. (%)	210 (77.2)	210 (78.9)	0.63
"Very much better" or "much better" on Patient Global Impression of Improvement — no./total no. (%)†	248/270 (91.9)	238/262 (90.8)	0.68
Secondary			
Change in Urogenital Distress Inventory score	−100.2±50.1	−98.4±51.4	0.68
Change in Incontinence Severity Index score‡	−6.0±3.3	−5.7±3.4	0.40
Change in MESA score			
Stress incontinence	−61.5±22.0	−60.2±24.7	0.50
Urgency incontinence	−19.7±21.4	−22.2±22.4	0.19
Change in Incontinence Impact Questionnaire score	−35.9±23.2	−37.3±23.7	0.49
Change in SF-12 score§	5.0±10.8	7.3±12.0	0.02
Change in Patient Global Impression of Severity score¶	−1.8±0.9	−1.8±0.9	0.68
Score of moderate or severe on the Patient Global Impression of Severity at 12 mo — no./total no. (%)‖	19/271 (7.0)	15/266 (5.6)	0.51
Overall patient satisfaction score at 12 mo**	79.5±30.4	82.2±28.6	0.28
Positive provocative stress test at 12 mo — no./total no. (%)††	36/225 (16.0)	26/222 (11.7)	0.19

* Plus–minus values are means ±SD. Change was calculated as the score at 12 months minus the score at baseline. For all change scores except the SF-12 score, higher scores indicate worse function, so the larger the negative value, the greater the improvement; for the SF-12 score, the larger the positive value, the greater the improvement.
† Data on the Patient Global Impression of Improvement were missing for 2 women in the urodynamic-testing group and 4 in the evaluation-only group.
‡ Data on the change in the Incontinence Severity Index score were missing for 2 women in the urodynamic-testing group and 3 in the evaluation-only group.
§ Data on the change in the SF-12 score were missing for 5 women in the urodynamic-testing group and 5 in the evaluation-only group.
¶ Data on the change in the Patient Global Impression of Severity score were missing for 4 women in the urodynamic-testing group.
‖ The Patient Global Impression of Severity score at 12 months was missing for 1 woman in the urodynamic-testing group.
** The overall patient satisfaction score at 12 months was missing for 4 women in the urodynamic-testing group and 7 in the evaluation-only group.
†† Data on the provocative stress test at 12 months were missing for 47 women in the urodynamic-testing group and 44 in the evaluation-only group.

urodynamic-testing and evaluation-only groups (Table 3). After urodynamic testing, women in the urodynamic-testing group were significantly less likely to receive a diagnosis of an overactive bladder with incontinence, an overactive bladder without incontinence, or suspected intrinsic sphincter deficiency, and they were significantly more likely to receive a diagnosis of voiding-phase dysfunction than were those in the evaluation-only group (Table 3). This change in preoperative diagnosis after urodynamic testing did not result in different distributions of overall surgical treatments. The surgical treatments that were performed in the urodynamic-testing and evaluation-only groups, respectively, were as follows: retropubic midurethral sling in 64.7% and 64.6%, transobturator midurethral sling in 29.0% and 28.1%, mini-sling in 2.0% and 1.4%, traditional sling in 3.4% and 4.9%, retropubic urethropexy in 0.0% and 0.7%, and urethral-bulking injection in 1.0% and 0.4% (Table 8 in the Supplementary Appendix).

Approximately 93% of women in both study groups underwent midurethral-sling surgery. After urodynamic testing, 12 patients for whom a retropubic midurethral sling was planned received a transobturator midurethral sling and 6 for whom a transobturator midurethral sling

10.1056/NEJMoa1113595 NEJM.ORG

Table 3. Clinical Diagnosis.*

Diagnosis	After Office Evaluation		P Value	After Urodynamic Testing (N=294)†	P Value‡
	Urodynamic-Testing Group (N=315)	Evaluation-Only Group (N=315)			
Stress urinary incontinence — no. (%)	315 (100)	315 (100)	>0.99	292 (99.3)	NA
Overactive bladder with incontinence — no. (%)	131 (41.6)	108 (34.3)	0.06	74 (25.2)	<0.001
Overactive bladder without incontinence — no. (%)	99 (31.4)	94 (29.8)	0.67	61 (20.7)	0.002
Voiding dysfunction — no. (%)	7 (2.2)	9 (2.9)	0.61	35 (11.9)	<0.001
Suspected intrinsic sphincter deficiency — no./total no. (%)§	61/314 (19.4)	54/315 (17.1)	0.46	37/294 (12.6)	0.003

* Physicians completed a checklist of five possible diagnoses, which were not considered to be mutually exclusive. NA denotes not applicable.
† Data on diagnoses are for the urodynamic-testing group only and were missing for 21 women in this group.
‡ P values are for the comparison of diagnoses before and after urodynamic testing.
§ Data on the diagnosis of suspected intrinsic sphincter deficiency were missing for 1 woman in the urodynamic-testing group.

was planned received a retropubic midurethral sling. There were no other major changes in decision making.

DISCUSSION

In this randomized trial involving women with uncomplicated, primary stress-predominant urinary incontinence (as confirmed by stress urinary leakage during an office evaluation), the rate of successful treatment at 12 months among women who underwent office evaluation only was noninferior to the rate among those who underwent urodynamic testing in addition to office evaluation. Our findings suggest that for women with uncomplicated stress urinary incontinence, a basic office evaluation as described in this report (i.e., a positive result on a provocative stress test, a normal postvoiding residual volume, an assessment of urethral mobility, and confirmation of the absence of bladder infection) is a sufficient preoperative workup.

Although some professional organizations recommend routine urodynamic testing before surgery for stress urinary incontinence,[21,22] the National Institute for Health and Clinical Excellence advised against routine urodynamic testing before surgery "in women with a clearly defined clinical diagnosis of pure stress urinary incontinence."[8] This recommendation has been criticized by authors who note that only 5% of their patients with urinary incontinence at a tertiary care center in the United Kingdom had an isolated diagnosis of "pure stress urinary incontinence."[2] Our eligibility criteria more broadly reflect the characteristics of women with stress incontinence who are seen in clinical practice. We included women who had symptoms of both stress and urgency urinary incontinence, as long as the stress symptoms were predominant. However, this study did not address the role of urodynamic testing in patients with more challenging issues, such as urge-predominant incontinence, previous surgery for incontinence, neurologic disease, or planned concomitant surgery for pelvic-organ prolapse.

Our finding that preoperative urodynamic tests failed to improve the rate of treatment success, as compared with the success rate associated with a basic office evaluation, may be explained by several factors. First, the diagnosis of stress incontinence, as made by office evaluation, was confirmed urodynamically in 97% of the women in the urodynamic-testing group. Second, the factors identified on preoperative urodynamic testing that traditionally have been considered to increase the risk of a poor outcome after surgery for stress incontinence (e.g., intrinsic sphincter deficiency, detrusor overactivity, and voiding dysfunction) may not be predictive of a poor outcome.[23] A randomized clinical trial of the Burch colposuspension versus the autologous fascial sling showed that typical urodynamic measures did not predict the likelihood of successful treatment of stress incontinence[24] or the risk of postoperative voiding dysfunction.[25]

Even if some urodynamic measures predict worse outcomes in some patients than in others, they may not be indications for a change in surgical management. In a recent randomized trial of retropubic and transobturator midurethral-sling surgery, low Valsalva leak-point pressures and low maximum urethral-closure pressures were associated with an odds of surgical failure that was increased by a factor of 2, but these associations were observed in both surgical groups and therefore did not suggest changes in surgical management.[26] In our study, approximately 93% of the participants in both groups received a transobturator or retropubic midurethral sling, and midurethral slings are routinely used in patients with either stress incontinence or both stress and urge (mixed) incontinence. Whereas the urodynamic findings caused physicians in many cases to change their clinical diagnosis (e.g., fewer diagnoses of overactive bladder and more diagnoses of voiding dysfunction), these changes in diagnoses were not associated with overall changes in surgical management or surgical outcome, and we therefore question the clinical importance of such diagnostic changes.

Unlike many prior studies that evaluated surgery for stress incontinence, in which cure of stress incontinence was used as the primary outcome measure, we chose a broader measure of lower urinary tract function to capture any potential benefit of urodynamic tests. It is possible that urodynamic tests could improve global outcomes by altering diagnosis and treatment in a manner that reduces the risk of postoperative adverse effects such as urinary urgency, urinary frequency, urgency incontinence, or voiding dysfunction. For this reason, we chose for our primary outcome measures the scores on the Urogenital Distress Inventory and the Patient Global Impression of Improvement — measures that broadly assess bladder storage and emptying function — and took into account the patient's own assessment of her urinary condition. However, even when we assessed efficacy with the use of more specific measures of incontinence severity, there was no demonstrable benefit of preoperative urodynamic studies over standard office evaluation.

The strengths of this study include the clearly defined, large study population, the fact that the outcome evaluators were unaware of the study assignments, and the generalizability that was afforded by the participation of 11 centers and 53 surgeons. This generalizability should be qualified by the fact that more than 90% of our surgeons were fellowship-trained and therefore may be more experienced at clinical evaluation than others with less training. Some modest between-group differences at baseline may have contributed to an imbalance that favored the noninferiority conclusion, but statistical adjustment for these differences did not materially alter our results.

In conclusion, with respect to success of treatment at 1 year, this study showed that a basic office assessment for women with uncomplicated stress-predominant urinary incontinence who have stress incontinence on office evaluation is noninferior to a preoperative evaluation that also includes urodynamic testing. These results argue against routine preoperative urodynamic testing in patients with uncomplicated stress urinary incontinence.

Supported by cooperative agreements (U01 DK58225, U01 DK58229, U01 DK58234, U01 DK58231, U01 DK60379, U01 DK60380, U01 DK60393, U01 DK60395, U01 DK60397, and U01DK60401) from the National Institute of Diabetes and Digestive and Kidney Diseases and by the Eunice Kennedy Shriver National Institute of Child Health and Human Development.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

REFERENCES

1. Urogynecologic surgical mesh: update on the safety and effectiveness of transvaginal placement for pelvic organ prolapse. Silver Spring, MD: Food and Drug Administration, July 2011.
2. Agur W, Housami F, Drake M, Abrams P. Could the National Institute for Health and Clinical Excellence guidelines on urodynamics in urinary incontinence put some women at risk of a bad outcome from stress incontinence surgery? BJU Int 2009;103:635-9.
3. Digesu GA, Hendricken C, Fernando R, Khullar V. Do women with pure stress urinary incontinence need urodynamics? Urology 2009;74:278-81.
4. Schierlitz L, Dwyer PL, Rosamilia A, et al. Effectiveness of tension-free vaginal tape compared with transobturator tape in women with stress urinary incontinence and intrinsic sphincter deficiency: a randomized controlled trial. Obstet Gynecol 2008;112:1253-61.
5. Ingenix. National Fee Analyzer 2011: charge data for evaluating fees nationally. Eden Prairie, MN: Ingenix, 2011.
6. Okorocha I, Cumming G, Gould I. Female urodynamics and lower urinary tract infection. BJU Int 2002;89:863-7.
7. Glazener CM, Lapitan MC. Urodynamic investigations for management of urinary incontinence in adults. Cochrane Database Syst Rev 2002;3:CD003195.
8. Urinary incontinence: the management of urinary incontinence in women. NICE clinical guideline 40. London: Na-

tional Institute for Health and Clinical Excellence, 2006:1-36.
9. Nager CW, Brubaker L, Daneshgari F, et al. Design of the Value of Urodynamic Evaluation (ValUE) trial: a non-inferiority randomized trial of preoperative urodynamic investigations. Contemp Clin Trials 2009;30:531-9.
10. Herzog AR, Diokno AC, Brown MB, Normolle DP, Brock BM. Two-year incidence, remission, and change patterns of urinary incontinence in noninstitutionalized older adults. J Gerontol 1990;45:M67-M74.
11. Schäfer W, Abrams P, Liao L, et al. Good urodynamic practices: uroflowmetry, filling cystometry, and pressure-flow studies. Neurourol Urodyn 2002;21:261-74.
12. Abrams P, Cardozo L, Fall M, et al. The standardisation of terminology in lower urinary tract function: report from the standardisation sub-committee of the International Continence Society. Urology 2003;61:37-49.
13. Shumaker SA, Wyman JF, Uebersax JS, McClish D, Fantl JA. Health-related quality of life measures for women with urinary incontinence: the Incontinence Impact Questionnaire and the Urogenital Distress Inventory. Qual Life Res 1994;3:291-306.
14. Yalcin I, Bump RC. Validation of two Global Impression questionnaires for incontinence. Am J Obstet Gynecol 2003;189:98-101.
15. Albo ME, Richter HE, Brubaker L, et al. Burch colposuspension versus fascial sling to reduce urinary stress incontinence. N Engl J Med 2007;356:2143-55.
16. Sandvik H, Hunskaar S, Seim A, Hermstad R, Vanvik A, Bratt H. Validation of a severity index in female urinary incontinence and its implementation in an epidemiological survey. J Epidemiol Community Health 1993;47:497-9.
17. Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996;34:220-33.
18. Crystle CD, Charme LS, Copeland WE. Q-tip test in stress urinary incontinence. Obstet Gynecol 1971;38:313-5.
19. Bump RC, Mattiasson A, Bø K, et al. The standardization of terminology of female pelvic organ prolapse and pelvic floor dysfunction. Am J Obstet Gynecol 1996;175:10-7.
20. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373-83.
21. Abrams P, Andersson KE, Birder L, et al. Fourth International Consultation on Incontinence: recommendations of the International Scientific Committee: evaluation and treatment of urinary incontinence, pelvic organ prolapse, and fecal incontinence. Neurourol Urodyn 2010;29:213-40.
22. Ghoniem G, Stanford E, Kenton K, et al. Evaluation and outcome measures in the treatment of female urinary stress incontinence: International Urogynecological Association (IUGA) guidelines for research and clinical practice. Int Urogynecol J Pelvic Floor Dysfunct 2008;19:5-33.
23. Fletcher SG, Lemack GE. Clarifying the role of urodynamics in the preoperative evaluation of stress urinary incontinence. ScientificWorldJournal 2008;8:1259-68.
24. Nager CW, FitzGerald M, Kraus SR, et al. Urodynamic measures do not predict stress continence outcomes after surgery for stress urinary incontinence in selected women. J Urol 2008;179:1470-4.
25. Lemack GE, Krauss S, Litman H, et al. Normal preoperative urodynamic testing does not predict voiding dysfunction after Burch colposuspension versus pubovaginal sling. J Urol 2008;180:2076-80.
26. Nager CW, Sirls L, Litman HJ, et al. Baseline urodynamic predictors of treatment failure 1 year after mid urethral sling surgery. J Urol 2011;186:597-603.